



1976

The Effect of Male-female Stimulus Variables on the Reliability of the Motivation Index of the Thematic Apperception Test

Roger S. Arnold
Loyola University Chicago

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

 Part of the [Psychology Commons](#)

Recommended Citation

Arnold, Roger S., "The Effect of Male-female Stimulus Variables on the Reliability of the Motivation Index of the Thematic Apperception Test" (1976). *Dissertations*. 1603.
https://ecommons.luc.edu/luc_diss/1603

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



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).
Copyright © 1976 Roger S. Arnold

THE EFFECT OF MALE-FEMALE STIMULUS VARIABLES ON THE
RELIABILITY OF THE MOTIVATION INDEX OF THE
THEMATIC APPERCEPTION TEST

by

Roger S. Arnold

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

June

1976

ACKNOWLEDGEMENTS

I would like to express my thanks to Dr. Ann Heilman, the director of this dissertation, to Dr. Leroy Wauck and Dr. Jeanne Foley, members of the committee for their suggestions, criticisms and encouragement in this work. I am very grateful to Dr. Bede Smith for his help in analyzing and scoring the data. In addition, I would like to thank Mrs. Danielle Uhlmansiek who graciously offered her skills to type the manuscript; to the students of Loyola University who volunteered to be subjects for the study; to Dr. Magda Arnold, Dr. and Mrs. Donald Tyrell for their counsel and encouragement; to Dr. Emil Posavac, my academic advisor; to Mr. and Mrs. John Dickmann for their interest and stimulation; and to the members of the Society of the Divine Word, particularly Fr. Robert Mallonee, Fr. John Fincutter, and Fr. John Donaghey, with whose help the time and resources were offered to make the study possible in the first place.

LIFE

Roger S. Arnold was born on June 28, 1933, in Pittsburgh, Pennsylvania.

He received a Bachelor of Arts degree in Philosophy from Divine Word College, Epworth, Iowa in June, 1957, a Bachelor of Theology from Divine Word Theologate, Techny, Illinois in June, 1961, a Master of Arts degree in Canon Law from the Gregorian University, Rome, Italy, in June, 1963, and a Master of Arts degree in Psychology from Loyola University, Chicago, Illinois.

He taught Canon Law, Moral and Pastoral Theology at Divine Word Theologate, Techny, Illinois from 1963 to 1969. He has also taught Psychology at Loyola University from 1974 to present.

He entered the doctoral program in Psychology of Personality at Loyola University in 1974 and expects to complete it in June, 1976.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
LIFE	iii
LIST OF TABLES	v
CONTENTS OF APPENDICES	vi
 Chapter	
1. INTRODUCTION	1
II. REVIEW OF RELATED LITERATURE	4
III. METHOD	41
IV. RESULTS	45
V. DISCUSSION	50
SUMMARY	64
REFERENCES	71
APPENDIX A	87
APPENDIX B	88

LIST OF TABLES

Table	Page
1. Ranges and Median Motivation Index (MI) Scores	47
2. Ranges and Median Motivation Index (MI) Scores from Combined Data for Male and Female Subjects	48
3. Nonparametric Spearman Rank Correlation Coefficients for Motivation Index Scores	49
4. Relationship of Correlation Coefficients Derived from Comparison of the Male Cards and Female Cards (Rho 1) and the Correlation Coefficients Derived from the Comparison of the First Five Male and Female Cards with the Second Five Male and Female Cards (Rho 2)	58

CONTENTS OF APPENDICES

	Page
APPENDIX A Relationship of Horner's Criteria for Fear-of-Success Stories and the Scoring Manual of Story Sequence Analysis	87
APPENDIX B Descriptions of the Individual Male and Female Stimulus Cards used in this Study	88

CHAPTER I

INTRODUCTION

The phenomenon of story telling is probably as old as the human race. Stories have been told for many different purposes: to entertain, to teach, to encourage, to persuade, to confront, to challenge, to evoke feelings.

Stories have also been studied from various viewpoints: content, purpose, structure, language, style. The province of these investigations is usually that of the art critic or literary scholar.

Given the fact that telling stories is a common human activity, it is not surprising that they have come under the scrutiny of the psychologist. Since stories are the product of human psychological activities (perception, memory, imagination, emotion, insight, judgment decision) they offer abundant material for psychological investigation. Naturally, each one of the psychological functions mentioned above can be studied in its own right, at least to some extent. In telling or writing a story, however, the individual must exercise these functions in a unified way to produce a single structured result. And despite the controversies that have existed in this area it is generally accepted that the way an individual organizes psychological activities to produce and deal with various situations is the field of personality investigation.

From a literary viewpoint, interest is focused on the personalities of the characters in the story. By contrast, the psychologist tries to

find out what the story reveals about the personality of the storyteller. Although there are many different opinions as to how this should be done, the inference itself is generally considered valid.

To investigate stories scientifically a method is required which incorporates the following elements: (a) a systematic, standardized procedure for eliciting an adequate sample of stories, (b) a way of extracting from the stories the data relevant to personality structure and function, and (c) a system for measuring the results of the analysis. Even after the method, instrument, technique or test has been set up there still remains the task of establishing its validity and reliability. Does it measure the personality characteristics it purports to measure and does it produce consistent results?

The most common procedure for eliciting stories has been to present the subject or client with a series of pictures and ask the individual to make up a story about each one. The stories can then be analyzed and scored according to specified criteria. Unfortunately, not only are there a variety of methods with widely divergent principles of analysis and scoring, there are also recurrent difficulties with regard to validity and reliability.

The main goal of the present study is to add to the research on the reliability of the method of Thematic Apperception Test analysis called Story Sequence Analysis (Arnold, 1962) with particular regard to sex differences. In addition, the following issues with regard to sex differences will be investigated: thematic preoccupation, the effect of the stimulus cards on the production of scorable stories, and motivational level (i.e., positive/negative motivational dispositions). Lastly, the

protocols will be examined for evidence of Horner's (1968, 1970) hypothesis that women show more "fear of success" than men.

CHAPTER II

REVIEW OF THE RELATED LITERATURE

The literature regarding the TAT is voluminous. Fortunately, there are many reviews, manuals, and edited symposia which survey the field from a clinical as well as research and methodological point of view or which describe specific scoring systems.

Basic bibliographies on all aspects of the test are contained in the Third, Fourth, Fifth, and Sixth Mental Measurements Yearbooks edited by Buros (1949, 1953, 1959, 1965).

The complete reviews (Bell, 1948; Mayman, 1946; Symonds, 1949; Tomkins, 1947) are somewhat dated but nonetheless valuable. More recent publications containing review material either have been selective and incomplete (Bellak, 1950; Bellak & Ort, 1952; Eron, 1959; Jensen, 1959; Watson, 1951; Wyatt & Veroff, 1956) or have dealt with special topics (Kagan, 1960; Murstein, 1959; Murstein, 1961).

The review in this study will cover only those aspects of the TAT which are relevant to the topic. These issues are the following: general background in the genesis of the TAT; administration; stimulus conditions; sex differences; methods of scoring and interpretation; reliability; and validity.

Historical Antecedants

Using the responses of individuals to pictures as a way of studying psychological phenomena has a number of forerunners in the history of

psychology.

Binet and Simon (1905) found that they could distinguish three successive stages in the verbal responses of children to pictures. At age 3, there was a mere enumeration of objects. By age 7 a child could describe the picture rather adequately in terms of its more obvious qualities. As he grew older he could describe the action, but not until the age of 15 was the child really able to comprehend and interpret fully what was happening. Thus Binet was able to use these developmental differences in deriving his age scale of intelligence.

Early in the century, Brittain (1907) and Libby (1908) employed the storytelling technique for studying imagination in adolescents. These pioneer studies, which all dealt with such general facets of mental life as intelligence, apperception, and imagination, were not concerned with the picture-story method as an index to the individual personality.

A more direct precursor to present day methods was the set of pictures Schwartz (1932) used as an adjunct in the psychiatric interviewing of delinquents. This consisted of eight pictures representing situations most frequently encountered in the histories of delinquents. The subject was first asked to describe what he saw, then to tell what the boy in the picture was thinking. On the basis of the responses, the examiner then questioned the subject further, following up every lead suggested by the child's answers. The child was then asked to tell what he would think and do if he were the boy in the picture. The test was never widely used by psychologists probably because of its circumscribed area of application and the difficulty of standardizing the procedure due to the complex follow-up questions which had to be asked.

One additional early test should be noted, i.e. the Four-Picture Test developed by van Lennep in 1930 (1948, 1951). Four vaguely drawn colored pictures of human figures are presented to the subject who is instructed to tell a story in which all four pictures appear. Analysis of results is based on both formal and content factors. The test, however, has not gained wide popularity in the United States.

The Thematic Apperception Test of H. A. Murray

The main credit for the innovation of using the picture-story method to study personality belongs to H. A. Murray, who in the 1930s was the moving spirit in developing new projective techniques at Harvard Psychological Clinic for the systematic appraisal of college students. Of these techniques, most of which have fallen into disuse, initially the most promising and eventually the most far-reaching in influence was the Thematic Apperception Test (hereafter referred to as the TAT).

Murray saw the utility of the technique in uncovering the particular images, strivings, and sentiments which are formed in the course of an individual's development. The rationale is explained in Explorations in Personality (Murray, 1938):

The test is based upon the well recognized fact that when a person interprets an ambiguous social situation he is apt to expose his own personality as much as the phenomenon to which he is attending. Absorbed in his attempt to explain the objective occurrence, he becomes naively unconscious of himself and of the scrutiny of others and, therefore, defensively less vigilant. To one with double hearing, however, he is disclosing certain inner tendencies and cathexes: wishes, fears, and traces of past experience [p. 52].

On the basis of their experience with 50 students at Harvard, Morgan and Murray concluded that of all the short procedures and tests which were tried, the TAT gave the best understanding of the "deeper layers of personality" (1935).

Because of Murray's initial reported successes with the procedure, and because he had assembled at the Harvard Psychological Clinic many researchers who were, or were later to become, leading figures in clinical psychology, the TAT took hold where previous attempts at using the picture-story method had attracted only momentary attention.

Since 1938 the scope of inquiry has expanded, the pace of research has quickened. The TAT was employed in the study of a wide variety of psychopathological syndromes: hysteria, anxiety hysteria, and obsessive-compulsive neurosis (Balken & Masserman, 1940), schizophrenic psychosis (Balken & Masserman, 1943); Bennett, 1941; Harrison, 1940; Rapaport, 1943; Rosenzweig & Sarason, 1942), head injury (Rautman & Brower, 1945), psychopathic delinquency (Kutash, 1943), stuttering (Richardson, 1944), and mental deficiency (Sarason, 1943a, 1943b). The TAT has proven a useful instrument in the exploration of such diverse areas as child development (Sanford, 1941), social attitudes and sentiments (Berman, 1943), assessment of military personnel (Murray & Stein, 1943), and culture and personality (Henry, 1947; Oppenheimer, 1945). Paralleling these developments, the TAT in its practical applications has spread out into the neighboring field of counseling and into industrial psychology.

Despite early success and rapid growth in its use, the TAT has been plagued with theoretical and methodological problems. Harrison (1965) gave a comprehensive list of these difficulties:

The TAT has been said to be valid and not valid, reliable and not reliable, predictive of overt behavior and not predictive of overt behavior, useful in research and not useful in research, easily faked and not easily faked, good for diagnosis and not good for diagnosis, economical and not economical, capable of plumbing the unconscious depths and shallower than other projectives. There is also radical disagreement between the subjective, or qualitative, approach on the one hand and the objective, or quantitative, approach on the other [pp. 563-564].

At first sight this is a rather bleak picture. Of what use could the TAT be for clinical practice or personality research if professional opinion about it is so polarized. Yet it continues to be used by both researchers and clinicians with great regularity.

Harrison (1965) accounts for the discrepancies as follows:

Some of the reasons for the contradictions aside from bias, are different kinds and degrees of experience with the test, different scoring and analytical methods, and widely divergent approaches and research methods. It is fallacious to approach the TAT indiscriminately as many test academicians do, as if it were a standard test, comparable to machine made ability and personality tests--to consider it a homogeneous psychometric instrument with set conditions of administration, scoring, tabular norms, and reliabilities--and if such requirements are not met, then to decide that it does not measure up to scientific standards. The TAT is in no sense standardized--in the number and kind of pictures, in the mode of administration or the instructions, in the methods of analysis, or in the uses to which it is put. ...It is not really a test in any usual sense of the word but, like the clinical interview, is more a personality probing or an evaluative method [p. 564].

Although Harrison described the difficulties in using the TAT very well he did not give much of a solution to them. It is certainly not of much comfort to the personality researcher who needs standardized, reliable, and valid instruments to carry on his work. It probably does not help the clinician either to offer him another "technique" which he is to use "intuitively" to "probe" and evaluate his client's personality. Harrison seemed convinced that the TAT has value but that it can be used only in a general, impressionistic way.

Arnold's Story Sequence Analysis

Fortunately, another approach has been proposed to resolve many of the difficulties mentioned above. Arnold (1962) developed a theory and method for TAT interpretation and scoring that shows promise in making it a more valid, reliable, and scientific procedure.

In her TAT sequence analysis, each story is summed up in the

import, i.e., what the storyteller is saying about his life situation. This is a simple condensation of the moral of the story, without any interpretation. Each import, as indicated by the story and outcome, may be either constructive or not constructive. It is constructive when the story import indicates that achievement is the result of effort, initiative, virtue, or the outcome of a definite plan that accepts limitation and adapts to circumstances; when loss, harm, or danger is overcome by positive action; when ill-intentioned action is punished, rejected, or renounced; when others are met with good will, good fellowship, or humor.

The sequence analysis gives us a picture of a man's motivational pattern, his personality structure or "self-ideal" in action. Thus according to Arnold (1962) it can be used to predict whether he will act constructively or not. The normal person will reveal in the sequence analysis how he is ordering his life, not merely what are the raw materials that are so ordered. In the neurotic, the sequence analysis shows the preoccupation with particular areas of conflict, and in the psychotic, the difficulty, even the impossibility, of acting effectively.

Arnold (1962) contended that the inclusiveness of so many studies employing the TAT may be not due to the test itself but to inadequate systems of interpretation and scoring:

A story is not a collection of themes nor is it a string of memory images. A story is a creative reorganization of past sense impressions, a new product of human imagination, very different from personal memories recalled in the original sequence and pattern. The story has a meaning which cannot be discovered from the meaning of the individual themes into which it can be analyzed. It is another truism that a structured whole is not the sum of its parts. Whatever score we may assign to such parts, and however we may manipulate or categorize the elements into which we have divided the story, we are disregarding the story so long as we deal with themes rather than plot and outcome. If it were possible to score

what the story is saying, and include all there is to the story, we would discover a new dimension in the TAT which might allow very different correlations with behavior [p. 13].

Since, Arnold (1962) maintained, it is the story imports that are scored and not story elements, the score does not depend on the picture on which the story is based. The picture sets the motif or theme of the story, but the story plot and outcome are provided by the storyteller. Since it is the import that is scored as positive or negative, and the import abstracts the plot and outcome, each picture has the same chance of yielding a positive or negative score. For this reason, the "stimulus value" of the picture is inconsequential.

Scoring story imports instead of story themes also equalizes the story length. Whether long or short, every story is reduced to an import which is usually contained in one sentence. Imports make it possible to score and evaluate meager as well as rich records.

Administration

The test materials have been published by Murray (1943). Although he experimented with a large number of pictures before he finally selected those to be included in the test, the present standard test contains 31 cards: 10 to be used with both men and women, 10 with men or boys alone, 10 with women or girls alone, and one blank card. The test itself consists of 19 pictures and the blank card which are divided into two sets of ten. Murray (1943) suggested the test be given in two sittings of about an hour each and separated by a day or more.

The instructions to the subject may vary in form with the situation and the subject but their content is standardized (Murray, 1943, pp. 3-4).

As a general rule the acceptance of only one story to a picture

is recommended in the interests of economy of time. However, Lasaga y Travieso and Martinez-Arango (1946) suggest that in the case of disturbed patients one should permit violations of this rule in order to obtain the range of fantasy to a given stimulus.

Traditionally, stories have been transcribed by the examiner from the dictation of the subject but they can also be written by the subject. Harrison (1965) gave the relative merits of both methods:

This arrangement [transcription], while time-consuming and burdensome for the examiner, has several advantages over the written method. It allows for the examiner to draw out nonproductive subjects, and it makes possible valuable behavioral observations...The question is whether for most research projects or for the more indolent clinicians enough additional information is gained in the oral method to justify the inconvenience. The oral method is usually regarded as the method of choice for individual work; for research, the advantages of group administration outweigh the disadvantages [p. 565].

Most of the research, however, shows little or no difference between the oral and written methods for the variables studies (Bernstein, 1956; Clark, 1944; Eron & Ritter, 1951; Lindzey & Heinemann, 1955; Lindzey & Silverman, 1959; Winchester, 1948). In addition, there is no appreciable difference for screen projection or individual handling of cards in group administration (Lindzey & Silverman, 1959) or for single or double sessions (Garfield, Blek, & Melker, 1952).

Order of presentation of the cards makes little or no measurable difference (Bellak, 1944; Kannenberg, 1948; Lowe, 1952; Lubin, 1955; Mason, 1952).

The presence or absence of an examiner may have some effect, according to Bernstein (1956). Analysis of variance showed that when the examiner was absent there was no involvement and the stories were sadder.

When comparisons are made on the effects of different examiners, the studies are in conflict. Turner and Coleman (1962), concluded that there were only slight effects produced with different personal characteristics, different amounts of test experience, and different attitudes toward the test. Similar results were obtained by Sumerwell et al. (1958), while Garfield et al. (1952) reported that the sex of the examiner exerted no influence on the plot level, activity of the hero, mood, or outcome.

Stimulus conditions

The pictures were originally chosen according to following criterion given by Murray (1943):

In estimating the effectiveness of a picture we waited in each case until the personality of the subject taking the test had, with the help of other methods, been thoroughly studied and understood; and then we rated the picture according to the amount of information contributed to the final diagnosis by the story it had evoked. The average of such ratings was accepted as a measure of the stimulating power of the picture....

Experience has shown that in the long run the stories obtained are more revealing and the validity of the interpretations is increased if most of the pictures include a person who is of the same sex as the subject. This does not mean that it is necessary to have two completely different sets of pictures, since some pictures of proved value contain no human figures; others portray an individual of each sex, and in others the sex of the one figure shown is questionable. In fact, eleven of our pictures (including the blank card) have been found suitable for both sexes [p. 2].

However, Murray does not say why or how the stories are "more revealing" and the "validity of interpretations is increased" if most of the pictures include a person who is of the same sex as the subject.

Arnold (1962), by contrast, maintained that the content of the picture is of little consequence for the over-all results of the test:

Since the play of imagination is the important factor in a story, it is the action, plot, and outcome that matter, not the picture about which the story is told or even the characters who appear in it. We

use the M [male] pictures of Murray's TAT set for both men and women because these pictures are dramatic and lend themselves easily to imaginative exploration. Actually, any general kind of picture could be used. We have obtained good imports and sequences from pictures found in Life and from short descriptive sentences, for instance: "Tell me a story about a boy and a violin." A person who has a rich and fertile imagination can tell good stories about almost anything. However, for those whose imagination is sluggish, pictures illustrating a dramatic situation are a decided help [pp. 51-52].

This would explain why a large number of studies show card differences for productivity, degree of involvement, emotional tone, nature of stories, interpersonal relations, outcome, and expression of needs (Bijou & Kenny, 1951; Coleman, 1947; Eron, 1948; Eron, 1950; Eron, 1953; Eron, Terry, & Callahan, 1950; Gurel & Ullman, 1958; Kutash, 1943; Lindzey & Goldberg, 1953; Lowe, 1952; Lubin, 1960; Murstein, 1958; Newbigging, 1955; Sarason & Sarason, 1958; Starr, 1961; Terry, 1952; Ullman, 1957). Nevertheless, according to Arnold (1962), imaginative productivity, as such, is not the essential element in interpreting and scoring TAT stories.

In addition, the assumption has been made that the similarity between the storyteller and the central figure in the picture encourages identification, even though clinical experience has shown that subjects may make identifications across age and sex lines. The desirability of close physical similarity between subject and picture was not supported by the investigation of Weisskopf and Dunleavy (1952), who found that converting the Murray figures into crippled or obese figures did not increase productivity for persons with these handicaps. Similarly, replacing the face of the central figure with a photograph of the subject did not increase productivity (Weisskopf-Joelson & Money, 1953). In working with nuns, Lasaga y Travieso and Martinez-Arango (1946) did not find that using figures of nuns improved the diagnostic value of the test. McIntyre (1954) obtained the same kind of results for similarity of age and sex to

subject population in his investigation of short filmstrips as a projective medium.

The evidence prompted Harrison to the following conclusion concerning stimulus effects:

While the examiner should develop sophistication about stimulus effects, there is a tendency in some quarters to overstate the role of the stimulus and underestimate the role of personality projection. Under ordinary conditions of administration, more story variance is accounted for by projection or expression of personality than by the stimulus complex. Most of the administrative and stimulus variations demonstrated in the laboratory are mainly of academic interest and are not of primary importance in the usual testing situation [p. 568].

Therefore Harrison is in substantial agreement with Arnold's evaluation, i.e. that the personality of the storyteller comes through regardless of the stimulus picture.

Methods of Interpretation and Scoring

The issues discussed in this section have been touched upon earlier but some further elaboration at this point is necessary.

The theory of personality used by TAT interpreters from Murray to the contributors to the symposium edited by Kagan and Lesser (1961) has usually been psychoanalytic theory. This proposes that instinctive drives or impulses are the real motivating forces of the individual which are modified by ego-processes in overt behavior, but reveal themselves in fantasy. This is possible because in fantasy the primary process stemming from these drives is concealed from the ego by various defense mechanisms (projection, identification, etc.). Early TAT workers, particularly, assumed that TAT stories, like fantasy in general, express the storyteller's needs or drives.

Later research findings, however, did not support the theory (cf. Arnold, 1962, pp. 4-11). It was found that the drives and affects

that are assumed to be projected upon the story characters are not a sure guide to the kind of motivation that leads to action in everyday life. Arnold (1962) suggests that if the drives and affects are isolated thematically they may indicate preoccupation but not genuine striving. The storyteller's motivation is illustrated not by the story themes but by the story outcome and the way the story is told.

This brings us to the assumption underlying the scoring methods developed thus far. It is the notion that the story is an aggregate of themes, and these themes must be isolated before scoring is possible. Themes may be categorized in a variety of ways, from Murray's needs and press, Tomkins' vector and level analysis, to Kagan's affect states; but in every case, the scoring is based on isolated parts of the story. This assumption is based on Freud's notion of fantasy, as exemplified by his dream analysis: fantasy images are conceived as the end product of a causal chain stemming from drives aroused by the original traumatic experience. For Freud, every fantasy is a wish fulfillment, an image of the object cathected by the drive. Influenced by Freud, most modern psychologists seem to think of fantasy as a series of personal memories, strung together haphazardly, but connected and cleverly disguised by the superego.

This view of the TAT story as a patchwork of personal memories has not changed, despite the changing view of what it is that is expressed in it. At least this is the conclusion that seemed to be implied by the unanimous agreement of TAT experts to score each story theme individually.

Arnold (1962) viewed this as the cause for most of the uncertain results of TAT research on personality:

Despite Henry's insistence that "there is no single element in a given TAT record that has any meaning in and of itself" (1961, p. 117), the themes are scored one by one; and the variables supposedly revealed in these themes remain similarly isolated. Veroff (1961), for instance, reports that he found no correlation between n Achievement, n Affiliation, n Power, in a nationwide survey of a representative sample of 1619 men and women. Since McClelland's system is one of the best worked-out scoring methods, it is not likely that we may expect better success from other systems if they are based on the same assumptions.

If it were true that the themes represent the storyteller's needs, drives, or affects when they are not blocked by ego-processes, and that ego-processes have a similar connection with themes, we should be able to use TAT themes for predicting behavior at least in those cases where fantasy expression is not blocked by ego-defenses. In such cases, the storyteller's actions should duplicate the hero's actions; the storyteller should have the same desires, the same anxieties and inhibitions he ascribed to the hero.... Unfortunately for prediction, but perhaps fortunately for this hypothesis, we never can tell whether themes revealed in the TAT accompany behavior or are an alternative to behavior; whether themes missing in the TAT indicate lack of the corresponding need, its blocking by ego-defenses, or its being acted out in reality. It almost seems as if we had to agree with Lazarus, who insists that we will never be able to predict behavior from the TAT alone.

Before we subscribe to such a pessimistic conclusion, we might consider another possibility. Perhaps neither drives nor ego-processes are revealed in the story themes. Perhaps motivating tendencies, whatever they are, shape the story action and are expressed in the story outcome [pp. 11-12].

There are several interpretive techniques which use the outcomes in addition to story themes (McClelland et al., 1953; Eron, 1951; Hartman, 1951). However, the outcome is not just another theme like aggression, affiliation, achievement and the like. The outcome caps the plot, and the plot integrates the various themes into a unified whole.

Therefore Arnold (1962) proposed that each story be condensed into an import that leaves out incidental details but preserves the kernel--the meat of the story. When all the imports are read in sequence, a picture of the individual emerges that does portray his attitudes and intentions for action. Every story makes a point, expresses

a conviction. It describes an action that may be headed for success or failure, may exemplify cooperation or hostility, may be an attempt to cope with adversity or betray spineless acceptance of what may come. In every case, emotions may influence the action, but the outcome is primarily an expression of the storyteller's convictions, gathered from experience and reflection. The plot sets a problem, the outcome solves it. Both the type of problem a man sets for himself and the kind of outcome he prefers are characteristic for him. It is then possible to score such story imports objectively and to arrive at a final score which will indicate the individual's positive or negative motivation.

Arnold's (1962) scoring method has several advantages. Since it is the story imports that are scored and not story elements, the score does not depend on the picture on which the story is based. The picture (or description) sets the motif or theme of the story, but the story plot and outcome are provided by the storyteller. Since it is the import that is scored as negative or positive, and the import abstracts the plot and outcome, each picture has the same chance of yielding a positive or negative score. Therefore, according to Arnold (1962), the "stimulus value" of the picture is irrelevant.

Scoring story imports instead of story themes also equalizes the story length. Whether long or short, every story is reduced to an import which is usually contained in one sentence. Imports make it possible to score and evaluate meager as well as rich records.

Finally, the method reveals positive and negative motivation as expressed in adjustment and achievement. This motivation is stable according to Arnold:

In several cases, we have given repeated tests, using different pictures, in intervals up to one year. The positive or negative attitudes revealed in successive sets of stories were similar and obtained similar scores even though the actual imports were quite different [p. 15].

The critical process, then, in Arnold's method is extracting the import of each story. The import is a condensation of the story which preserves the kernel of the story but leaves out the incidentals. Each story makes a point, and when this point is extracted in the import, and the imports are strung together in a sequence, a picture emerges of the person's motives and intentions to actions:

If each story is an imaginative exploration of various problems and their possible solutions, we must try to isolate what it is the storyteller is trying to say. What he says about the picture will reveal his convictions: what could be called the "moral" of the story. When this moral is applied to the storyteller's subjective circumstances, we arrive at the import (the meaning or significance) of the story. Once the import of each story is set down in sequence, it becomes possible to follow the storyteller's trend of thought, which reveals his habitual dispositions, the way he evaluates human actions, and the circumstances of man's life. The story import will show how the storyteller thinks people usually act and how he feels they should act; what actions he thinks right and wrong; what will lead to success, in his opinion, and what to failure; what can be done when danger threatens, and what things to strive for. In short, the story imports, taken in sequence, give a connected statement of the storyteller's principles of action, his motivational pattern (Arnold, 1962, p. 51).

Skill in taking imports from stories can be learned through instruction and practice. It is an objective distillation of the central meaning of the story in that it is abstracted as well as possible without adding from subjective interpretations. However, because of the examiner's focus on a particular problem around which the story is spun, imports as well as the sequence in which they are embedded may vary somewhat from examiner to examiner. Some stories are straightforward and clear as to what the storyteller is trying to say. With others, the point of the story is less well made, outcomes are implied rather than

stated clearly, stories border on monologues, and the import is difficult to score. Here skill in abstracting has to be developed, and when the stories are written with definite plots and outcomes, the meaning of the import is consistent when taken by different examiners although the particular wording of it may vary considerably. When the import is scored using the categories that have been devised, the variations in the wording of the import become unimportant and the import is preserved in a unified score.

There are two aspects of scoring: classifying the content of the story and giving it quantitative value. Imports fall into four broad, thematic categories: I. Achievement, success, happiness, active effort and their opposites; II. Right and Wrong; III. Human Relations; and IV. Reaction to Adversity. These categories were derived empirically from the records of high and low achievers, from effective and ineffective teachers, from offenders and non-offenders among Navy recruits, and efficient and inefficient executives in a government project. Over a period of ten years these categories of imports were organized into a scoring system with various headings and sub-headings to include the types of imports found in stories.

Quantitatively, each story has four possible scores ranging from +2 to -2. These scores derive from the types of imports found in positively motivated persons as contrasted with those found in negatively motivated persons and were derived empirically from the contrasting groups mentioned above. The algebraic sum of all imports is transformed into a score based on the number of cards used and is called the Motivation Index. The M.I. is calculated according to the following formula:

$$\frac{n^o}{n^p} \quad 200$$

Where n^o is the number of units obtained, the algebraic sum of the scores of the imports, and n^p is the number of units obtainable or possible. The n^p is four times the number of the cards used: e.g. for a 20-story sequence, the n^p is 80. Arnold provides a table to make this transformation easy (Arnold, 1962, pp. 146-147).

Since it is the story import that is scored and not the picture about which the story is told, this formula can be used for any set of 20 consecutive stories--whether TAT cards or other pictures are used or whether stories are elicited by a verbal description (e.g. "tell me a story about a boy and a violin")--provided only that they are told spontaneously and in sequence and are scored according to the criteria.

Since each story is scored by reference to the same scoring criteria, and since the import derived from the story is scored and not the picture about which the story is told, Arnold (1962) has argued that every "item" in the test must be equal to every other "item."

This has been shown to be the case by Fagot (1961) who found that the two positive and two negative scores were evenly distributed among all the cards used in a sample of 252 records of twelve stories each. For this investigation, Fagot used the records of 99 teachers (Burkhard, 1958), 100 college students (Garvin, 1960) and 53 seventh-grade children (Arnold, 1961).

Since the "items" measured by the intensity scale are equivalent, even one story should theoretically give a correct picture of the

storyteller's intensity and direction of motivation if he were completely consistent and the scoring were completely reliable. Unfortunately, not many individuals are so completely consistent in positive attitudes that they always tell stories with highly positive imports. On the other hand, it is fortunate that there are still fewer people who have such strong (and strongly entrenched) negative attitudes that they would reveal them in every story they tell. Each story, then, is considered as a sample of the storyteller's motivating attitudes, his habitual motivation. The intensity score is a measurement of this motivation. Demanding several stories is equivalent to applying repeated measurements. Instead of measuring various dimensions of the quality we want to measure, as is done in intelligence tests and most personality tests, many measures of one and the same dimension are taken, which is the storyteller's motivation, positive or negative. Highly consistent people will reveal essentially the same type of motivation in every sample and show high "reliability"; as do, for instance, the good and poor teachers. But in a group that represents the whole range of consistency, one has to take quite a few samples before he can be sure that he can strike an average that will fairly represent the individual's degree of consistency. Arnold (1962) does not recommend to use fewer than ten TAT cards or show fewer than ten pictures from other sources. She proposed that there is a progression in thought from one story to another which requires several stories that give positive, though conventional, imports and only gradually reveals a negative attitude as the person comes to grips with his/her own problems. Another person may start out with stories that explore several solutions to his particular problem--some positive, some negative--until, toward the end

of the sequence, the basic positive orientation prevails. There is an optimum number of samples, i.e., one must have enough stories so that further samples will neither significantly increase or decrease the final score. Ideally, there should be a point where additional stories would not change the Motivation Index even for inconsistent storytellers. Too few stories may give a spurious measure of consistency, while too many stories will extend the testing time and exhaust the storyteller without materially improving the consistency score. For clinical purposes Arnold (1962) recommended 13 stories and for research purposes 10 seem to be sufficient.

Reliability in general

Anastasi (1968) says that reliability is that aspect of a test which refers to the consistency of scores obtained by the same individuals on different occasions or with different sets of equivalent items. The concept of test reliability has been used to cover not one but several aspects of score consistency.

The most obvious method of finding the reliability of a test is by means of a retest, or repetition of the identical test on a second occasion. The reliability coefficient in this case is simply the correlation between the scores obtained by the same subjects on the two administrations of the test.

Although apparently simple and straightforward, this technique presents difficulties when applied to most psychological tests. Practice will probably produce varying amounts of improvement in the retest scores of different individuals. If the interval between the retests is fairly short, the subjects may recall many of their former responses. Thus, the scores on the two administrations of the test are not inde-

pends on the order in which the tests are obtained and the correlation between them will be spuriously high. Only tests that are not appreciably affected by repetition lend themselves to the retest technique.

One way of avoiding the difficulties encountered in the retest reliability is through the use of equivalent forms of the test. The subjects can then be tested with one form on the first occasion and with another, comparable form, on the second. The correlation between the scores obtained on the two forms represent the reliability coefficient on the test.

It is also possible to arrive at a measure of test reliability by various split-half procedures. In this method, two scores are obtained for each individual by dividing the test into comparable halves. Split-half reliability provides a measure of equivalence, or adequacy of item sampling.

Projective tests like the TAT leave a good deal to the judgment of the scorer. With such tests, there is as much a need for a measure of scorer reliability as there is for the more usual reliability coefficients. Scorer reliability can be found by having a sample of test papers independently scored by two examiners. The two scores thus obtained by each subject are then correlated in the usual way, and the resulting correlation coefficient is a measure of scorer reliability.

Reliability of the TAT

Establishing reliability has been a perennial problem with regard to the TAT and the results of reliability studies have been generally inconclusive. TAT research workers themselves lack agreement as to what constitutes a measure of reliability, nor, apart from scorer reliability, have they provided a succinct statement as to what the

status of TAT reliability might be.

Henry and Farley (1959) noted some of the difficulties in estimating reliability with regard to projective techniques (including the TAT). They are convinced that stability across time is the crucial test of projective instrument reliability. Nevertheless, too short an interval between successive administrations introduces the memory factor, and too long an interval introduces the possibility of actual changes taking place in the subject.

They dismiss split-half, alternate form, and item consistency methods of obtaining estimates of reliability as inappropriate to projective testing, since patterning and sequence of responses, which, for them, is a crucial element in projective instrument interpretation is thereby lost.

Harrison (1965) strikes an equally pessimistic note when he says:

The conventional reliability methods that have been evolved for psychometric tests, such as measures of ability and achievement, cannot be applied in unmodified form to projective techniques like the TAT which are based on entirely different principles. Thus split-half reliability is not suited because, unlike psychometric items, the pictures were designed to yield psychologically different, not equivalent, data. Not only can the pictures not be made to fit into psychologically equivalent pairs, but there are not now available, and are not likely to be available in the future, equivalent sets which could be expected to give the same kind of material [p. 589].

Unfortunately, from the studies reviewed below the pessimism seems warranted.

Tomkins (1942) administered a new picture to a single subject five days a week, over a period of ten months. He also gave the entire test three times at three-month intervals and once when the subject was under the influence of alcohol. He found that the main themes all appeared on the first full administration of the test and that in spite

of attempts to make the stories different, they reoccurred on subsequent administrations.

Sen (1953) found that correlations between judgments made by the same judge at intervals nine months apart, based on TAT records of 100 British Civil Service candidates, was .59. Correlations between independent ratings made by two judges, although moderately high, were not extremely encouraging.

Haber and Alpert (1958) explored test-retest reliability of n Achievement scores using a carefully pretested equivalent set of pictures. College students (N of 26) were tested under controlled, relaxed conditions and then again with an equivalent set of pictures, after a three-week interval. The correlation coefficient was .54. When six pictures which were strongly suggestive (highly structured and cued in terms of n Achievement) were used, the estimate of reliability was .74; when six weakly cued pictures were used, a substantial drop in reliability was found, the coefficient dropping to .54. When another group (N of 54) was first tested under relaxed, then under achievement conditions, the correlations of equivalent forms dropped to .45.

The reliability estimates in this study seem to be related not only to the specific stimulus situation and their ambiguity, cue-value, etc., but also the specific motivating cues in the situation in general.

This notion is also suggested by Birney (1959) who, using sets of weakly cued pictures, and scoring for four n Achievement measures, reported coefficients ranging from .03 to .56 on the basis of a test-retest study over a two-year period, considerably lower than those reported in other studies.

Kagan (1959) reported on the basis of exploration of the stabil-

ity of eight TAT scoring categories over almost a 6 year period using 86 children tested at three intervals, that only physical aggression by the hero and achievement themes showed better-than-chance stability. His findings suggest that the more likely it is that a stimulus produces a given theme, the more likely would this remain stable.

Other reports, where observations on temporal stability were incidental, have indicated consistency over time but do not give statistical data (Fisher, 1958; Greenwald, 1959; Symonds & Jensen, 1961).

Where statistics are available, they more often report moderate to high correlations (Lindzey & Herman, 1955; Morgan, 1953; Tomkins, 1947; Weisskopf & Dunleavy, 1952), but sometimes the consistency is poor (Bradley, 1957; Krumboltz & Farquar, 1957).

Many investigators are of the opinion that internal consistency as usually calculated is not appropriate for the TAT, as a result, there have been relatively few split-half coefficients reported (Bialick, 1951; Calogeras, 1958; Child et al., 1956; Lesser, 1957; Lindzey & Herman, 1955; Sanford, 1943). With the exception of the Child study, which had a range from $-.10$ to $.44$, the resulting coefficients fell between $.40$ and $.85$. Most of the results were obtained from varying numbers of the Murray cards with needs and press as the usual measure.

There is sufficient evidence which indicates that scorer reliability is reasonably good when there is pretraining and supervision of the judges and scorers, and when the scoring system is specific with systematic rules.

In a review of several studies of interjudge reliability, Feld and Smith (1958) reported scoring reliabilities ranging from $.66$ to $.96$ with a median of $.89$.

Rank-order correlations between scores of different scorers ranged from .72 to .91 in a study reported by Veroff, Atkinson, Feld, and Gurin (1960). These scores were based on a specific and clear-cut analysis relevant to n Achievement, n Affiliation, and n Power.

Sarason and Sarason (1958), using Eron's scales for emotional tone and outcome, found an average interrater reliability, for 12 cards, of .87 for emotional tone and .79 for outcome; this is of the same order as obtained in earlier studies using these scales (Eron, Terry, & Callahan, 1950; Liccione, 1955).

In general, specific scoring systems have yielded relatively high scorer reliability figures whereas more holistic procedures of interpretation have tended to yield, as would be expected, lower reliability figures.

Reliability of the Motivation Index of Story Sequence Analysis

There has been very little direct investigation of the reliability of the Motivation Index; that which has been done has been mainly incidental.

Burkhard (1958) in her study on the motivational differences between good and poor teachers found high consistency scores, the good teachers in a positive direction, the poor teachers in a negative direction. The split-half correlation between the odd and even numbered stories was .86.

She then isolated the highest and lowest protocols by excluding the midrange of consistency scores and found that the first half of the sequence analysis (6 stories) had a correlation of .86 with the second half (6 stories).

Garvin (1960) followed the same procedure in his study on motivation and school achievement and found a correlation of .79 between the odd and even numbered stories. In contrast, the correlation of the first half of the sequence analysis with the second half dropped to .61.

Arnold (1974) carried out a test-retest reliability under both normal and "fake good" conditions. He found moderate but statistically significant correlations in both cases. The correlation between the pretest and the posttest of the normal conditions group was .58 (significant at the .01 level); the correlation between the pretest and the posttest of the "fake good" conditions groups was .46 (significant at the .05 level).

There have been no alternate-form reliability studies on the Motivation Index of Story Sequence Analysis.

As for scorer reliability, Burkard (1958) found 97 and 94 per cent agreement between herself and two other scorers in scoring 1,200 stories; Petrauskas (1958) found 80 and 82 per cent agreement between himself and two other scorers in scoring 780 stories; Arnold (1974) found 78 per cent agreement with one other scorer in scoring 624 stories.

Validity

Since the present study does not deal with validity as such there is no need to present a thorough review of the literature.

For the basic concepts the reader is referred to Anastasi (1968) [pp. 99-157].

Good surveys of validity studies on the TAT are provided by Harrison (1965) [pp. 591-597] and by Zubin, Eron, and Schumer (1965) [pp. 420-441].

Arnold (1962) presented adequate information on validity [Chap-

ters 10-14]. She and her doctoral students have reported some very successful scholastic and vocational predictions. On the basis of the Story Sequence Analysis they have been able to distinguish sharply between such criterion groups as achieving and nonachieving high school and college students (Brown, 1953; Garvin, 1960; McLandish, 1958; Snider, 1954) and more and less successful teachers (Burkard, 1958), seminarians (Quinn, 1961), Federal administrators (Steggert, 1961), and Navy enlisted personnel (Petrauskas, 1959).

Sex Differences in Productivity and Thematic Preoccupation

The systematic investigation of sex differences has been neglected except for a couple of studies by Lindzey and collaborators (Lindzey & Goldberg, 1953; Lindzey & Silverman, 1959), who obtained significant differences; but, except for the greater verbal responsiveness of women, the differences were not consistent from one college to another.

Almost all writers agree on the fact of TAT sex differences, but again because of the specificity of scoring and interpretive methods and because of population sampling, establishing firm trends presents some difficulty. A few writers have not been able to report any real differences (Cox & Sargent, 1948; Myers, 1958; Rosen & Neugarten, 1960); the most likely explanation is to be found in their scoring procedures. That TAT sex differences exist is indicated by Shneidman and Farberow (1958) whose judges could determine sex with significant accuracy from the protocols of adult subjects in 75 per cent of the cases.

There is agreement that girls and young women are more productive and fluent and show more imagination and psychological involvement than males of comparable age (Abel, 1945; Aron, 1949; Ericson, 1947; Lindzey & Goldberg, 1953; Lindzey & Silverman, 1959; Roquebrune, 1959;

Weisskopf, 1950), with only Foulds (1953) dissenting. Except by Kagan (1959), males are reported to show more themes of aggression (McDowell, 1952; Sanford, 1943a; Sarason, 1943b; Symonds, 1949; Whitehouse, 1949). Females show more affect (Aron, 1949; Kagan, 1961; Sarason, 1943), relate sadder stories (Eron et al., 1950; Newbigging, 1955), and are more concerned with nurturance (Kagan, 1959; Lindzey & Goldberg, 1953). Usually their stories reflect a greater interest in affiliation and personal relations (Sanford, 1943; Webster, 1953; Whitehouse, 1949; but this finding is not supported by Lindzey and Silverman (1959). In the case of institutionalized defectives, themes of loneliness were more prevalent in the stories of girls (Abel, 1945; Sarason, 1943a). Sanford (1943) found that girls manifested greater need for curiosity, deference, exhibition, and avoidance of blame and humiliation than boys.

Sex Differences in Achievement Motivation

Atkinson and his coworkers (Atkinson & Feather, 1966) have carried out a very systematic and impressive experimental analysis of the contemporary determinants of achievement-oriented activities. As a result of this work, there is a large body of data about achievement motivation in men.

Data for women, on the other hand, have been scarce. Atkinson (1958) filled more than 800 pages with a compilation of available theory and data on achievement motivation. The question of sex differences was treated only in a footnote. It is admittedly a long footnote in which he refers to the issue of sex differences as "perhaps the most persistent unresolved problem in research on n Achievement" [p. 33].

The major sex difference--at least the one that has received the greatest amount of attention--has been that women, unlike men, fail to

show an increase in their achievement-imagery score when they are exposed to experimental conditions that arouse achievement motivation by stressing "intelligence and leadership ability" (Veroff, Wilcox, & Atkinson, 1953). Under neutral conditions the scores of women are as high or higher than those of men. The evidence from the more recent studies on women's motivation to avoid success provides some understanding of this problem. The earlier studies, which will be considered briefly, have given a generally inconsistent pattern of results.

A study by Field (1951) suggests that achievement motivation in women can be aroused by referring to their social acceptability rather than to their "intelligence and leadership ability."

On the other hand, Angelini's (1955) data on Brazilian university women argue that "intelligence and leadership" arousal is effective provided the sample used is made up of highly competitive women who value intellectual accomplishment. The implication is that women at large American coed universities are more socially than intellectually oriented.

Lesser, Krawitz, and Packard (1963) tested Angelini's hypothesis within American society. They conducted their study at Hunter High School for girls in New York City. The school places great emphasis on the intellectual accomplishments of women. The results of the study were disappointing. Despite the fact that these girls were highly competitive and valued intellectual accomplishments, no overall increase in achievement imagery was found under arousal conditions stressing intelligence and leadership ability. However, an interesting pattern of interaction was noted. The impact of the arousal condition on the TAT responses of the girls who were doing well at the school, compared with

those who were not (with IQ scores matched), varied depending on whether the dominant stimulus figure on the TAT cue was male or female. The "achievers" showed an increase in achievement-motivation score under arousal conditions only to pictures of females and the "underachievers" only to pictures of males. Assuming that most of the girls who go to Hunter value achievement and see it as a relevant goal, the explanation offered for these results in terms of differential perception of social role among the girls is probably not sufficient. For instance, why should the achievers--the girls who do well and presumably value achievement more than those who do not do well--show an increase in achievement motivation to female pictures, most of which depict women involved with traditional activities?

Other studies have attempted to relate such factors as individual value orientation, achievement relevance of goals, sex of the TAT stimulus figure, nature of arousal conditions (French & Lesser, 1964), and sex-role orientation (Lipinski, 1965) to achievement-motive scores and performance. The results, however, have been so inconsistent that, instead of resolving the problem of achievement motivation in women, they have only further emphasized the vast complexity of the issue.

One possible solution to these discrepancies was proposed by Horner (1970):

Under achievement-oriented conditions that stress "leadership or intellectual ability," women may inhibit expression of their achievement motivation on the TAT because of the concurrent arousal of anxiety about failure and anxiety about success. Thus women's TAT scores may be not an accurate or valid measure of the strength of their achievement motive and cannot be expected to relate to performance in the same way as men's scores do [pp. 57].

In her experimental analysis Horner (1970) assumed that fear-of-success may be a relatively stable personality disposition (or motive) in

women, that it is aroused in competitive achievement situations especially where important men are present (e.g. prospective dates, boy friends), and that it may underlie many of the major sex differences that have appeared in the research on achievement motivation (see Bardwick, 1971; Horner, 1968).

Her procedure was similar to the standard TAT procedure developed by McClelland and his colleagues (1953) to assess need Achievement except that stories were written to verbal leads rather than to pictures. For women the verbal cue that was of particular significance was "After first-term finals, Anne finds herself at the top of her medical school class." The female cues were presented to female subjects only and the male cues (John at the top of his medical school class) to male subjects only. The subjects were predominantly college freshman and sophomores.

Horner (1968) found in her initial study that approximately 65 per cent of her female subjects wrote fear-of-success stories to the Anne cue. In contrast, 9 per cent of the male subjects wrote fear-of-success to the John cue. She indicates that thus a pattern of sex differences in the production of fear-of-success imagery has been maintained in subsequent studies involving white men and women except that in recent years there has been an increase in the amount of fear-of-success imagery expressed by men--men are taking an increasingly negative view of success (Horner, 1972, Table 1).

Feather and Raphelson's (1974) study was aimed at filling a gap in Horner's original procedures by requiring both male and female subjects to respond to the Anne and John cues in the four different possible combinations (male subjects-male cue, male subjects-female cue, female subjects-female cue, female subjects-male cue). The complete procedure was aimed at discovering whether male and female subjects had

shared conceptions about the consequences of male and female success--if males as well as females would be able to identify the negative consequences of female success when writing to the Anne cue, and if females as well as males would write very few fear-of-success stories to the John cue. If such evidence appeared it would suggest that the projective procedure would be tapping societal conceptions or stereotypes about the male and female sex roles in regard to what are considered to be appropriate achievements and occupations for the two sexes.

Compared with earlier results (Horner, 1968), the proportion of fear of success stories written by both males and females to the John cue was greater--agreeing with the trend noted by Horner (1972). Moreover, this trend has also been noted in another recent study (Feather & Simon, 1973).

The main drawback in these "fear of success" studies has been the lack of adequate story sampling. Therefore Feather and Raphelson (1974) recommend that:

Future studies should also be directed to closer scrutiny of the assessment procedure itself. For example, it would be an obvious advantage as far as reliability is concerned to obtain more than one story from each subject if the TAT projective procedure is employed [p.200].

Zuckerman and Wheeler (1975) gave a thorough review of the subsequent research on various aspects of the fear-of-success phenomenon. They traced the development of the theory on fear-of-success, reviewed Horner's (1968) original experiment, reviewed the problems raised by subsequent studies on fear-of-success, and mentioned the new directions of the fear-of-success research.

Relationship of the literature to the present study

The reliability of the TAT, despite the many studies concerned

with it, remains a problem. The key difficulty with regard to most systems seemed to be the method of scoring. As long as the story is scored according to the number of "themes" no story can adequately be compared to any other story. The unreliability seems to be intrinsic to the system. This, of course, does not mean that the system is clinically useless but it does make it less appropriate for research purposes.

In Arnold's (1962) system at least this source of unreliability seems to be eliminated. It is scored in such a way that the stories (i.e., the imports of the stories) are psychometrically equivalent. Each story gets one score and one score only. This does not mean that there are no other sources of unreliability. The main source is the inconsistency of the individual's motivation. If the person's attitudes and convictions are not firm there can be considerable variation in the Motivation Index from test to test. Other sources of unreliability are on the side of the interpreter. The import may be improperly derived and the numerical score misapplied. These difficulties, however, can be obviated by adequate training in the scoring system and by careful checking. It is recommended that all imports and scores be done by two independent examiners. The discrepancies can then be resolved in conference.

What seems necessary at this point is to evaluate the reliability of Story Sequence Analysis from different aspects given the dearth of studies in this area.

The review also brings to light the ambiguity of the term "theme." Again for most scoring systems the theme is the essential element. For Arnold it indicates the storyteller's preoccupations but

not his motivational attitudes. The latter is assessed by the outcome. With regard to stimulus variables, then, this question could be posed: Does the kind of card affect the motivational attitude or merely the thematic preoccupation?

Another difficulty raised in the literature is the puzzling inconsistency between achievement data on male subjects and that on female subjects. A possible solution to this problem seems to depend on the revision of the scoring systems. Motivation tended to be measured by achievement themes or lack of them. Arnold's system, by contrast, derives the Motivation Index not only from achievement theme scores but also from three other areas: Right and Wrong, Human Relations, and Reaction to Adversity. This gives the interpreter a broader base from which to judge the over-all motivational level.

Lastly, there is Horner's hypothesis concerning the fear-of-success motive in women. Is it in fact a culturally determined phenomenon or is it simply one motive among many which may or may not be characteristic of women?

Purpose

The general aim of the study was to propose and test hypotheses with regard to the questions raised above.

The main goal was to investigate the reliability of the Motivation Index of Story Sequence Analysis for male and female subjects under alternate-form (male and female cards) conditions. More specifically the conditions were as follows: (a) male subjects wrote stories to the male cards on the first testing and to female cards on the second testing; (b) male subjects wrote stories to the female cards on the first testing and to male cards on the second; c) female subjects

-- female cards first, male cards second; (d) female subjects -- male cards first, female cards second. The reliability coefficient for each of these was calculated. Then the over-all reliability coefficient for males between the first and second testing was found and was compared to that of females for the first and second testing. Lastly, the composite reliability coefficient between first and second testing for males and females was found. These comparisons indicated whether the sex of the subject and the stimulus conditions have a differential effect on the reliability of the Motivation Index.

The next task was to examine the relative effect of the sex of subject and stimulus conditions on the thematic preoccupations evidenced in the stories. This was done by answering the following questions:

(a) Do male subjects, over-all, tell more achievement stories than female subjects? Do males tell more achievement stories to male cards than to female cards? (b) Do female subjects, over-all, tell more human relations stories than male subjects? Do female subjects tell more human relations stories to female cards than to male cards? (c) Do male subjects tell more reaction to adversity stories than female subjects? Do they tell more stories with this theme to male or female cards?

Another aim of the study was to investigate the differences in motivational level between male and female subjects as measured by the Motivation Index. This was done by making three comparisons: (a) Do males differ significantly from females with regard to Motivation Index derived from the composite score of the two testings? (b) Do males differ significantly from females on the Motivation Index from stories told to the male pictures? (c) Do males differ significantly from females on the Motivation Index from stories told to the female pictures? This should

make it possible to find out if differences in Motivation Index are really sex differences or are simply a result of stimulus variables.

The last objective is exploratory rather than correlational or experimental. Horner's (1968) hypothesis for the phenomenon of fear-of-success in women was based on an extremely small sample of stories. This study will have a sample of 22 stories from each subject which can then be inspected for evidence of this hypothesis. If sufficient indication for this phenomenon is found on the basis of the larger samples, studies can be designed to test it more accurately.

Hypotheses

The hypotheses to be tested in this study can now be expressed concretely under three main headings: reliability, thematic preoccupation, and motivational level.

1. Reliability

Hypothesis I. The Motivation Index scores of Story Sequence Analysis are significantly reliable under alternate-form (male and female cards) conditions. This main hypothesis is divided into seven subhypotheses.

a. There is a significant correlation between the Motivation Index scores of male subjects writing stories to male TAT cards on the first testing and female cards on the second testing.

b. There is a significant correlation between the Motivation Index scores of male subjects writing stories to female cards on the first testing and male cards on the second testing.

c. There is a significant correlation between the Motivation Index scores of female subjects writing stories to female cards on

the first testing and male cards on the second testing.

d. There is a significant correlation between the Motivation Index scores of female subjects writing stories to male cards on the first testing and female cards on the second testing.

e. There is a significant correlation between the composite Motivation Index scores for male subjects under alternate-form conditions. By composite scores is meant that the male card scores are correlated with the female card scores of both male groups without regard to the order of testing.

f. There is a significant correlation between the composite Motivation Index scores for female subjects under alternate-form conditions.

g. There is a significant correlation between the composite Motivation Index scores for male and female subjects under alternate-form conditions.

2. Thematic preoccupation

Hypothesis II. There are no significant differences in thematic preoccupation with regard to male and female subjects or stimulus variables. This hypothesis is divided into four subhypotheses.

a. Male subjects do not tell significantly more stories with achievement themes than female subjects.

b. Male subjects do not tell significantly more stories with achievement themes to male cards than to female cards.

c. Female subjects do not tell significantly more stories with human relationship themes than male subjects.

d. Female subjects do not tell more stories with human

relationship themes to female cards than to male cards.

3. Motivational level

Hypothesis III. There is no significant difference in motivational level between male and female subjects. This is divided into three subhypotheses.

a. The Motivation Index derived from the combined scores of the two testings of male subjects does not differ significantly from that of female subjects. By combined score is meant the average of the two Motivation Index scores derived from the male and female cards.

b. The Motivation Index scores of male subjects do not differ significantly from those of the female subjects when writing stories to male cards.

c. The Motivation Index scores of female subjects do not differ significantly from that of male subjects when writing stories to female cards.

4. Fear-of-success themes

There was no specific hypothesis with regard to the presence or absence of such themes. The protocols were simply examined for evidence of this phenomenon using Horner's criteria (1968). The relationship between Horner's criteria and the relevant categories in Arnold's (1962) scoring manual is given in Appendix A.

CHAPTER III

METHOD

Subjects

The subjects of this study were 60 college students, the majority of whom were freshmen. They were enrolled in the introductory course in psychology at Loyola University, Chicago, Illinois. They all volunteered to be tested and received academic credit for their participation. They were divided into four groups, with two groups composed respectively of 13 and 14 male subjects and two groups composed respectively of 18 and 15 female subjects.

Measures

The measures for testing the hypotheses in sections 1 (Reliability) and 3 (Motivational level) were the Motivation Index scores derived from stories told to two separate sets of TAT cards. The set of male stimulus cards consisted of Murray's (1943) TAT pictures identified by the following numbers: 1, 3BM, 6BM, 7BM, 8BM, 9BM, 12BM, 13BM, 18BM, 16 [blank card], and 17BM, presented in that order. The female stimulus cards consisted of Murray's TAT cards identified by the following numbers: 2, 6GF, 3GF, 7GF, 12GF, 17GF, 18GF, 13GF, 9GF, 16 [blank card], and 8GF, presented in that order. The written descriptions of each card are provided by Murray (1943) and are given in Appendix B.

The Motivation Index scores were derived according to Arnold's

(1962) criteria. An import was first abstracted from each story. Then, according to Arnold's (1962) scoring manual, each import is given an intensity score of +2, +1, -1, or -2. The algebraic sum of the intensity scores is then converted into a Motivation Index score using the transformation tables in Arnold's (1962) scoring manual.

The measures for the hypotheses in section 2 (Thematic preoccupation) were established in this way. Once the import was abstracted it was identified as belonging to one of four categories: (a) Achievement, (b) Right and Wrong, (c) Human Relationships, and (d) Reaction to Adversity. Another way of saying this is that each import was categorized according to its basic theme. The basic measure, then, was the number of themes in each category.

The measures for fear-of-success stories were those of Horner (1968) adapted to Arnold's (1962) scoring system. According to Horner (1962), fear of success is scored as present if the responses contain any of the following themes: negative consequences because of success, activities away from future success, direct expression of conflict about success, and bizarre or inappropriate responses. The relationship of these criteria to Arnold's (1962) scoring categories is given in Appendix A. The measure, then, was the number of stories indicating fear-of-success themes.

Procedure

The TAT was administered twice to each group with an interval of one week between testings.

For the first group (14 male subjects and 18 female subjects) the first testing consisted of Murray's (1943) "male" stimulus cards.

The second testing consisted of Murray's "female" stimulus cards.

For the second group (13 male subjects and 15 female subjects) the first testing consisted of Murray's (1943) "female" stimulus cards and the second testing consisted of Murray's "male" stimulus cards.

The stimulus cards were presented using slides of the pictures projected on a screen in a group setting. There were 10 stimulus cards for each testing in addition to the blank card which was placed next to the last in each presentation.

The subjects wrote the stories on lined paper (8 1/2" by 11").

Before the first picture was shown these instructions were given:

This is a test of your creative imagination. You will see eleven pictures, one after another. As you look at each picture, write as dramatic a story as you can about it. Tell what has led up to the scene shown in the picture and what is happening now. What are the thoughts and feelings of the people in the picture? What will be the outcome?

Since we are interested in your creative imagination, be sure to tell a story with a plot and an outcome. Do not just describe the picture. Try to write a story and not a piece of conversation.

You will have seven minutes for each story. Be sure to write something about each picture. If you can't think of anything for one of the pictures, write that down too.

One of the pictures is blank. When you get to this card, imagine any picture there you want, describe it briefly and again, write a story with a plot and an outcome about the picture you imagined. Are there any questions about the procedure?

Before showing the second picture, the instructions were repeated in a briefer version:

Remember that we are interested in your creative imagination, and be sure to tell a story with a plot and outcome. Tell what has led up to the scene in the picture, what is happening now and how it will end. Write a story, not a piece of conversation. Are there any questions about what you are to do?

Before showing the third picture a last reminder of the procedure was given:

Remember, we want to know what goes on in the picture, what led up to it, what is happening now, and how it will all end. Are there any questions?

For the second testing the instructions were given only before the first picture.

To preserve anonymity the following procedure was carried out. Each subject was issued a blank envelope and a 3" by 5" index card with this information: the subject's name, sex, sex denomination of stimulus cards, and the testing session number. The information card was then sealed in the blank envelope. There was no information on the protocols themselves. After the testing, the envelope was stapled to the protocol. When all the testing was completed the protocols were given to a third party to be arranged randomly. The randomized protocols were then numbered together with accompanying envelopes from 1 - 120. Then the envelopes were detached from the protocols until the scoring was completed.

The imports were derived and scored by two independent examiners. When the imports and scores were compared it was found that there was 97% agreement on the content of the imports and 76% agreement as to whether the import should be scored positively or negatively. The examiners then met to discuss the reason for each discrepant score. After the differences were resolved a final score was agreed upon. Thus the differences were resolved before any statistical analysis of the data was done. The protocols were then put in their respective groupings for the appropriate statistical analysis.

CHAPTER IV

RESULTS

Statistical analysis

Since the rating scale of the Story Sequence Analysis scoring system has no equal intervals, it cannot be treated with parametric statistical procedures (cf. Siegel, 1956). The intensity of the scored import is indicated by +2, +1, -1, -2 (Arnold, 1962). These levels of positive or negative motivation cannot be considered on a metric scale, because when the rater changes from positive to negative scoring, it is not a step equal to changing from 1 to 2 on the same side of the scale. This intrinsic inequality of steps in the rating scale limits the type of statistics that can be used. Therefore, in this analysis, nonparametric methods that depend on rank ordering were employed to test the hypotheses in sections 1 and 3.

With the hypotheses in section 1, the nonparametric statistic used was the Spearman rank correlation coefficient, rho. It is a measure of association which requires that both variables be measured in at least ordinal scale so that the individuals under study may be ranked in two ordered series.

Since the data for the hypotheses in section 2 met most of the assumptions necessary for treatment by parametric procedures, the t test was used.

To test the hypotheses in section 3, the Mann-Whitney U test was employed. This was used as an alternative to the parametric t test when

the measurement in the research was weaker than interval scaling.

The level of significance for the nonparametric tests was established at the .05 level, and .01 for the parametric tests.

General descriptive characteristics of the data

The possible Motivation Index scores range from 0 to 200 with a median of 100. A score of 0 means that the intensity scores for each import were -2. By contrast, a score of 200 indicates that the intensity scores for each import were +2. A score of 100 shows that the algebraic sum of the intensity scores for each import was 0, i.e., there was an equal number of negative and positive scores.

Arnold (personal communication, 1974), found that the actual median in early Story Sequence Analysis studies was 104.

The range and median Motivation Index score for each group and combination of groups is given in Tables 1 and 2. The medians are generally well below those found by Arnold and her collaborators in the early studies.

Reliability

The overall hypothesis was that the Motivation Index scores would be significantly reliable under alternate-form (male and female cards) conditions. The correlation coefficients for the individual hypotheses in section 1 are given in Table 3. In no case did the correlation reach statistical significance. The general hypothesis, therefore, was not supported by the data.

Thematic preoccupation

The general hypothesis was that there would be no significant differences in thematic preoccupation with regard to male and female

Table 1

Ranges and Median Motivation Index (MI) Scores

Groups	Male Stimulus Cards		Female Stimulus Cards	
	Range	MI	Range	MI
I. Males - Male cards first	150	46.0	123	57.5
II. Males - Female cards first	136	73.0	155	59.0
III. Females - Female cards first	122	82.0	118	59.5
IV. Females - Male cards first	105	82.0	127	82.0
Composite Scores for I and II	159	68.0	150	59.0
Composite Scores for III and IV	122	82.0	127	73.0
Composite Scores for I, II, III, and IV	159	73.0	150	60.0

Table 2

Ranges and Median Motivation Index (MI) Scores from Combined^a Data
for Male and Female Subjects

Subjects	Range	MI
Male	136.0	61.5
Female	103.5	75.0
Male and Female	136.0	68.0

^aThe combined data were derived by taking the averages of the two Motivation Index scores from the male and female cards respectively.

Table 3
Nonparametric Spearman Rank Correlations Coefficients
for Motivation Index Scores

Subjects	Rho
Group I. Males	
1st Testing - Male Cards	.00
2nd Testing - Female Cards	
Group II. Males	
1st Testing - Female Cards	.33
2nd Testing - Male Cards	
Group III. Females	
1st Testing - Female Cards	.05
2nd Testing - Male cards	
Group IV. Females	
1st Testing - Male Cards	.06
2nd Testing - Female Cards	
Composite Scores of Groups I and II	.21
Composite Scores for Groups III and IV	.06
Composite Scores for Groups I, II, III, and IV	.12
None of the correlation coefficients was statistically significant at the .05 level.	

subjects or stimulus variables. The individual hypotheses in Section 2 were tested with the parametric statistic t . The results were as follows:

a) Male subjects did not tell significantly more stories with achievement themes than female subjects -- $t(1, 58) = 1.62$. The probability of this is greater than .01, therefore the null hypothesis was accepted.

b) Male subjects did not tell significantly more stories with achievement themes to male cards than to female cards -- $t(1, 26) = 3.38$. The probability of this is less than .01, therefore the null hypothesis was rejected.

c) Female subjects do not tell significantly more stories with human relationship themes than male subjects -- $t(1, 58) = 2.72$. The probability of this is less than .01, therefore the null hypothesis was rejected.

d) Female subjects do not tell significantly more stories with human relationship themes to female cards than to male cards -- $t(1, 32) = 3.55$. The probability of this is less than .01, therefore the null hypothesis was rejected.

Since the null hypothesis was rejected in three out of four cases, the data can be said to show that sex variables had a significant effect on thematic preoccupation.

Motivational level

The basic hypothesis was that there would be no significant differences between male and female subjects with regard to motivational level.

The results of the individual hypotheses were as follows:

a) The Motivation Index scores derived from the combined scores of the two testings of male subjects did not differ significantly from that of female subjects. This comparison yielded a z score of 1.77 and indicated that females scored higher than males. The probability of this is less than .05, therefore the null hypothesis was rejected.

b) The Motivation Index scores of male subjects did not differ significantly from those of the female subjects when writing stories to male cards. This comparison yielded a z score of 1.24. The probability of this is greater than .05, therefore the null hypothesis was accepted.

c) The Motivation Index scores of female subjects did not differ significantly from those of male subjects when writing stories to female cards. This comparison yielded a z score of 1.05. The probability of this is greater than .05, therefore the null hypothesis was accepted.

The data, therefore, indicated the female subjects had a higher motivational level than the male subjects but that this difference did not show up on the male cards alone or the female cards alone.

Fear-of-Success Themes

There were no specific hypotheses with regard to these themes so no statistical analysis was done. As was said above, the themes were simply counted using Horner's (1968) criteria for fear of success.

To the male cards, male subjects (N = 27) wrote a total of 116 stories with achievement themes and of these, 88 (or 75%) indicated fear of success; female subjects (N = 33) wrote a total of 121 stories with achievement themes and of these, 85 (or 70%) indicated fear of success.

To the female cards, male subjects wrote a total of 74 stories

with achievement themes and of these 45 (or 60%) indicated fear of success; female subjects wrote a total of 80 stories with achievement themes and of these 43 (or 53%) indicated fear of success.

Thus, both male and female subjects wrote more fear-of-success stories to male cards. However, male subjects wrote more fear-of-success stories to both male and female cards than did the female subjects.

CHAPTER V

DISCUSSION

General descriptive characteristics of the data

One of the difficulties in assessing the meaning of the general characteristics of the data (i.e., central tendency and variability) is that the Motivation Index of Story Sequence Analysis has never been standardized. Arnold (1962) gives the reason for this lack:

To work out the Motivation Index from standard scores would take many years of research because it would involve a careful sampling not only on the basis of sex, age, intelligence, geographic distribution, socio-economic level and the like, but especially on the basis of achievement or effectiveness in life. When obtained these scores would allow us to compare the consistency found in a normative population. There is a reasonable doubt, however, whether this additional comparison would offer a worthwhile return for the labor required [pp. 148-149].

Therefore, although the mathematical median of the Motivation Index is 100, there is no way of knowing whether this represents the true central tendency of any population. Arnold (personal communication, 1976) stated, however, that in her early studies the actual median was 104. This indicates that the average individual was characterized by mixed attitudes and motives - half positive, half negative.

In recent studies with college students the picture seems to have changed. In four samples (Arnold, 1974; Smith, 1971) the median Motivation Index scores were 62, 69, 77, and 81 respectively. The over-all median was 72. In the present study the over-all median was 68. This could mean one of two things; either the "normal" median Motivation Index score is below the mathematical median of 100 or that the

college students in recent samples are poorly motivated.

Smith (1971) found that after group psychotherapy a sample of male college students improved their Motivation Index scores from a median of 77 to a median of 135. Unfortunately, no follow-up study was done to find out whether the students showed comparable improvement in scholastic achievement.

Therefore, no firm judgment can be made at this time about the general motivational level of college students because of the lack of adequate normative data. But, because of the generally low scores, it might be safe to say that many of the students in this sample, especially the males, are not functioning to their full potential.

Reliability

The discussion of the reliability of the Motivation Index covers two areas: (a) interjudge reliability and (b) reliability under male/female stimulus card conditions.

Interjudge reliability. As was said above, the stories were analyzed by two independent examiners. It was found that there was 97% in abstracting the imports. This is probably the most reliable element in the whole Story Sequence Analysis procedure. There was much more discrepancy between the examiners in assigning the intensity score for each import from which the Motivation Index is derived. There was 76% agreement as to whether the import should be scored positively or negatively.

There seem to be two reasons for the discrepancies. Before the two sets of scores were compared, the mean score of the data assigned by examiner 1 was 61.02; the average score of examiner 2 was 78.36. After the discrepancies were resolved in conference, the average Moti-

vational Index score was 67.78. It is interesting to note that this is very close to the mathematical mean of the two uncorrected scores, i.e. 69.69. Therefore, one of the reasons for the discrepancies was a consistent examiner bias or "halo" effect. Examiner 1 tended to score imports more negatively and examiner 2 tended to score them more positively. However, in the long run, the biases cancelled each other out.

It was also found that many of the discrepancies were due to inadvertence on the part of the examiners to the nuances of the scoring system and were easily resolved. The main cause for this seemed to be fatigue. Assigning the intensity scores is tedious, time consuming work so the examiner is quite vulnerable to effects of fatigue.

Very few of the discrepancies could be attributed to ambiguities in the scoring system itself except for scores in Category II, B, which deals with "bad human relationships" (e.g., quarrels, enmity). Differences in this area were most difficult to resolve but it is not certain at this point what can be done to remedy the problem.

Therefore, it is recommended that in future research two examiners always be used, not only to correct for examiner bias, but also as a check on mistakes due to inadvertence or fatigue.

Reliability of the Motivation Index under male/female stimulus conditions.

Hypothesis I, a through g stated that there would be significant correlations between Motivation Index scores across male/female stimulus variables. In other words, the male and female sets of pictures could be used as alternate-forms of the test. This was not verified by the data. Most of the correlation coefficients fluctuated around .00 and only one was as high as .33. This result was both surprising and puzzling since test-retest correlations of the Motivation Index are usually statistically

significant even though they are generally modest.

Since the correlations were consistently low, the cause cannot be attributed to sex variables (the correlations for both male and female subjects were equally low) nor to the order of testing (whether the subjects wrote the stories to male pictures first or second made no difference).

There are a number of possible explanations for this lack of reliability. First, it may have been due to motivational inconsistency on the part of the subjects themselves. The Motivation Index scores tended to be in the low, intermediate range (46 to 82). This generally means that the subjects oscillated between positive and negative motivational attitudes, neither consistently positive nor consistently negative. This does not seem to hold up, however, since one would expect this oscillation to be distributed equally on both male and female cards and, thus, would not affect the reliability.

Secondly, the inconsistency may have been caused by the discrepancies in the scoring. This does not seem to be the case either since whatever discrepancies appeared were easily resolved by the examiners.

Thirdly, it is possible that male stimulus cards tended to evoke themes in different motivational areas than female cards. As was seen from the results of Hypothesis III, the male cards did tend to elicit more achievement themes than the female cards and that the female cards tended to evoke more human relationship themes (Categories I and III respectively). This effect was true of both male and female subjects. Since motivational attitudes need not be the same in both areas (e.g., an individual could have positive attitudes towards work and achievement and negative attitudes in human relationships), it is possible that the unreliability was the result of inconsistent thematic preoccupation.

Arnold (1962) noted:

It is quite possible that a man may have positive attitudes to work and achievement.... If a set of pictures is narrowly confined to one theme, that of work and achievement, we only tap a man's motivating attitudes in this one area. But he is not influenced only by his motivation at work. Unwittingly, but nevertheless effectively, his attitudes towards others, toward right and wrong, his habitual reaction to adversity of all kinds, will influence even his workaday life. The same executives who may be highly positive in the work and achievement area may have negative attitudes in other areas of life....[pp. 202-203].

Working on this supposition, namely, that the male cards consistently tended to tap motives in the area of achievement and success and that female cards consistently tended to tap motives in the area of human relationships for both male and female subjects, the investigator did another correlation. To distribute the effects of the cards evenly he derived the Motivation Index scores in the following way. The intensity scores of the first five male cards and the first five female cards for each subject were combined to derive a Motivation Index score. Then the second five male cards and the second five female cards for each subject were combined to derive a second Motivation Index score. This yielded two scores for each subject, but in this case each one was derived from half male and half female cards. He then compared the resulting correlations to those previously done. The results are given in Table 6.

This time all the correlations (although modest as usual) were statistically significant but one. Apparently, female stimulus cards systematically elicit a preponderance of stories with human relationship themes. Male stimulus cards, even though the preponderance is not as great, tend to elicit stories with achievement themes. Therefore, on the basis of the evidence, sets of male and female cards cannot be used as alternate-forms of the same test.

Table 4

Relationship of the Correlation Coefficients Derived from the Comparison of the Male Cards and Female Cards (Rho 1) and the Correlation Coefficients Derived from the Comparison of the First Five Male and Female Cards with the Second Five Male and Female Cards (Rho 2)

Subjects	Rho 1	Rho 2
Group I. Males		
1st Testing - Male Cards	.00	.57*
Group II. Males		
1st Testing - Female Cards	.33	.68**
Group III. Females		
1st Testing - Female Cards	.05	.45*
Group IV. Females		
1st Testing - Male Cards	.06	.33
Composite Scores for Groups I and II	.21	.67***
Composite Scores for Groups III and IV	.06	.40*
Composite Scores for Groups I, II, III and IV	.12	.56***

* $p = < .05$

** $p = < .01$

*** $p = < .001$

Thematic Preoccupation

Some aspects of the results of this section have already been dealt with above. The most significant finding is that male subjects overall did not write significantly more stories with achievement themes than did female subjects, at least under the nonarousal conditions used in this study. This goes against the cultural stereotype that men are more interested in success and careers than are women. However, it was also found that both male and female subjects wrote more achievement stories to the male stimulus cards than to the female stimulus cards. This is consistent with the cultural stereotype. Thus men and women express more of their achievement needs and attitudes to the male cards than to the female cards.

Nevertheless, female subjects, as expected, maintained their interest in human relationship themes. They wrote significantly more human relationship stories, overall, than male subjects and wrote more human relationship themes to the female cards than to the male cards.

It would appear, then, that although college women were as interested in achievement and success as college men in this sample, they still retained their traditional interest in home and family. This fact may have important implications in resolving some of the conflicts that many women feel at present. The question does not seem to be merely a matter of helping women express and fulfill their desire for achievement but also providing ways of reconciling that with their attitudes toward human relationships.

Sex Differences in Achievement Motivation

The main significant difference found here was that the female

subjects had higher overall Motivation Index scores. In comparing the scores of women derived from the male cards alone to those of the men, the difference was not statistically significant but the women's scores were higher. This was likewise the case with the female cards taken alone.

It appears, then, that the women in this sample were even more positively motivated than the men. The results here are consistent with those of earlier studies (Veroff, Wilcox, & Atkinson, 1953). Under neutral conditions the scores of women are as high or higher than those of men. However, the most striking differences are not between the men and the women but between the low achievers and the high achievers. To illustrate this the investigator chose four protocols: (a) the male subject with the lowest Motivation Index score; (b) the female subject with the lowest Motivation Index score; (c) the highest male subject score; (d) the highest female subject score.

The imports and respective intensity scores of the four protocols are as follows:

a) Male Subject #20 - Motivation Index score 2.5

Male cards:

1. You're exhausted and bored with what you're doing so you decide to quit and destroy the means for doing it. [-2, I.B.1.b.]
2. A person that can't cope with the world will do themselves in. [-2, IV.A.5.b.]
3. If someone close to you is done in by accident you'll take it out on those responsible. [-2, IV.A.5.b.]
4. You get good advice from an elder but disregard it, leave and take a menial job. [-2, III.C.3.b]
5. After watching a professional in action you decide what he does is gross and switch to another type of study. [-2, I.A.5.b.]
6. If a group of pushers find drugs they try them themselves then divide them and go their own ways. [-2, II.A.1.a.]

7. If you pass out from drinking and someone tries to help you they'll be unsuccessful. [-2, III.C.3.b.]
8. Your parents are done in by others; you're captured and raised by those who did in your parents. [-2, IV.A.2.a.]
9. A person who drinks too much will pass out and be thrown into the gutter. [-2, III.F.2.c.]
10. No story.
11. If you're escaping from the police, a beautiful blond will offer you her favors and you'll have a great time. [-2. II. A.1.a.]

Female cards:

1. You try to convince somewhat that they're doing something wrong but they walk away and you're depressed about it. [-2, III.E.1.a.]
2. Someone close to you is killed by a certain thing; you're upset and decide never to use the thing again. [-1, IV.5.a.]
3. If you struggle against an attack it's to no avail and you're done in. [-2, IV.A.4.a.]
4. At first you're bored and embarrassed by sex information but later you decide that sex is interesting and you're going to experiment with it. [-2, II.A.3.a.]
5. If someone constantly puts you down, you'll do them in and get away with it. [-2, II.A.1.a.]
6. You're ill and have to struggle to get to where you want to go; you almost make it but the effort does you in. [-2, IV.A.4.b.]
7. If someone stands in your way you'll do them in and get what you want. [-2, II.A.1.a.]
8. People are frightened by a natural phenomenon and in the panic that ensues, half of them die and the place where they lived is deserted. [-2, IV.A.5.b.]
9. If you're envious of someone you'll do them in even if they try to escape. [-2, III.B.3.b.]
10. If someone is cute you go after him even if he's nuts. [-2, III.A.1.f.]
11. All of a sudden your wishes come true. [-2, I.B.1.a.]

b. Female Subject #16 - Motivation Index score 25

Male cards:

1. An elder wants you to be great and threatens you with deprivation of goodies if you don't practice, but you're tired so you give up the goodies to avoid the work. [-2, I.A.5.b.]
2. If people torment you the only way to block out the confusion and anxiety is to escape through drugs but this will eventually do you in. [-2, IV.A.3.a.]
3. Homosexual problems will cause you to be rejected by your parents. [-2, III.B.1.d.]

4. An old person, depressed by the memories of lost youth will do themselves in. [-2, IV.A.5.b.]
5. Even though you escape one danger you'll fall into another. [-2, IV.A.4.d.]
6. Lying around dreaming of what you want to do will not get anything done; you've got to go out and do it. [-1, I.B.5.a.]
7. If you think you're strange and dangerous and someone close to you tries to rid you of that thought, you'll end up hurting them and regretting it. [-2, III.E.3.c.]
8. You've run away from home but since you didn't make provisions, when you get hungry you have to return. [-1, III.C.3.e.]
9. If you get hurt by someone and complain to the authorities, they just scold you and do nothing about it. [-1, III.C.3.d.]
10. You imagine yourself after death to be in a state of being surrounded by nothing and unable to see anything. [-2, 1.A.7.]
11. Even professionals can't figure out what's wrong with you so they let you go to get employment that suits your state of mind. [-1, I.D.3.a.]

Female cards:

1. You're disgusted at someone's failure; it's more than you can take so you tell them you'll discuss it later. [-1, III.F.1.d.]
2. After legitimate confrontation from an elder you feel the world is cold and cruel so you leave the scene. [-2, III.F.2.c.]
3. If someone bothers you, you'll tell them to leave you alone. [-1, III.F.1.d.]
4. You are more interested in sex than in sex information and when your elders have finished talking to you you'll go out and experiment. [-1, III.C.2.a.]
5. At the very time you're doing something good for someone, you lose them. [-2, III.E.3.c.]
6. When you discover the results of violence, elders tell you you're only dreaming. [-1, III.E.1.a.]
7. Your elders want you to marry for wealth but you'll marry for beauty. [-2, III.A.1.f.]
8. You're tired of going through all these changes and decide to do yourself in. [-2, IV.A.5.b.]
9. When you think you're being saved from danger you'll end up in more danger. [-2, IV.A.4.b.]
10. You're tired from partying and if you're awakened accidentally you'll go back to bed. [-1, I.B.7.b.]
11. You think about getting away from your present drudgery but it'll continue till you die. [-2, 1.A.2.c.]

c. Male Subject #26 - Motivation Index score 138.5

Male cards:

1. After years of work and practice you'll reach your goal. [+2, I.B.1.a.]
2. If you get professional help for you illness you'll recuperate. [+1, IV.A.1.f.]
3. If you're cruel to someone you'll regret it but since it's too late to undo the damage you'll be brokenhearted. [-2 III.A.2.c.]
4. An elder is angry with you because you won't accompany him to visit a relative; you convince him to go himself, help him with the fare, and he has a happy reunion. [+1, III.A.3.a.]
5. Your fear sometimes causes you to run from people who mean no harm. [-1, IV.A.5.a.i.]
6. But if someone convinces you that your fears are groundless, you'll realize that others are not as bad as you thought they were. [+1, III.C.2.b.]
7. You leave a person who is ill after a moment of intimacy not realizing it'll be the last one. [+1, IV.A.2.a.]
8. You're anxious to start school so as to get out into the world of people. [+1, I.A.4.]
9. You're annoyed with someone who startles you but when you find out it's a friend, you get reacquainted. [+1, III.A.1.a.]
10. You wonder what it will be like when you're old but you realize this is foolish; you feel that you must live your life to the fullest now. [+1, III.F.1.b.]
11. If you can keep you mind off the danger of a situation, you'll perform smoothly and later be grateful that you're safe. [+2, IV.A.1.a.]

Female cards:

1. When your work is profitable, you're happy that your efforts weren't wasted. [+2, I.A.4.a.]
2. After the death of a loved one you feel you can't go on but an elder convinces you that you can so you do make the best of your life. [+1, IV.A.2.e.]
3. People surprise you sometimes with good things you didn't expect. [+1, III.F.1.a.]
4. You're worried about someone who is away but they return and you're happy and thankful. [-1, III.A.4.d.]
5. If you find someone in trouble you'll help them. [+2, III.E.2.a.]
6. You're reprimanded by your elders for something wrong; they explain why and you understand and agree to change. [+1, III.C.3.b.]
7. If someone tries to tempt you to do wrong you won't be swayed and convince them it's not the right thing to do. [+2, II.B.4.a.]

8. You're sometimes astonished at a stroke of good fortune. [-1, I.B.1.a.]
9. You meet an old friend and get reacquainted. [+1, III.F.5.d.]
10. To prepare for marriage a couple has to try to understand each other and love each other and help each other through thick and thin. [+2, III.A.3.a.]
11. If you work hard in school, you'll get good grades. [+2, I.B.1.a.]

d. Female Subject #22 - Motivation Index score 128.5

Male cards:

1. You don't like what you're doing but will keep doing it and eventually enjoy it and do well even though you won't be great. [+1, I.A.1.a.]
2. A woman will at first be depressed by academic failure but will eventually get over it and continue her work and get a Ph.D. [+1, I.A.1.b.]
3. When a son leaves his mother she'll be sad but she'll help him go out on his own and she'll do something useful and be happy with it. [+1, III.C.4.a.]
4. The mutual problems and sadness in life will bring a father and son even closer together. [+2, III.A.3.c.]
5. If someone misses out on the good things in life as a child, he'll end up bitter and hateful. [-2, III.F.4.c.]
6. You've been helped by people whose life-style you don't want to share, so you go on and succeed at your chosen profession. [+2, I.D.1.d.]
7. You've been ill but now you're better and the people around you have been sensitive to your needs. [+2, III.A.1.a.]
8. If you lose someone close you'll be sad but others will comfort you. [+1, III.E.1.a.]
9. If a psychotic does himself in it's because others have forced him. [-2, III.C.2.b.]
10. Since there are people who love you and whom you love, you are happy even when things are dreary on the outside. [+1, III.A.1.d.]

Female cards:

1. Even though you feel trapped by lack of education you can encourage others to get it. [+1, III.E.2.b.]
2. You're inexperienced so if someone makes a pass at you, you'll be frightened; this will make a lasting mark on you but you'll eventually marry your fiance. [+1, IV.A.2.c.]
3. You mourn the loss of someone close and even though you'll overcome your grief you'll still miss them. [+1, IV.A.2.c.]
4. If someone tries to teach you something you already know you won't be interested and you'll dream. [-1, I.B.1.b.]

5. If a person presents a good facade to the world what craziness is inside the person will eventually take over. [-1, III.F.1.d.]
6. Material things and worries have become your whole world; you think of doing yourself in but decide to leave your present life and get the perspective back by giving up all that you've made. [+1, IV.A.2.e.]
7. You're sad when someone close to you dies but since you're strong you'll make a life of your own. [+1, IV.A.2.e.]
8. You're just beginning your career and even though you're afraid to make mistakes this feeling of uneasiness will disappear once you get started and you'll perform well and become famous. [+1, IV.A.2.e.]
9. When you help someone, others will join in helping you. [+2, III.E.2.b.]
10. Even though natural forces are sometimes destructive you're awed by their force and majesty. [+1, III.F.1.a.]
11. You're discouraged in teaching others and wonder why you do it but you keep thinking of ways to excite the students and you know you'll continue. [+1, III.E.1.b.]

One can see from even a cursory reading of these imports that the main differences show up not between men and women but between the poorly motivated individuals and those that are well motivated. The subjects with the high scores show many similar attitudes: if you work hard difficulties can be overcome whether in school or outside of it; relationships between people are generally good and when problems arise they can be dealt with and resolved; in situations of loss and danger something can be done to assuage the pain and overcome the danger.

The subjects with the low scores also showed many similarities: work is distasteful, a bore, to be gotten out of; human relationships are haphazard, destructive, a cause of grief, anxiety and depression; authorities can't be trusted; trying to overcome loss or danger just leads to more trouble.

Therefore positively motivated men and women tend to be more like each other than either of them are to their poorly motivated counterparts.

The only difference with regard to achievement that one notices is that the male subject sees it as coming about with less frustration: "After years of work and practice you'll reach your goal; if you work hard in school you'll get good grades." The female subject sees herself successful but only after setbacks and failure. "A woman will at first be depressed by academic failure but eventually gets over it and continues her work and gets a Ph.D.; Even though you feel trapped by lack of education you can encourage others to get it."

This attitude was common in the protocols of the women--you can succeed academically and professionally but it will be against the odds.

Fear-of-Success Themes

The findings with regard to this problem were striking. Male subjects wrote more fear-of-success stories to both the male and female cards than did the female subjects. In addition, the percentage of fear-of-success stories was higher for the male cards than for the female cards for both sexes. This is puzzling since it goes against Horner's (1968) original findings.

Feather and Raphelson (1974) found that the proportion of fear-of-success stories written by both male and females to the John cue was greater. This study supports their findings in that more of the fear-of-success stories were written to the male cards. What is surprising is that the male subjects seem to fear success more than female subjects. It may be that some of the cultural stereotypes are breaking down so that women now feel more competent to succeed in academic and professional life. This may be having a kind of backlash effect in that the more women succeed in areas traditionally dominated by men the more insecure men become.

The phenomenon certainly deserves closer scrutiny.

Zuckerman and Wheeler (1975) reviewed Horner's (1968) study and subsequent research using the fantasy-based measure of fear of success and pointed out that the findings obtained by Horner are not as robust as originally assumed. The writer would agree with their conclusions.

Considerations for Future Research

From the results of this study it appears that the problem of the reliability of the Motivation Index of the TAT is by no means solved. It would be interesting to find out what the correlation coefficients would be between two administrations of female cards alone, male cards alone and a mixed set of male and female cards together. This could be done for both male and female subjects.

There is a dearth of data on reliability for populations other than college students. Most college students are still in the adolescent period, a period not generally associated with consistency and stability. Is it possible that the protocols of older, more settled individuals would produce more reliable results?

It is not certain whether the Motivation Index scores are still safe predictors of actual academic achievement. The validity studies done two decades ago may not be reliable guides for what is happening academically at the present time.

It would also be interesting to investigate further why the achievement desires and attitudes of women are elicited more readily by the male stimulus cards than by the female stimulus cards.

SUMMARY

The main goal of this study was to investigate the alternate-form reliability of the Motivation Index of Story Sequence Analysis for male and female subjects using a set of male stimulus cards as one form of the test and a set of female cards as the other.

The next task was to examine the relative effect of the sex of the subject and stimulus conditions on the thematic preoccupations evidenced in the stories.

The third aim was to investigate the differences in motivational level between male and female subjects as measured by the Motivation Index.

The last objective was exploratory rather than correlational or experimental. The data were examined for corroboratory evidence of Horner's (1968) hypothesis that women write more fear-of-success stories than men.

The subjects were 60 college students, most of whom were freshmen. There were 27 male subjects and 33 female subjects.

They were divided into four groups, with two groups composed respectively of 13 and 14 male subjects and two groups composed of 18 and 15 female subjects.

The TAT was administered twice to each group with an interval of one week between testing.

For the first group (14 male subjects and 18 female subjects) the first testing consisted of Murray's (1943) "male stimulus cards. The second testing consistent of Murray's "female" stimulus cards.

For the second group (13 male subjects and 15 female subjects) the order was reversed.

The Motivation Index scores were derived according to Arnold's (1962) criteria. These were the measures used to test reliability and motivational level.

To assess thematic preoccupation each import was categorized according to its basic theme. The measure, then, was the number of themes in each category.

The measures for fear-of-success stories were those of Horner (1968) adapted to Arnold's (1962) scoring system.

The following conclusions were drawn from the evidence presented in this study.

1. The female stimulus cards cannot be considered an alternate or equivalent form of the TAT when used in conjunction with the male stimulus cards.
2. Male stimulus cards tended to elicit significantly more achievement themes than did the female cards for both male and female subjects.
3. Male subjects did not write significantly more achievement themes than female subjects.
4. Female subjects, on the other hand, did write significantly more stories with human relationship themes than did the male subjects.
5. Female subjects obtained significantly higher Motivation Index scores than male subjects. This was true only of the combined scores derived from both male and female cards. When the scores derived from the male cards taken alone or the female cards taken alone the difference between male and female subjects was not significant, but the female subjects scored higher in both instances.

6. The phenomenon of fear-of-success themes, originally thought to be primarily characteristic of women, was not found to be so. In this study it was the men who told more fear-of-success stories.

The study did not deal directly with the interjudge reliability but these findings were of interest. The interjudge reliability with regard to the formulation of the imports of stories according to Arnold's (1962) method was very good (97% agreement between two examiners). The interjudge reliability with regard to assigning the intensity score to the imports was not as satisfactory but could be resolved by a thorough analysis of the discrepant scores by the examiner. Therefore it is recommended that two examiners always be used in assigning the intensity scores.

REFERENCES

- Abel, T.M. Responses of the negro and white morons to the Thematic Apperception Test. American Journal of Mental Deficiency, 1945 49, 463-468.
- Anastasi, A. Principles of psychological testing. New York: The Macmillan Company, 1968.
- Angelini, A.L. Un novo methodo para avaliar a motivacao humano. Unpublished doctoral dissertation, Universidade de Sao Paulo, Brazil, 1955. (Results summarized in Atkinson, 1958).
- Arnold, M.B. Emotion and personality, 2 Vols. New York: Columbia University Press, 1960.
- Arnold, M.B. Story sequence analysis. New York: Columbia U. Press, 1962.
- Arnold, M.B. Motives as causes. Journal of Phenomenological Psychology, 1971, 1, 185-192.
- Arnold, R.S. "Faking good" on the Thematic Apperception Test. Unpublished master's thesis, Loyola University, Chicago, 1974.
- Aron, B. A manual for analysis of the Thematic Apperception Test: a method and technique for personality research. Berkeley, Calif.: Willis E. Berg, 1949.
- Atkinson, J.W., & Feather, N.T. A theory of achievement motivation. New York: Wiley, 1966.
- Atkinson, J.W., & Litvin, G.H. Achievement motive and test anxiety con-

ceived as motive to approach success and motive to avoid failure.

Journal of Abnormal and Social Psychology, 1960, 60, 52-63.

Balken, E.R., & Masserman, J.H. The language of fantasy: III. The language of the fantasies of patients with conversion hysteria, anxiety state, and obsessive-compulsive neuroses. Journal of Psychology, 1940, 10, 75-86.

Balken, E.R., & Masserman, J.H. A delineation of schizophrenic language and thought in a test of imagination. Journal of Psychology, 1943, 16, 239-271.

Bardwick, J.M. The psychology of women. New York: Harper, 1971.

Bell, J.E. Projective techniques. Longmans, Green and Company, 1948.

Bellak, L. The concept of projection: an experimental investigation and study of the concept. Psychiatry, 1944, 7, 355-370.

Bellak, L. The Thematic Apperception Test in clinical use. In L.E. Abt and L. Bellak (Eds.), Projective psychology: clinical approaches to the total personality. New York: Knopf, 1950, Pp. 185-229.

Bellak, L., & Ort, E. The Thematic Apperception Test and other apperceptive methods. In D. Brower and L.E. Abt (eds.), Progress in clinical psychology, Vol. I. New York: Grune and Stratton, 1952. Pp. 149-172.

Bennett, G. Some factors related to substitute value of activities in normal and schizophrenic persons: a technique for investigation of central areas of the personality. Character and Personality 1941, 10, 42-50.

Berman, P.M. A method for diagnosing social relations from the Thematic Apperception Test. Unpublished master's thesis, Radcliffe College, 1943.

Bernstein, L. The examiner as an inhibiting factor in clinical testing.

Journal of Consulting Psychology, 1956, 20, 287-290.

Bialick, I. The relationship between reactions to authority figures on the TAT and overt behavior in an authority situation by hospital patients. Unpublished doctoral dissertation, University of Pittsburgh, 1951.

Bijou, S.W., & Kenny, D.T. The ambiguity value of TAT cards.

Journal of Consulting Psychology, 1951, 15, 203-209.

Binet, A., & Simon, T. Application des methodes nouvelle au diagnostic du niveau intellectuel chez des enfants normaux d'hospice et d'ecole primaire. Annee Psychologique, 1905, 11, 245-336.

Birney, R.C. The reliability of the achievement motive. Journal of Abnormal and Social Psychology, 1959, 58, 266-267.

Bradley, M.O. The test-retest reliability of the Thematic Apperception Test. Unpublished master's thesis, Fordham University, 1957.

Brittain, H.L. A study in imagination. The Pedagogical Seminary, 1907, 14, 137-207.

Brown, J.E. Personality dynamics of high and low academic achievers in high school: a modified Thematic Apperception Test reliability study. Unpublished master's thesis, St. Louis University, 1953.

Burkard, M.I. Characteristic differences, determined by TAT sequential analysis, between teachers rated by their pupils at the extremes in teaching efficiency. Unpublished doctoral dissertation, Loyola University, Chicago, 1958.

Buros, O.K. (Ed.) The third mental measurements yearbook. New Brunswick, N.J.: Rutgers University Press, 1949.

- Buros, O.K. (Ed.) The fourth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1953.
- Buros, O.K. (Ed.) The fifth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1959.
- Buros, O.K. (Ed.) The sixth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1965.
- Calogeras, R.C. Some relationships between fantasy and self-report behavior. Genetic Psychology Monographs, 1958, 58, 273-325.
- Child, I.L., Frank, Kitty, F., & Storm, T. Self-ratings and TAT: their relations to each other and childhood background. Journal of Personality, 1956, 25, 96-114.
- Clark, R.M. A method of administering and evaluating the Thematic Apperception Test. Genetic Psychology Monographs, 1944, 30, 3-55.
- Coleman, W. The Thematic Apperception Test. I. Effects of recent experience. II. Some quantitative observations. Journal of Clinical Psychology, 1947, 3, 257-264.
- Cox, B., & Sargent, H.D. The common responses of normal children to ten pictures of the Thematic Apperception Test series. American Psychologist, 1948, 3, 363.
- Ericson, M. A study of the Thematic Apperception Test as applied to a group of disturbed children. American Psychologist, 1947, 2, 272.
- Eron, L.D. Frequencies of themes and identifications in the stories of schizophrenic patients and non-hospitalized college students. Journal of Consulting Psychology, 1948, 12, 387-395.
- Eron, L.D. A normative study of the TAT. Psychological Monographs, 1950, 64, No. 9 (Whole No. 315).

- Eron, L.D. Chapter Seven. In Shneidman, E.S., et al., (Eds.) Thematic test analysis. Grune and Stratton, 1951.
- Eron, L.D. Responses of women to the Thematic Apperception Test. Journal of Consulting Psychology, 1953, 17, 269-282.
- Eron, L.D. Thematic Apperception Test. In O.K. Buros (Ed.) Fifth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1959. Pp. 306-310.
- Eron, L.D., Terry, D., & Callahan, R. The use of rating scales for emotional tone of TAT stories. Journal of Consulting Psychology, 1950, 14, 473-478.
- Eron, L.D., & Ritter, A.M. A comparison of two methods of administration of the Thematic Apperception Test. Journal of Consulting Psychology, 1951, 15, 55-61.
- Fagot, H.J. Personal communication, 1961. (Cited in M.B. Arnold, Story sequence analysis, 1962, p. 149).
- Feather, N.T., & Simon, J.G. Fear of success and causal attribution of outcome. Journal of Personality, 1973, 41, 525-542.
- Feather, N.T., & Raphelson, A.C. Fear of success in Australian and American student groups: Motive or sex-role stereotype? Journal of Personality, 1974, 42, 190-201.
- Field, W.F. The effects of thematic apperception on certain experimentally aroused needs. Unpublished doctoral dissertation, University of Maryland, 1951.
- Fisher, R.L. The effect of a disturbing situation upon the stability of various projective tests. Psychological Monographs, 1958, 72, No. 14 (Whole No. 467).

- Foulds, G. A method of scoring the T.A.T. applied to psychometrics. Journal of Mental Science, 1953, 99, 235-246.
- French, E.G., & Lesser, G.S. Some characteristics of the achievement motive in women. Journal of Abnormal and Social Psychology, 1964, 68, 119-128.
- Garfield, S.L., Blek, L., & Melker, F. The influence of method of administration and sex differences on selected aspects of TAT stories. Journal of Consulting Psychology, 1952, 16, 140-144.
- Garvin, J.A. A Thematic Apperception Study of non-intellective factors related to academic success on the college level. Unpublished doctoral dissertation, Loyola University, Chicago, 1960.
- Greenwald, A.F. Affective complexity and psychotherapy. Journal of Projective Techniques, 1959, 23, 429-435.
- Gurel, L., & Ullman, L.P. Quantitative differences in response to TAT cards: the relationship between transcendence score and number of emotional words. Journal of Projective Techniques, 1958, 22, 399-401.
- Haber, R.N., & Alpert, R. The role of situation and picture cues in projective measurement of the achievement motive. In J.W. Atkinson (Ed.) Motives in fantasy, action, and society. Princeton, N.J.: Van nostrand, 1958. Pp. 644-663.
- Harrison, R. Studies in the use and validity of the Thematic Apperception Test with mentally disordered patients. II. A quantitative validity study. III. Validation by the method of "blind analysis." Character and Personality, 1940, 9, 122-138.
- Harrison, R. Thematic apperceptive methods. In B.B. Wolman (Ed.) Handbook

- of clinical psychology. New York: McGraw-Hill, 1965. Pp. 562-620.
- Hartman, A.A. Chapter Nine. In Shneidman, E.S., et al. (Eds.) Thematic test analysis. New York: Grune and Stratton, 1951.
- Henry, W.E. The thematic apperception technique in the study of culture-personality relations. Genetic Psychology Monographs, 1947, 35, 1-134.
- Henry, W.E., & Farley, J. The validity of the Thematic Apperception Test in the study of the adolescent personality. Psychological Monographs, 1959, 73 (Whole No. 487).
- Henry, W.E. Discussion of Dr. Veroff's paper. In: J. Kagan and G.S. Lesser, (Eds.) Contemporary issues in Thematic apperceptive methods. Springfield, IL: Thomas, 1961.
- Horner, M.S. Sex differences in achievement motivation and performance in competitive and non-competitive situations. Unpublished doctoral dissertation, University of Michigan, 1968.
- Horner, M.S. Femininity and successful achievement: A basic inconsistency. In J.M. Bardwick, et al. (Eds.) Feminine personality and conflict, Belmont, Calif.: Wadsworth, 1970.
- Horner, M.S. Toward an understanding of achievement-related conflicts in women. Journal of Social Issues, 1972, 28, 157-175.
- Jensen, A. The reliability of projective techniques: review of the literature. Acta Psychologica, 1959, 16 (1), 108-136.
- Kagan, J. The stability of TAT fantasy and stimulus ambiguity. Journal of Consulting Psychology, 1959, 23, 266-271.
- Kagan, J. Thematic apperceptive techniques with children. In A.I. Rabin and M.R. Haworth (Eds.) Projective techniques with children.

New York: Grune and Stratton, 1960. Pp. 105-148.

Kagan, J., & Lesser, G.S. Contemporary issues in thematic apperceptive methods. Springfield, Ill.: Thomas, 1961.

Kannenberg, K.M. A comparison of results obtained from the Thematic Apperception Test under two conditions of administration. American Psychologist, 1948, 3, 363.

Krumboltz, J.D., & Farquar, W.W. Reliability and validity of the n achievement test. Journal of Consulting Psychology, 1957, 21, 226-227.

Kutash, S.B. Performance of psychopathic defective criminals on the Thematic Apperception Test. Journal of Criminal Psychopathology, 1943, 5, 319-340.

Lasaga y Travieso, J.I., & Martinez-Arango, C. Some suggestions concerning the administration and interpretation of the TAT. Journal of Psychology, 1946, 22, 117-163.

Lesser, G.S. The relationship between covert and fantasy aggression as a function of maternal response to aggression. Journal of Abnormal and Social Psychology, 1957, 55, 218-221.

Lesser, G.S., Krawitz, R., & Packard, R. Experimental arousal of achievement motivation in adolescent girls. Journal of Abnormal and Social Psychology, 1963, 66, 59-66.

Libby, W. The imagination of adolescents. American Journal of Psychology, 1908, 19, 249-252.

Liccione, J.V. The changing family relationships of adolescent girls. Journal of Abnormal and Social Psychology, 1955, 51, 421-426.

- Lindzey, G., & Goldberg, M. Motivational differences between males and females as measured by the Thematic Apperception Test. Journal of Personality, 1953, 22, 101-117.
- Lindzey, G., & Heinemann, S.H. Thematic Apperception Test: individual and group administration. Journal of Personality, 1955, 24, 34-55.
- Lindzey, G., & Herman, P.S. TAT: a note on reliability and situational validity. Journal of Projective Techniques, 1955, 18, 36-42.
- Lindzey, G., & Silverman, M. Thematic Apperception Test: techniques of group administration, sex differences, and the role of verbal productivity. Journal of Personality, 1959, 27, 311-323.
- Lipinski, B.G. Sex-role conflict and achievement motivation in college women. Unpublished doctoral dissertation, University of Cincinnati, 1965.
- Little, K.B. & Shneidman, E.S. Congruences among interpretations of psychological test and amnesic data. Psychological Monographs, 1959, 73 (Whole No. 476).
- Lowe, W.F. Effect of controlling the immediate environment of responses to the Thematic Apperception Test. Perceptual and Motor Skills, 1952, 4, 98.
- Lubin, B. Some effects and stimulus properties on TAT stories. Journal of Projective Techniques, 1961, 24, 11-16.
- Lubin, N.M. The effect of color in the TAT on productions of mentally retarded subjects. American Journal of Mental Deficiency, 1955, 60, 366-370.
- Mason, B.B. An experimental investigation of repetition and variation in administration upon the Thematic Apperception Test. Perceptual and Motor Skills, 1952, 4, 98.

- Mayman, M. Review of the literature of the TAT. In D. Rappaport, Diagnostic psychological testing: the theory, statistical evaluation and diagnostic application of a battery of tests. Vol. 2. Chicago: Year Book Medical Publishers, 1946. Pp. 496-506.
- McCandlish, L.A. An investigation of a new method of T.A.T. analysis. Unpublished doctoral dissertation, Loyola University, Chicago, 1958.
- McClelland, D.C., Atkinson, J.W., Clark, R.A., & Lowell, E.L. The achievement motive. New York: Appelton-Century-Crofts, 1953.
- McDowell, J.V. Developmental aspects of phantasy production on the Thematic Apperception Test. Unpublished doctoral dissertation, Ohio State University, 1952.
- McIntyre, C.J. Sex, age, and iconicity as factors in projective film tests. Journal of Consulting Psychology, 1954, 18, 337-343.
- Morgan, C.D., & Murray, H.A. A method for investigating fantasy. The Thematic Apperception Test. Archives of Neurology and Psychiatry, 1935, 34, 289-306.
- Morgan, H.H. Measuring achievement motivation with "picture interpretations." Journal of Consulting Psychology, 1953, 17, 289-292.
- Murray, H.A. Explorations in personality. New York: Oxford University Press, 1938.
- Murray, H.A. Thematic Apperception Test Manual. Cambridge: Harvard University Press, 1943.
- Murray, H.A., & Stein, M. Note on the selection of combat officers. Psychosomatic Medicine, 1943, 5, 386-391.
- Murray, H.A. Thematic Apperception Test. In A. Weider (Ed.), Contributions

toward medical psychology: theory and diagnostic methods. Vol. 2.

New York: Ronald, 1953. Pp. 636-649.

Murstein, B.I. Nonprojective determinants of perception on the TAT.

Journal of Consulting Psychology, 1958, 22, 195-198.

Murstein, B.I. A conceptual model of projective techniques applied to

stimulus variables with thematic techniques. Journal of Consulting Psychology, 1959, 23, 3-14.

Murstein, B.I. The role of the stimulus in the manifestation of fantasy.

In J. Kagan and G.S. Lesser (Eds.), Contemporary issues in thematic apperceptive methods. Springfield, Ill.: Thomas, 1961. Pp. 229-273.

Myers, R.L. An analysis of sex differences in verbalization and content of responses to the Rorschach and to the Thematic Apperception Test.

Unpublished doctoral dissertation, Temple University, Philadelphia, 1958.

Newbigging, P.L. Influences of a stimulus variable on stories told to

certain TAT pictures. Canadian Journal of Psychology, 1955, 9, 195-206.

Oppenheimer, F. Pamela: a case study in status symbols. Journal of Abnormal and Social Psychology, 1945, 40, 187-194.

Petrauskas, F.B. A TAT and Picture-Frustration study of naval offenders

and non-offenders. Unpublished doctoral dissertation, Loyola University, Chicago, 1958.

Quinn, T.L. Differences in motivational patterns of college student brothers as revealed in the TAT, the ratings of their peers and the ratings of their superiors: a validation study. Unpublished doctoral dissertation, Loyola University, Chicago, 1961.

- Rapaport, D. The clinical interpretation of the Thematic Apperception Test. Bulletin of the Menninger Clinic, 1943, 7, 106-113.
- Rautman, A.L., & Brower, E. War themes in children's stories. Journal of Psychology, 1945, 19, 191-202.
- Richardson, L.H. The personality of stutterers. Psychological Monographs, 1944, 56, 1-44.
- Roquebrune, G. Aspects genetique et typologiques des resultats obtenus a une epreuve projective. Enfance, 1959, No. 1, 29-47. (Psychological Abstracts, 34, 4352).
- Rosen, J.L., & Neugarten, B.L. Ego functions in the middle and late years: a thematic apperception study of normal adults. Journal of Gerontology, 1960, 51, 62-67.
- Rosenzweid, S., & Sarason, S.B. Fantasy in personality and its study by test procedures. Journal of Abnormal and Social Psychology, 1942, 37, 1012.
- Sanford, R.N. Some quantitative results from the analysis of children's stories. Psychological Bulletin, 1941, 38, 749.
- Sanford, R.N. Thematic Apperception Test. In R.N. Sanford, et. al., Physique, personality and scholarship. Monograph Soc. Res. Child Development, 1943, 8, No. 1, (Ser. No. 34), 258-301.
- Sarason, B.R., & Sarason, I.G. The effect of type of administration and sex of subject on emotional tone and outcome ratings of TAT stories. Journal of Projective Techniques, 1958, 22, 333-337.
- Sarason, S.B. The use of the Thematic Apperception Test with mentally deficient children. I. A study of high grade girls. American Journal of Mental Deficiency, 1943, 47, 414-421. (a)

- Sarason, S.B. The use of the Thematic Apperception Test with mentally deficient children. II. A study of high grade boys. American Journal of Mental Deficiency, 1943, 48, 169-173. (b)
- Schwartz, L.A. Social-situation pictures in the psychiatric interview. American Journal of Orthopsychiatry, 1932, 2, 124-133.
- Sen, A. A preliminary study of the Thematic Apperception Test. British Journal of Statistical Psychology, 1953, 6, 91-100.
- Shneidman, E.S., Joel, W., & Little, K.B. Thematic test analysis. New York: Grune and Stratton, 1951.
- Shneidman, E.S., & Farberow, N.L. TAT heroes of suicidal and nonsuicidal subjects. Journal of Projective Techniques, 1958, 22, 211-228.
- Siegel, S. Non-parametric statistics. New York: McGraw-Hill, 1956.
- Smith, W.B. Evaluating the effects of group marathon sessions using the MMPI, 16PF, and TAT. Unpublished master's thesis, Loyola University, Chicago, 1972.
- Snider, L.B. A research method validating self-determination. In M.B. Arnold and J.A. Gasson (Eds.), The human person. New York: Ronald, 1954, Pp. 222-263.
- Standards for educational and psychological tests and manuals. Washington, D.C.: American Psychological Association, 1966.
- Starr, S. The relationship between hostility-ambiguity of the TAT cards, hostility fantasy and hostile behavior. Unpublished doctoral dissertation, Washington State University, 1961.
- Steggert, F.X. An analysis of some personal and executive characteristics of participants in a university program of executive development for federal personnel. Unpublished doctoral dissertation, Loyola University, Chicago, 1961.

- Sumerwill, H.C., Campbell, M.M., & Sarason, I.G. The effect of differential motivating instructions on the motivational tone and outcome of TAT stories. Journal of Consulting Psychology, 1958, 22, 385-388.
- Symonds, P.M., & Jensen, A.R. From adolescent to adult. New York: Columbia University Press, 1961.
- Terry, D. The use of a rating scale of level of response in TAT stories. Journal of Abnormal and Social Psychology, 1952, 47, 507-511.
- Tomkins, S.S. Limits of material obtainable in the single case study by daily administration of the TAT. Psychological Bulletin, 1942, 39, 490.
- Tomkins, S.S. The Thematic Apperception Test. Grune and Stratton, 1947.
- Turner, G.C., & Coleman, J.C. Examiner influence on Thematic Apperception Test responses. Journal of Projective Techniques, 1962, 26, 478-486.
- Tyrell, D.J. When love is lost. Waco, Texas: Word Books, 1972.
- Ullman, L.P. Productivity and the clinical use of TAT cards. Journal of Projective Techniques, 1957, 21, 399-403.
- Van Lennep, D.J. The Four-Picture Test. In H.H. Anderson and G. Anderson (Eds.), An introduction to projective techniques. Englewood Cliffs, N.J.: Prentice-Hall, 1951, Pp. 149-180.
- Van Lennep, D.J. The Four-Picture Test. The Hague, Netherlands: Martinus Nijhoff, 1948.
- Veroff, J. Thematic apperception in a nationwide survey. In: J. Kagan and G.S. Lesser (Eds.), Contemporary Issues in thematic apperceptive methods Springfield, IL: Thomas, 1961.
- Veroff, J., Atkinson, J.W., Feld, S., & Gurin, G. The use of thematic

apperception to assess motivation in a nationwide interview study.

Psychological Monographs, 1960, 74 (Whole No. 499).

Veroff, J., Wilvox, S., & Atkinson, J. The achievement motive in high school and college-age women. Journal of Abnormal and Social Psychology, 1953, 48, 103-119.

Watson, R.I. The Thematic Apperception Test. In R.I. Watson, The clinical method in psychology, New York: Harper and Row, 1951, Pp. 436-523.

Webster, H. Derivation and use of the masculinity-femininity variable. Journal of Clinical Psychology, 1953, 9, 33-36.

Weisskopf, E.A. A transcendence index as a proposed measure in the TAT. Journal of Psychology, 1950, 29, 379-390.

Weisskopf, E.A., & Dunleavy, G.P. Bodily similarity between subject and central figure in the TAT as an influence on projection. Journal of Abnormal and Social Psychology, 1952, 47, 441-445.

Weisskopf, E.A., & Money, C. Facial similarity between subject and central figure in the TAT as an influence on projection. Journal of Abnormal and Social Psychology, 1953, 48, 341-344.

Whitehouse, E. Norms for certain aspects of the Thematic Apperception Test on a group of nine and ten year old children. Persona, 1949, 1, 12-15. (Psychological Abstracts, 24, 1210).

Wyatt, F. The scoring and analysis of the Thematic Apperception Test. Journal of Psychology, 1947, 24, 319-330.

Wyatt, F., & Veroff, J.B. Thematic apperception and fantasy tests. In D. Brower and L.E. Abt (Eds.), Progress in clinical psychology. Vol. 2. New York: Grune and Stratton, 1956, Pp. 32-57.

Zubin, J., Eron, L.D., & Schumer, F. An experimental approach to projective techniques. New York: Wiley, 1965.

Zuckerman, M., & Wheeler, L. To dispel fantasies about the fantasy-based measure of fear-of-success. Psychological Bulletin, 1975, 82, 932-946.

APPENDIX A

APPENDIX A

Relationship of Horner's Criteria for Fear-of-Success Stories and the Scoring Manual of Story Sequence Analysis.

A. Horner's Criteria for Fear-of-Success Stories

Fear of success is scored as present if the responses contain any of the following themes: negative consequences because of success, activities away from future success, direct expression of conflict about success, denial of effort or responsibility for attaining success, and bizarre or inappropriate responses.

B. The Scoring Categories in Arnold's Story Sequence Analysis which Correspond to the above Criteria.

- [-1, I.A.1.] Lesser goals are preferable; because they: a) require less effort, b) do not affect personal worth, c) are best.
- [-2, I.A.2.e.] Success is expected but failure is experienced instead.
- [-2, I.A.4] Failure as outcome; failure is: a) expected, b) experienced, just happens, etc., c) not admitted, d) caused by other people and things, e) result of chance, fate, etc.
- [-2, I.A.5.] Goal is not firmly pursued: a) it seems foolish, unrealistic, b) relinquished because of pain, danger, etc., c) becomes more difficult to reach, d) is wondered about.
- [-2, I.A.6.] No goal is indicated.
- [-1, and -2, I.B.] Whole category.
- [-1, and -2, I.D.] Whole category.
- [-1, and -2, I.E.] Whole category.
- [-1, and -2, I.F.] Whole category.

APPENDIX B

APPENDIX B

Descriptions of the Individual Male and Female Stimulus Cards used in this study (Murray, 1943, pp. 18-20).

Male Stimulus Cards

1. (1) A young boy is contemplating a violin which rests on a table in front of him.
2. (3BM) On the floor against a couch is the huddled form of a boy with his head on his right arm. Beside him on the floor is a revolver.
3. (6BM) A short elderly woman stands with her back turned to a tall young man. The latter is looking downward with a perplexed expression.
4. (7BM) A gray-haired man is looking at a younger man who is sullenly staring into space.
5. (8BM) An adolescent boy looks straight out of the picture. The barrel of a rifle is visible at one side, and in the background is the dim scene of a surgical operation, like a reverie-image.
6. (9BM) Four men in overalls are lying on the grass taking it easy.
7. (12BM) A young man is lying on a couch with his eyes closed. Leaning over him is the gaunt form of an elderly man, his hand stretched out above the face of the reclining figure.
8. (13BM) A little boy is sitting on the doorstep of a log cabin.
9. (18BM) A man is clutched from behind by three hands. The figures of his antagonists are invisible.
10. (16) Blank card.
11. (17BM) A naked man is clinging to a rope. He is in the act of climbing up or down.

Female Stimulus Cards

1. (2) Country scene: in the foreground is a young woman with books in her hand; in the background a man is working in the fields and an older woman is looking on.
2. (6GF) A young woman sitting on the edge of a sofa looks back over her shoulder at an older man with a pipe in his mouth who seems to be addressing her.

3. (3GF) A young woman is standing with downcast head, her face covered with her right hand. Her left arm is stretched forward against a wooden door.
4. (7GF) An older woman is sitting on a sofa close beside a girl, speaking or reading to her. The girl, who holds a doll in her lap, is looking away.
5. (12GF) The portrait of a young woman. A weird old woman with a shawl over her head is grimacing in the background.
6. (12GF) A bridge over water. A female figure leans over the railing. In the background are tall buildings and small figures of men.
7. (18GF) A woman has her hands squeezed around the throat by another woman who she appears to be pushing backwards across the banister of a stairway.
8. (13GF) A little girl is climbing a winding flight of stairs.
9. (9GF) A young woman with a magazine and a purse in her hand looks from behind a tree at another young woman in a party dress running along a beach.
10. (16) Blank card.
11. (8GF) A young woman sits with her chin in her hand looking off into space.

APPROVAL SHEET

The dissertation submitted by Roger S. Arnold has been read and approved by members of the Department of Psychology.

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

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

5/17/76

DATE

Ann E. Heilman

ADVISOR'S SIGNATURE