

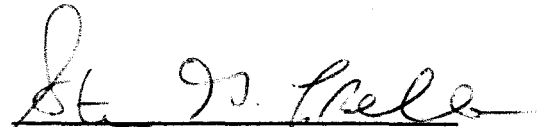
COMPARING THE RESPONSES OF AFRICAN-AMERICAN  
AND WHITE SUBJECTS ON THE MURRAY AND  
THOMPSON THEMATIC APPERCEPTION TEST

Mary Ana Sullivan

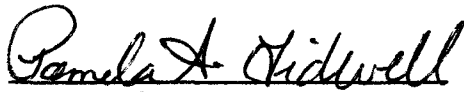
Certificate of Approval:



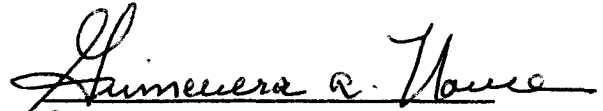
Allen K. Hess, Chair  
Head, Professor, Psychology



Steven G. LoBello, Associate  
Professor, Psychology



Pamela S. Tidwell, Assistant  
Professor, Psychology



Guinevera A. Nance  
Vice Chancellor  
Academic Affairs

COMPARING THE RESPONSES OF AFRICAN-AMERICAN  
AND WHITE SUBJECTS ON THE MURRAY AND  
THOMPSON THEMATIC APPERCEPTION TEST

Mary Ana Sullivan

A Thesis  
Submitted to  
the Graduate Faculty of  
Auburn University at Montgomery  
in Partial Fulfillment of the  
Requirements for the  
Degree of  
Master of Science

Montgomery, Alabama

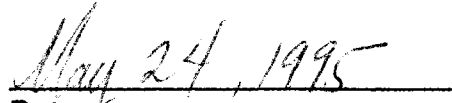
May 24, 1995

COMPARING THE RESPONSES OF AFRICAN-AMERICAN  
AND WHITE SUBJECTS ON THE MURRAY AND  
THOMPSON THEMATIC APPERCEPTION TEST

Mary Ana Sullivan

Permission is granted to Auburn University at Montgomery to make copies of this thesis at its discretion, upon the request of individuals or institutions at their expense. The author reserves all publication rights.

  
\_\_\_\_\_  
Signature of Author

  
\_\_\_\_\_  
Date

Copy sent to:

Name

Date

## VITA

Mary Ana Sullivan, daughter of Anita and Jim Sullivan, was born in Birmingham, Alabama, on May 1, 1970. In May of 1988, she graduated from Houston Academy in Dothan, Alabama. From August, 1988 to December, 1992, she attended the University of Alabama where she obtained a Bachelor's of Science in New College with a depth study focus in Psychology and Business. Between December, 1992 and August, 1993, she held the position of the Director of the Indian River's Crisis Line in Tuscaloosa, Alabama. She entered the Graduate School at Auburn University at Montgomery in September, 1993. In August, 1995, she will attend the University of Southern Mississippi to pursue her Ph.D. in Counseling Psychology.

## ACKNOWLEDGMENTS

I would like to express my thanks the members of my thesis committee, Dr. Allen Hess, Dr. Steven LoBello, and Dr. Pamela Tidwell. I appreciate the their contributions to this research and the guidance they have given me throughout this process. Also, I would like to thank Dr. Shelia Mehta for her contributions to this research.

I would like to thank Mrs. Hess for her contributions to the statistical analysis of this project. Additionally, I would like to thank Kathy Sellers, Anita Sullivan, and Patt Howard for the many hours they contributed to the scoring of the responses.

Special thanks to my mother, Anita Sullivan, and my father, Jim Sullivan, for the continuing support and encouragement they have given to me throughout this process and throughout my life.

THESIS ABSTRACT  
COMPARING THE RESPONSES OF AFRICAN-AMERICAN  
AND WHITE SUBJECTS ON THE MURRAY AND  
THOMPSON THEMATIC APPERCEPTION TEST

Mary Ana Sullivan

Master of Science, May 24, 1995  
(B.S., University of Alabama, 1992)

72 Typed Pages

Directed by Allen Hess

The purpose of this study was to investigate the hypothesis that subject's responses would be more productive on the Thematic Apperception Test (TAT) when they were presented with stimuli which depicted people of their own race as opposed to another race. Specifically, this study examined the responses of African-American and white subjects to Murray's TAT (M-TAT), a version depicting white people, and to Thompson's TAT (T-TAT), a version depicting African-American people. Sixty-four under-graduate students from introductory psychology courses from the University of Auburn at Montgomery served as subjects and the number of subjects was balanced on gender and race. There were two testing sessions for each subject in which they responded to six TAT cards, one session used the M-TAT version and the other the T-TAT version. The presentation of the M-TAT and the T-TAT was counterbalanced to control

for order effects. This was a 2x2x2 mixed factor design with Race, Sex, and Form of TAT serving as the independent variables. The dependent variables were word counts and type-token ratios which were calculated for the total number of words, nouns, adjective, adverbs and verbs. The overall analysis found little support for the hypothesis. Test x Race interactions were found on the Verb Type-token Ratios, the Adjective Type-token Ratios, and the Adjective Type-token Ratios for the initial session. However, the results of the Test x Race interaction for the Adjective Type-token Ratios were opposite of the expected results. Additionally, when alpha inflation is adjusted for none of the Test x Race interactions were significant. Limitations of the study and recommendations for future research are discussed.

D. The Verbal Facility Test

E. Word Classification

F. Part of Speech Classification



## TABLE OF CONTENTS

LIST OF TABLES . . . . .	x
I. INTRODUCTION . . . . .	1
II. LITERATURE REVIEW . . . . .	7
Precursors to the Thematic Apperception Test	
The Thematic Apperception Test	
Thompson's Version of the TAT	
III. STATEMENT OF PROBLEM . . . . .	24
IV. METHOD . . . . .	26
Permission to Conduct Research	
Subjects	
Materials	
Scoring	
Procedure	
V. RESULT . . . . .	31
Analysis of Word Counts	
Analysis of Type-token Ratios	
Analysis of Initial Session Data	
Pearson's Correlation	
VI. DISCUSSION . . . . .	45
REFERENCES . . . . .	48
APPENDICES . . . . .	54
A. Human Research Committee Permission to Conduct Research	
B. Statement of Informed Consent	
C. Demographic Information	

## LIST OF TABLES

1.	Analysis of Covariance for the Word Count Totals . . . .	33
2.	Analysis of Covariance for the Type-token Ratios . . . .	37
3.	Analysis of Covariance of the Initial Session Data for the Word Count Totals . . . . .	40
4.	Analysis of Covariance of the Initial Session Data for the Type-Token Ratios . . . . .	41
5.	Pearson Correlation Matrix for the Word Count Variables. . . . .	43
6.	Pearson Correlation Matrix for Type-token Ratios . . . . .	44

## COMPARING THE RESPONSES OF AFRICAN-AMERICAN AND WHITE SUBJECTS ON THE MURRAY AND THOMPSON THEMATIC APPERCEPTION TEST

The assessment of human characteristics has been a subject of interest since ancient times (McReynolds, 1975). Paul McReynolds (1975) states his definition of assessment as being “. . . the process whereby one person attempts to know, understand, or ‘size up’ another person; I will further limit the term to the delineation of conscious, deliberate systems for assessing persons” (p. 480). Accurate assessment of individuals has a great number of benefits for it allows characteristics of an individual, which are not obvious, to be known.

### History of Testing

The value of assessing individuals through testing was first officially recognized by the Chinese. The Chinese were conducting civil service testing as early as 2200 B.C.E. However, this early testing was unsystematic and it merely involved the emperor examining his officials. The interest of the Chinese in assessing individuals led to the development of more systematic procedures and by 1115 B.C.E. formal examination procedures had been established to evaluate a person’s aptitude for certain jobs (DuBois, 1966).

Although the earliest documented psychological testing was in the form of aptitude assessment, there was also early interest in personality assessment. It is not certain, but astrology was probably the first tool used in personality assessment and the personal horoscope was most likely the first psychological test used to assess personality (McReynolds, 1975).

The origin of horoscopal astrology is found in Babylon in the fourth century B.C. (Lindsey, 1971). The theory behind this notion was that the position of the planets at the time of a person's birth determined or influenced that individual's personality. Although there is no scientific validity for this system of personality assessment, the importance of the idea of a systematic method for assessing personality should be recognized.

Another precursor to modern personality assessment is physiognomy. Physiognomy is the practice of determining personality from physical characteristics. This technique of personality assessment was practiced in the ancient world and peaked in the second century C.E. (Evans, 1969). The significance of this method of personality assessment is that, unlike horoscopal astrology, it involved actual examination of the individual.

It is plausible that Pythagoras was responsible for the first personality assessment procedure (McReynolds, 1975). He conducted personality assessments, including physiognomic evaluations, at his Croatian center, which was founded in 529 B.C. Therefore, McReynolds concludes that "we can plausibly conjecture that the first personality assessment battery is almost exactly 2500 years old" (McReynolds, 1975, p. 491).

The next major precursor to modern personality assessment came in the 19th century with the development of moral statistics by Lambert Quetelet. Quetelet was a Belgian mathematician who developed a new field which involved applying the use of statistical procedures to analyze events which continually occur in society; such as births, suicides, marriages, and murders (Quetelet, 1969). His book was a step towards making the measurement of psychological variables possible.

The work of the phrenologists and Quetelet made important contributions

to the development of modern psychological assessment. Although the foundation of the work of the phrenologist is invalid, their work contributed the interest in attempting to qualify personality characteristics by relating them to measurable aspects of the individual. Quetelet's work provided an important step in the development of psychological statistics. Both the phrenologist and Quetelet helped to establish the stage for the emergence of the work of Francis Galton and the development of the modern era of psychological assessment.

Sir Frances Galton is considered to be the Father of Testing. Galton's main area of interest was in studying human heredity and he recognized the need to measure the similarities and differences of related and unrelated individuals. Galton was the first to gather a large systematic body on individual differences in sensory discrimination, which he believed was an indication of a person's intelligence. Galton also made advances in the use of rating-scale and questionnaire methods and in the adaptation of statistical methods for analyzing test data.

The United States military recognized the value of assessing individuals with psychological tests and began using psychological testing in World War I. These tests were the Army Alpha and the Army Beta and they were derived from multiple choice questions developed by Arthur S. Otis (Anastasi, 1988).

Today psychological testing has diverged to form four major types of testing. These are achievement tests, aptitude tests, intelligence tests, and personality tests. Results from these tests are used in educational settings; for selection of industrial personnel; and for diagnosis of disabilities and personality disorders. The personnel needs of the Armed Forces during World War I and World War II resulted in the development of many psychological tests

in the areas of intelligence and personality testing. However, today the focus has shifted from test development to careful scrutiny of existing tests and attempting to validate these tests (Bellack, 1993).

Psychological tests can facilitate assessing an individual's personality and abilities. However, this form of assessment is only beneficial when the tests provide an accurate assessment of the individual. Because psychological tests are used to evaluate and label individuals, test bias is a critical issue.

### Test Bias

#### Defining Test Bias

Psychological tests have a great impact in today's society. Consequently, the validity of psychological tests is being scrutinized in great detail. Ideally, psychological tests are intended to be objective measurements of an individual's personality characteristics or of their abilities. However, a common assertion is that psychological tests are contaminated by systematic measurement bias. Measurement bias of a predictor occurs if "it either overestimates or underestimates an individual's criterion performance depending on his group membership and if it correlates more with group membership than with the criterion it is intended to predict" (Jensen, 1980, p. 48).

Today one of the most common criticisms of psychological tests is that they are culturally biased. Cultural bias occurs when tests "discriminate unfairly against racial or ethnic minorities or persons of low socioeconomic status" (Jensen, 1980, p. 2).

Beginning in the 1960s a considerable amount of research began to accumulate concerning the possibility of biases in prediction of mental tests due to ethnic differences (Hunter, Schmidt, & Rauschenberger, 1984). The bias

being referred to here is indicating systematic error as opposed to chance error. The main issues of interest concerning test bias deal with the validity coefficient and the relationship between group means on the test taken and the measurement criterion (Anastasi, 1988).

Bias on the validity coefficient is referred to as slope bias and slope bias occurs when a test's validity coefficient reaches statistical significance for one group but does not reach it for another group. Whenever there is a statistically significant difference between the validity coefficient of two groups, slope bias is present (Anastasi, 1988).

The second type of test bias deals with the relationship between group means and the measurement criterion. This form of test bias is called intercept bias. This form of bias also applies to systematic differences but intercept bias occurs when a test systematically over predicts or under predicts the criterion performance of a group. Concern over the possibility of test bias and its effects has led to the empirical evaluation of test bias.

### Examining Test Bias

The growing concern over the biases of mental tests has led to revisions of tests and the development of new instruments. The Minnesota Multiphasic Personality Inventory (MMPI) is a test which has been revised in an attempt to eliminate bias based upon sex and race (Archer, 1992). The MMPI-2 includes 394 unaltered items from the MMPI, 107 new items, and 66 rewritten items. One of the main reasons for rewriting some of the test questions was to eliminate sexist language and to modernize the test content (Archer, 1992).

The MMPI-2 has also attempted to eliminate racial bias by using a normative sample which was representative of American society (Graham, 1993). The use of a representative norming sample is an improvement over the

normative sample of the MMPI, which did not represent minority groups. The norming sample for the MMPI-2 compares reasonably well to the data from the 1980 U.S. Bureau of the Census on the factors of sex and race, however, the norming sample for the MMPI-2 consists of a higher percentage of subjects with an education level of a bachelors degree or higher and a higher percentage of subjects with professional occupations than represented in the 1980 U.S. Census data. The issue concerning the effects of the skew of educational level and occupation is still being debated (Archer, 1992).

Test bias is a major issue in today's society. However, when this issue is empirically investigated, there is either no support for test bias against minority groups or the test slightly discriminates in favor of the minority group. These results have been found in the prediction of college grades (Breland, 1979; Duran, 1983), law school grades (Linn, 1975), performance in Army and Air Force training programs (Gordon, 1953; Maier & Fuchs, 1973; Shore & Marion, 1972), and a wide variety of industrial criteria (Hunter, Schmidt, & Rauchenberger, 1984).

Thus, there is generally no support for the claim of test bias, based upon race, in psychological testing. However, there are some psychological tests where the issue of racial bias has not been thoroughly examined. The Thematic Apperception Test is an example of one of these tests.



## LITERATURE REVIEW

### Precursors to the Thematic Apperception Test

Initially, the field of experimental psychology was limited to the study of sensation and perception. However, in the closing decades of the nineteenth century experimental psychology began to broaden and investigate issues such as memory, attention, reaction, and feelings (Tomkins, 1947). These new areas of interest led to the development of tests that attempted to assess these qualities of personality.

In 1907, Brittain published a paper called "A Study of Imagination". In this experiment, Brittain presented a series of nine pictures to a group of boys and girls and they were instructed to write a story about the pictures. These stories were scored for the use of names for persons or animals, the use of the first person, use of details present in the picture, their imaginative quality, their unity, length, explanatory power, and their use of religious, moral, and social elements.

The analysis of these stories revealed some significant differences in the responses of boys and girls. For example, the stories told by the girls revealed more religious and moral elements than the stories told by boys and the stories told by girls were full of pity, sadness, and fear of being alone. Brittain construed that there are three aspects of life; physical activity, affective life, and imaginal activity; which are mutually interactive.

A similar study was conducted by Libby (1908). He investigated how the imagination and feelings of school children are related. In this study a picture was presented to the children and they were asked to write a composition about

the picture. His main interest was a developmental one and Libby believed he had demonstrated that the imagination of high-school students was not poorer than that of grade-school children.

Brittian's technique remained relatively unexploited until it was rediscovered by the psychiatrist Schwartz (Tomkins, 1947). Schwartz was working at the Clinic for Juvenile Research in Detroit and, due to the overcrowded conditions, he usually did not have sufficient time to devote to the psychiatric interview. As a result, Schwartz developed and used the Social Situation Picture Test as an aid in initiating the interview (Schwartz, 1932).

The Social Situation Picture Test consists of eight pictures representing different situations which were found to be common in the histories of male juvenile delinquents. The pictures depict a boy in different situations. The subjects were asked to describe what they saw in the picture and then describe what the boy in the picture was thinking. Then the examiner would question the subject concerning his response.

The Social Situation Picture Test was never widely used by psychologists (Tomkins, 1947). This may be due to the difficulty of standardizing the test due to the series of questions which followed the responses. Also, the limited range of the pictures may have contributed to the limited use of the test. Although none of these tests have been widely used, the method has been adopted and developed to produce the Thematic Apperception Test (TAT), a test which has met with wider use (Bellack, 1993).

#### The Thematic Apperception Test

The TAT was formally introduced by C. D. Morgan and H. A. Murray (Morgan & Murray, 1935). The original set of TAT pictures consists of 19 of pictures and one blank card. The pictures were divided into two series with the

second series being “purposely more unusual, dramatic, and bizarre than those of the first” (Murray, 1943, p.2).

The present set of pictures is the result of the third revision of the original set, which was distributed by the Harvard Psychological Clinic in 1936, and this set consists of 31 cards, 30 pictures and one blank card. The third set also distinguishes itself from the first set by having pictures which are twice the size of the original pictures. This change was probably made to help facilitate the subjects rapport with the pictures (Bellack, 1993).

The current pictures were selected because of their stimulating power. To test the stimulating power of the pictures, subjects were presented with TAT pictures and the subjects’ personalities were assessed. The stories were then rated based upon the amount of information they contributed to the final diagnosis.

The TAT involves showing the subject a set of pictures of people in different scenes and having the subject write a story with a beginning, middle, and an end which describes the scene. Morgan and Murray believed that written fiction is an expression of the author’s conscious or unconscious fantasies or experiences. Therefore, the concept behind the TAT is that when people tell stories of the pictures they are revealing conscious and unconscious aspects of their personality. Basically this process functions as a shortcut to the conscious and unconscious and thus assists in speeding up the progress of personality analysis. Identification and projection are key concepts in the TAT and they find their basis in psychoanalytic theory.

#### The Process of Identification

Identification is “the act or process of becoming like something or someone in one or several aspects of thought or behavior” (Brenner, 1973,

p. 41). The process of identification is one that depends upon the individual's experience and it is of great importance in the development of the ego. Brenner discusses the process of language acquisition as an illustration of how identification is dependent upon one's experience.

The importance of identification on the TAT was recognized by Morgan and Murray. Morgan and Murray believed that, for projection to readily occur, the subject must be able to easily identify with at least one person in the picture. This person is called "an evoker, that is, one who evokes empathy in another" (Morgan & Murray, 1935, p. 280). Therefore, according to Morgan and Murray there should be at least one person in each of the pictures who is of the same sex and approximately the same age as the subject.

This indicates that Morgan and Murray were sensitive to the variables of sex and age in the process of identification. Because of the importance of the sex and age of the evoker in facilitating identification, Morgan and Murray indicated that separate sets of pictures should be developed for males and females as well as for children, adolescents, and senior citizens (Morgan & Murray, 1935).

However, this did not necessitate the development of completely different sets of pictures for males and females. Some of the pictures contain members of both sexes and in others the sex of the figure is questionable. It has been found that eleven of the pictures are suitable for both sexes (Murray, 1943).

To meet the requirement that there should be at least one person in the picture of approximately the same age as the subject, Murray followed the principle that at least one person in the picture should not be old enough to be a contemporary of the subjects parents. Based upon this principle, the set of pictures is appropriate for subjects age fourteen to forty.

This illustrates the recognition of the idea that what a person identifies with is affected by such variables of the individual as sex and age. These aspects affect the person's experience and, therefore, the person's frame of reference. The need to facilitate identification on the TAT has led to the development of the Senior's Apperception Test and the Children's Apperception Test (Bellak, 1965; & Bellak, & Bellak, 1973). Identification facilitates the next step in a revealing TAT response, which is projection.

#### The Process of Projection

The concept of projection was central to the development of the TAT. The term projection was first presented by Freud when he wrote "The psyche develops the neurosis of anxiety when it feels itself unequal to the task of mastering (sexual ) excitation arising endogenously. That is to say it acts as if it had projected this excitation into the outer world " (Freud, 1894, p. 102).

Freud further elaborated the concept of projection in his paper on the case of Schreber where he described the role of projection in paranoia (Freud, 1905). Basically Freud proposed that the paranoiac has certain homosexual tendencies which are transformed from love to hate due to the unacceptability of such feelings. The feeling of hate is then projected onto the former love object due to the prohibition of the hate feeling by the super-ego. Thus, projection functions as a defense mechanism.

Although the concept of projection was originally formulated in connection with psychosis and neurosis, it is currently applied in a broader sense to human functioning. A more recent definition of projection is an attribution of "a wish or impulse of his own to some other person, or for that matter, to some non personal object of the outside world" (Brenner, 1973).

Projection is the foundation of the TAT. The TAT is "based upon a 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" (Morgan & Murray, 1935, p. 279). Identification with the evoker is the first step in the production of successful and revealing TAT protocol and identification with the evoker leads to the second step, which is projection.

### Reliability

The TAT presents three main concerns when examining the reliability of the test. The first of these is interpreter reliability or the extent to which analysts agree on the interpretation. The second issue is test-retest reliability. The final issue is that of internal consistency.

#### Interpreter Reliability

The extent to which raters interpretations produce the same ratings on the TAT depends upon such variables as the training and competence of the interpreter and the scheme used for making these interpretations. Consequentially, the inter-rater reliability for interpreters ranges from .30 to .96 (Tomkins, 1947). The reliability tends to increase as the rationale for interpreting the TAT responses is discussed and made explicit. Investigators at the Harvard Psychological Clinic have achieved interpreter reliabilities as high as .95 (Tomkins, 1947).

#### Test-Retest Reliability

The reliability of the TAT not only depends upon the reliability of the test itself but is also a function of the stability of the individual's personality. The reliability coefficient on the TAT, as with all tests, has been found to decrease as the time interval between test administration increases (Tomkins, 1947). The subjects for this study were 45 white females ranging in age from 18 to 20 years of age. The reliability coefficient was .80 for a two month interval, .60 for a six

month interval, and .50 for a ten month interval.

Another study found reliability ranging from .46 for insecurity to .57 for blandness over a six week period (Hurley, 1955). Johnston (1957) reported correlations ranging from .55 to .78 on the same variables for a one month interval. An average retest reliability reported once a year over a three year period was found to be .46 based upon Murray's need-press schema (Sanford, 1943).

#### Internal Consistency

Another factor is the internal consistency of the TAT. The internal consistency values for the studies on the TAT are generally low. When the first half of individual's responses were compared with the second half the reliability results range from .13 for Murray's ten major needs to .78 for need for achievement (Child, Frank, & Storm, 1956).

#### Validity

The information on the validity of the TAT is sparse in the TAT manual Murray (1943). Murray merely reports that there is a positive correlation which is greater than .40 for college men between their responses on the TAT and their overt behavior for the need for Abasement, need for Creation, need for Dominance, need for Exposition, need for Nurturance, need for Passivity, and the need for Rejection and Dejection. A negative correlation, ranging from -.33 to -.74, was reported for the need for Sex and no correlation was found for need for Aggression and need for Achievement.

Since the publication of the TAT manual a plethora of research has been conducted on the validity of the TAT (Murray, 1943). Although there is a great deal of research on this issue, the results cannot necessarily be generalized. There are many different scoring schemes for the TAT and the validity of each

approach needs individual examination.

Probably the most thoroughly researched method for scoring the TAT has been provided by McClelland, Atkinson, Clark and Lowell (1953). This scoring method outlines their strategy for studying achievement motives and other dimensions of motivation. However, this method requires a great deal of training and experience for it to be mastered by the examiner (Ryan, 1985).

Another issue which appears to affect TAT validity besides the scoring scheme is the instructional set. Lundy (1988) investigated the responses to TAT's produced following four instructional settings: neutral, following a personality test, emphasizing that the TAT is a personality test, and in a non-threatening but structured situation. Neutral instructions produced significant correlations between need for Achievement and achievement criteria ( $r=.29$ ,  $p < .01$ ) need for Affiliation and the affiliation criteria ( $r=.24$ ,  $p < .05$ ) and need for Power and power criteria ( $r=.34$ ,  $p < .01$ ). Additionally, significant differences were found between the neutral instructions and nonneutral instructions for achievement, affiliation, and power. This demonstrates the necessity of using neutral instructions in order to produce valid TAT profiles.

The usefulness of the data produced by the TAT is directly related to the validity of the assessment method and the skill of the interpreter. Some scoring methods have been shown to have reliability and construct validity while others have not been thoroughly researched. Therefore, criticisms concerning the issues of the reliability and validity of the TAT should be directed at specific methods of interpretation and should not be generalized to all methods of interpretation.

#### Biases of the TAT

The TAT is the fifth most commonly used psychological assessment



measure (Harrison, Kaufman, Hickman, & Kaufman, 1988). The MMPI and the Rorschach are the only personality measures which are used more frequently than the TAT. Additionally, it remains one of the most important projective personality techniques used for research and clinical purposes (Kaplan & Saccuzzo, 1989). Due to the popularity and range of TAT use, biased results may have significant and wide ranging effects.

The TAT would be biased if differences in results were due to group membership instead of differences in personality characteristics. One problem which may ensue from bias on the TAT is that research results may promote stereotypes. Another possible consequence is that the evaluation of individuals may result in inaccurate conclusions or treatment recommendations.

There has been a great deal of interest concerning the issue of gender bias on the TAT. A review of the research in this area indicates that the results have been inconclusive (Stewart, & Chester, 1982). However, the TAT cards are designated as male stimulus cards or female stimulus cards, meaning that some cards are more appropriate for females and some for males.

A more recent study found a significant interaction between gender and type of card in the categories of withdrawal and anxiety (Worchel, Aaron, & Yates, 1990). Men responded to a higher degree in the categories of the withdrawal and anxiety factor when presented with female pictures and women responded to a higher degree on these factors when presented with male pictures. Thus, the interaction between type of TAT card and the gender has a significant effect on TAT responses.

Race is another area of possible bias on the TAT. The factor of race is an issue because all of the TAT stimulus pictures, which have people in them, depict white people. This approach was appropriate during the development of

the TAT because it was developed on white males. However, one must question the issue of whether a test developed on one racial group can accurately be applied to another racial group, without some modification.

A study which found a racially sensitive variable, as determined by the TAT, is one which compared the responses of African-American boys to white boys on the Murray TAT (Mussen, 1953). Mussen found that African-American boys show a greater need for aggression on two out of the four measures of need for aggression in response to the Murray TAT than white boys. These two measures were defined as "Hero assaulted, injured" and "Hero hated, quarreled with". It was also found that African-American boys view the environment as basically inhospitable. One must ask the question of whether this result is due to a real difference between the groups or due to the stimuli. If the difference is due to the type of stimuli, the M-TAT may be incorrectly providing support for stereotypes. Research in this area has been sparse. However, the necessity of investigation in this area was recognized by Charles Thompson who developed his own version of the TAT (Thompson, 1949).

#### Thompson's Version of the TAT

The interest in facilitating identification and projection on the TAT led to the development of the Children's TAT and the Senior's TAT (Bellack, 1965; & Bellack, & Bellack, 1973). Less attention has been focused on the adequacy of testing people of different racial backgrounds from those depicted in the current TAT. TAT's have been developed for Indians (Chowdhury, 1960), Africans (Lee, 1953), and African-Americans (Thompson, 1949) in an attempt to investigate the differences in responses when the subject is of the same race as those depicted in the TAT pictures as opposed to being of a different race. The present study is concerned with the latter group.

In 1949, Charles E. Thompson presented a African-American version of the TAT which is called Thompson's Thematic Apperception Test (T-TAT). The importance of the sex and age of evoker had been recognized as crucial in facilitating identification and projection when Thompson began his work with the TAT. Therefore, it seemed natural that identification and projection would also be facilitated if the evoker was of the same race as the subject.

Thompson's version of the TAT was the next logical step in facilitating identification and projection on the TAT for African-American subjects.

Thompson's interest in this issue was sparked when he had an African-American client who produced a sparse TAT record. Following this result, Thompson asked the client to imagine that the people in the stimulus pictures were African-American. The result was a more productive record where the subject expressed a greater degree of "belongingness." Thompson concluded that African-American subjects would identify with stimulus pictures of African-American people to a greater degree than to stimulus pictures of white people.

#### Thompson's Study

To test this hypothesis, Thompson chose ten of the male stimulus pictures from Murray's TAT and gave them black features. The pictures he chose were numbers 1, 2, 4, 5, 6BM, 7BM, 8BM, 12M, 17BM, and 18BM. The subjects for this experiment were drawn from adult, African-American male veterans who were participating in the 1947-48 academic program at Dillard University in New Orleans, Louisiana. The subject population consisted of 26 volunteers.

The subjects were then randomly assigned four experimental groups. Group I (n = 7) was presented the Murray TAT (M-TAT) followed by the T-TAT with a time interval of seven days. Group II (n = 7) was presented the T-TAT

followed by the M-TAT with a time interval of seven days. The procedure for Group III (n = 6) and Group IV (n = 6) were the same as Group I and Group II except that the time interval was one day.

Because Tomkins (1947) found no difference in whether TAT responses were given orally or written, Thompson used the written procedure. The subjects were tested in groups for the precedent for the procedure of group administration of the TAT had already been established (Tomkins, 1947). No significant differences in protocol content had been found in the responses resulting from individual administration of the TAT versus group administration for such variables as achievement, affiliation, transcendence, number of themes, dominance, involvement, sex, quality, number of figures and compliance with instructions (Lindzey & Silverman, 1959; & Murray & MacKinnon, 1946).

The responses were scored based upon verbal productivity. This measure included scores “for the total number of words used, for verbs (defined as ‘a part of speech expressing action, occurrence or mode of being’), for nouns-excluding pronouns (defined as a ‘word that is the name of a subject of discourse, as a person, place, thing, quality, idea, action’), and for modifiers, adverbial and adjectival (defined as ‘to limit or restrict the meaning of , qualify’)” (Thompson, 1949, p. 474).

There was no significant relationship found between length of interval between tests and verbal productivity. However, there was a significant difference found between verbal productivity and test type. The mean difference of verbal productivity was found to be significant for type of test given with African-Americans being more productive on the T-TAT than the M-TAT. However, Thompson only tested African-American males. One can not assume

that just because African-American males are more verbally productive on a TAT which is matched to their race that white subjects will respond in the same manner. The possibility that white subjects will also respond in a more verbally productive manner to the T-TAT than the M-TAT must also be investigated.

#### Studies Supporting Thompson's Hypothesis

Other investigators have found that African-American stimulus pictures do lead to facilitating story production by African-American subjects (Bailey, & Green, 1977; Cowan, 1971; & Cowan, & Goldberg, 1967). Bailey and Green (1977) tested 135 African-American males with the M-TAT, T-TAT and an experimental TAT (E-TAT). The E-TAT is a modification of the M-TAT where the figures had been redrawn to depict African Americans with modern clothes and hair styles.

The subjects were divided into groups based upon age; 17 - 18 year-old high-school students (HS), 19 - 22 year-old college students (CS), and 25-45 year-old residents of Houston, Texas and the surrounding communities (R). The experimental design was a 3 x 3 factorial design based upon age and form of TAT. One problem with this study is that no attempt was made to control for such factors as intelligence or socioeconomic status. Also, the subjects only took one form of the TAT and, therefore, did not serve as their own controls.

After the test had been administered the subjects rated the TAT on a series of 1 - 10 scales on the topics of (a) the degree to which the test looked like people in general, (b) the degree to which subjects were able to relate their personal feelings to the test, (c) the degree to which the test series represented African-American or white people, (d) the degree to which the people, as opposed to the background, facilitated the writing of the subject's stories. No objective quantitative measure was conducted for productivity. The stories were

also rated, by trained raters, for content using Murray's (1943) needs for Achievement (n-Ach.), n-Sex, n-Aggression (n-Agg.), and n-Succorance (n-Succ.).

The results indicated a significant difference on the measure of representing people in general between the T-TAT and the E-TAT with the T-TAT being rated as depicting people in general. For the ability of the test to facilitate story writing the T-TAT was rated significantly higher than the M-TAT. However, no significant differences were found on the measures of content between the T-TAT, M-TAT, and E-TAT. As in Thompson's study, white subjects were not tested and, therefore, it can not be assumed that white subjects would rate the M-TAT higher than the T-TAT on the ability of the test to facilitate story writing.

Another study examining the effects of race and gender of TAT cards upon productivity was conducted by Cowan and Goldberg (1967). In this study, the responses of sixty-seven African-American females were compared for protocol length for the factors of gender of card and race of card. For this study an African-American version of the M-TAT was designed. The results indicated that the African-American version of the TAT produced greater protocol length. However, there was no significant difference in need for Achievement scores when the protocols were corrected for length differences.

One criticism of this experiment is that no control was made for individual factors, such as intelligence level, and, because each subject merely took one form of the TAT, these factors may have a critical influence on the productivity levels of protocols. Additionally, white subjects were not tested in this study.

#### Studies Opposing Thompson's Hypothesis

Further investigation of Thompson's rationale has been conducted

which find no support for Thompson's Hypothesis. Korchin opined that it "has been assumed, tacitly perhaps, that the test stimuli (Murray's TAT) could be used equally well with all types of subjects" (Korchin, Mitchell, & Meltzoff, 1950, p. 445). However, this assumption was not held by Morgan and Murray. As discussed previously, Morgan and Murray indicated the necessity for developing TAT cards for subjects of different ages and sexes.

One investigation of Thompson's rationale was conducted by Riess, Schwartz, and Cottingham (1950). Riess et al. used a three-fold experimental matrix varying race of subject, race of stimulus card, and race of administrator. The subjects were 30 white females and 30 black females. The subjects were individually presented with five white cards (1, 8GF, 3GF, 12F, and 13G) and five black cards (3BM, 4, 6GF, 7GF, and 5). The cards were presented in the following order: 1, 3BM, 4, 8Gf, 6GF, 3GF, 7GF, 5, 12F, and 13G. No rationalization was presented for the chosen order and no consideration was given for possible contamination due to order effects.

The failure to use two versions of the same card presents another possible source of contamination of the results due to the individual card for it has been found that different cards have different eliciting properties (Ryan, 1985). Also, Riess compared the length of response of African-Americans to whites on each form of the TAT instead of comparing the length of responses of each race on the two TAT forms. The results showed no significant differences between African-Americans and whites on the measure of story length.

Another experiment tested the hypothesis that African-Americans will be less productive, in terms of story-length, than white subjects on the M-TAT (Korchin, et al., 1950). The subjects were 80 white and 80 African-American males. The racial groups were divided into middle and lower socioeconomic

levels. The subjects were also matched based upon level of education; lower class, 0-9 years of schooling; middle class, 12-18 years of schooling; and for their father's occupation. How the father's occupation was rated was not indicated by the investigators.

The subjects were individually administered four M-TAT cards; 1BM, 2, 6BM, and 7BM. No main effect for race of subject or race of card was found and no interaction effect was found for these two factors. In this study, the Thompson version of these cards was not presented, therefore, there is no comparative measure for performance based upon TAT form and the comparison is between the two racial groups instead of within each racial group.

Korchin also states that when " cards are used in their original form, with white figures, it is our belief that for the majority of subjects in our culture, white and Negro alike, they represent 'universal' human beings onto which may be projected basic personality data" (Korchin et al., 1950, p. 450). No basis is provided for this view and one must consider the time in which it was expressed.

A further investigation of Thompson's rationale was conducted by Light (1955). Light administered the M-TAT and the T-TAT to 26 white college students. Group I was administered the M-TAT followed by the T-TAT and Group II was administered the T-TAT followed by the M-TAT. Cards 1, 2, 3GF, 3BM, 4, 5, 6GF, 6BM, 12F, and 18BM were used in this study. A two week interval occurred between the administrations. The results were analyzed for story length, thematic differences, and direct reference to "Negroes". Story length was determined by a cumulative word count. The method of analyzing theme differences and the direct references to African-Americans was not



indicated in the article.

The groups were also compared based upon those who said that they had responded to color and those who said they had not, 14 subjects and 12 subjects respectively. The examination of the protocols of the 14 subjects who did respond to color revealed that 29% of the T-TAT stories cast African-Americans as inferior individuals while only 9% of the M-TAT stories cast whites in an inferior light. This suggests differences in the treatment of the subject matter of these 14 subjects based upon the race of the subject in the stimulus picture.

Light found no significant differences between story length based upon the form of the card. A significant difference was found in the theme quality between those who claimed they responded to color and those who claimed they did not respond to color. The themes of those who responded to color were characterized by an increased use of racial stereotypes. However, only white subjects were used in this study.

## STATEMENT OF PROBLEM

Psychological tests have been subjected to increased scrutiny with regard to the test content. Specifically, this analysis has focused on whether or not the materials affect different groups due to group membership.

With regard to the TAT, one of the issues which arises is how the type of stimulus picture affects identification and, therefore, projection when the subject is of a race other than those depicted in the stimulus pictures. Fortuitously, a African-American TAT has been developed. However, research examining the possible difference due to stimulus type and the effects of the race has been sparse and inconclusive.

Some research has supported the importance of matching the race of the subject with the race of the stimulus picture ( Bailey & Green, 1977; Cowan, 1971; Cowan & Golberg, 1967; & Thompson, 1949). However, other studies have produced result which do not support these conclusions (Light, 1955; Riess, Schwartz, & Cottingham, 1950; Schwartz, Riess, & Cottingham, 1951).

Additionally, methodological problems arise in each of these studies. These problems include only one form of the TAT being used, only one race being tested, different forms of the M-TAT and T-TAT being used, and the variables of IQ and socioeconomic not being controlled. Therefore, the inconsistency of results and the lack of methodological controls warrants further investigation of this issue.

The present study examined this issue under stricter methodological control. Both African-American subjects and White subjects were tested and

each subject was tested on both the M-TAT and the T-TAT versions of six of the same TAT cards. Additionally the variables of IQ and socioeconomic status will be controlled.

Hypothesis 1: African-American subjects will identify to a greater degree with the T-TAT than the M-TAT as seen by scores on word counts and type-token ratios.

Hypothesis 2: White subjects will identify to a greater degree with the M-TAT than the T-TAT as seen by scores on word counts and type-token ratios.

## METHOD

### Permission to Conduct Research

Permission was obtained from the Human Research Committee at Auburn University at Montgomery to conduct this study, prior to the inception of data collection (Appendix A).

### Subjects

The subjects were 64 students from introductory psychology classes at Auburn University at Montgomery. There were thirty-two African-American subjects (sixteen females and sixteen males) and thirty-two white subjects (sixteen females and sixteen males). The recruitment of subjects was limited to American citizens between the ages of 18 and 23 and the mean ages for each group were as follows: African-American males  $M = 20.5$ , African-American females  $M = 19$ , white males  $M = 19$ , and white females  $M = 19$ . The subjects were randomly assigned to one of two groups categorized by the order of TAT form presentation. These groups were balanced based upon the factors of race and gender of the subjects.

### Materials

Six TAT slides were used in this experiment. There were two versions of each slide. One was the M-TAT version and the other the T-TAT version. Cards 1, 2, 4, 6BM, 13MF, and 18GF were used in this study. Cards 1, 2, 4, and 13MF are among the cards essential for testing males and females (Bellack, 1993). Additionally, the Thompson version of all of the cards chosen were judged by the committee to be good reproductions of the Murray cards, with the

only the characteristic of race being altered. Cards 6BM and 18GF were selected to balance the gender designation of the cards.

### Procedure

Each subject attended two testing sessions and at the inception of the initial session a consent form was signed by each subject (Appendix B). All subjects were tested by the same examiner who was a white female. The subjects were presented with both the M-TAT and the T-TAT versions of the six TAT cards and the order of the form of TAT presentation was counterbalanced. It was attempted to test the subjects in groups of eight, balanced based upon the factors of gender and race. However, due to no-shows this proved to be a difficult task. The result was that group size ranged from two to eight subjects. Studies examining the results of TAT responses under group administration and individual administration conditions have found no significant difference in the protocols (Eron & Ritter, 1951; Lindzey & Heinemann, 1955; Lindzey & Silverman, 1959; Sarason & Sarason, 1958; & Terry, 1952).

The subjects were also scheduled to be tested with a one week interval between the subject's first and second session. A one week interval was chosen because it was the interval used by Thompson (1938) and for convenience in testing. Due to no-shows for the first scheduling of the second session, nine subjects had their second testing session after a two week interval. Eight of these subjects were African-American males and one was a African-American female.

Slides of the TAT stimulus pictures were projected onto a screen for a 20 second period, after which the picture was no longer projected. The subjects then had five minutes to write a response to the picture. This manner of presenting the stimuli is optimal for group administration (Lindsey &

Heinemann, 1955).

The instructions were as follows:

This is a test of imagination. I am going to show you some pictures, one at a time; and your task will be to make up as dramatic a story as you can for each. Tell what led up to the event shown in the picture, describe what is happening at the moment, what the characters are feeling and thinking; and then give the outcome. Write your thoughts as they come to your mind. Each picture will be shown for 20 seconds and you will have five minutes to respond to each picture. Do you understand? Here is the first picture.

The same instructions were read at the inception of each session.

Demographic information was obtained following each of the sessions. A demographic questionnaire was filled out by the subjects following session one and the yearly income of the subject's parents was used to rate socioeconomic status(Appendix C). I.Q. information was obtained following the second session based upon an assessment of the subjects' verbal ability (Appendix D). The assessment of verbal ability was based upon the subjects' scores on the Verbal Facility Test (Ross, 1972). This test correlates .84 with the Wechsler Adult Intelligence Scale verbal intelligence quotient, .52 with the performance intelligence quotient, and .82 with the full scale intelligence quotient (Ross, 1972). The variables of IQ and socioeconomic status must be controlled when assessing differences in personality characteristics of groups (Bellack, 1993; Megaree, 1966; Rose, 1958; Rubin, 1964; & Veroff, Wilcox, & Atkins, 1953).

### Scoring

The rules set out by Mann (1944) were used to determine what constitutes a word and for classifying the word on part of speech (Appendix E and Appendix F). The subject's responses were scored for Verbal Productivity and Type-token Ratios.

The measure of verbal productivity was based upon word counts for the total number of words, nouns, adjective, verbs, and adverbs. The Type-token Ratio (TTR) has been described as a measure of "vocabulary diversity" (Cramblit, 1977, p. 476), a measure of "vocabulary richness" (Andolina, 1980, p. 373), and "a relatively simple, straight-forward measure of language deviance" (Manschreck, Maher, & Ader, 1981, p. 7).

The TTR is the ratio of the total number of different words, types, to the total number of words, tokens, in a given language sample (Johnson, 1941). For example, if one was calculating a TTR for nouns it would be the number of different nouns used in the language sample divided by the total number of nouns used in the language sample. One fact that must be accounted for is that type-token ratios have a negative acceleration. That is, the number of tokens increases at a faster rate than the number of types. Therefore for comparative purposes, the length of the samples being compared must be equivalent in terms of number of words (Chotlos, 1944).

In order to make samples of different magnitudes comparable, the samples were divided into 100 word segments, the TTRs were computed for each segment, and then the mean segmental TTR was computed by obtaining the average of the segmental TTRs. The Type-token Ratios were calculated for the total number of words, nouns, adjectives, verbs, and adverbs. Because the total number of words and the specific word produced are interrelated the

Type-token Ratio provides a control for such covariations.

Each of the subject's responses was scored by two examiners who reached a 93% agreement rate. Differences in score were not reconciled because there was a high percentage of agreement between the scores. Instead the data analyzed was from the scores obtained by the experimenter.



## RESULTS

First, data are presented concerning the dependent variables of word counts. The word counts were calculated for overall number of words and the number of nouns, adjectives, verbs, and adverbs that the subjects produced in their M-TAT stories and their T-TAT stories.

Next, the type-token ratios; which were computed for total number of words, nouns, adjectives, adverbs, and verbs of the subject's responses on the M-TAT and the T-TAT; are depicted. The next section concerns the analysis of the five word count variables and the five type-token ratios for the responses given on the first session only. Finally, a correlation matrix between the T-TAT variables and the M-TAT is presented.

Analyses of Covariance Variances (ANCOVA) were conducted for the between-subjects effects and within-subject effects for each of the dependent variables looking for main effects for Gender, Race, Test and Order and for the interaction effects. The Race x Test interaction is the analysis which tested the hypothesis. Additionally, the Analysis of Covariance was conducted with the Verbal Facility Test score and the Family Income as covariates.

Due to the large number of analysis conducted, the alpha level was adjusted based upon the Bonferroni method to control for Type I. error (see Keppel, 1982). The adjusted alpha resulted in a  $p = .003$  for significance. Both the analysis reaching significance at the standard alpha level and the adjusted alpha level are reported. The more liberal statistics of the non adjusted alpha

were used to insure that if there were any differences they would be detected. Additionally, the homogeneity of slopes was tested for all analysis and when there are problems it will be noted with the analysis.

#### Analysis of Word Count Variables

The ANCOVA of the word count variables yielded the results in Table 1.

#### Analysis of Total Word Count

In the analysis of the between-subject effects for the Total Word Count a significant effect was found for Gender,  $F(1,54) = 10.48, p < .005$ , with Females producing responses with significantly more words than Males (Female  $M = 562.63, SD = 5.68$ ; Males  $M = 468.43, SD = 5.68$ ). This difference was irrespective of race or form of TAT. This difference was also significant at the adjusted alpha level. Also, a significant effect was found for the Verbal Facility Test covariate,  $F(1,54) = 5.14, p < .05$ , indicating greater test productivity for more verbally facile people.

A significant within-subject effect was found on the Test x Order interaction for the Total Word Count,  $F(1,54) = 6.23, p < .05$ . (T-TAT/M-TAT: T-TAT  $M = 542.06, SD = 5.34$ ; M-TAT  $M = 503.39, SD = 5.98$ ; M-TAT/T-TAT: M-TAT  $M = 524.83, SD = 5.34$ ; T-TAT  $M = 491.85, SD = 5.98$ ). The analysis of simple effects did not produce any significant results. Additionally, there was a significant Family x Gender interaction,  $F(1,52) = 5.34, p < .05$ . No other significant effects were found on the analysis of the Total Word Count.

#### Analysis of Noun Count Totals

A significant effect for the Verbal Facility Test covariate was found in the analysis for the between-subjects effects of the Noun Count Totals,  $F(1,54) = 5.39, p < .05$ , indicating that more verbally facile individuals produce a greater

Table 1  
 Analysis of Covariance  
 for the Word Count Totals

	Total	Noun	Adj	Verb	Adv
Between subjects					
Gender	10.48***	n.s.	7.72	11.40****#	n.s.
Race	n.s.	n.s.	n.s.	n.s.	n.s.
Order	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Race	n.s.	n.s.	4.73*	n.s.	n.s.
Gender*Order	n.s.	n.s.	n.s.	n.s.	n.s.
Race*Order	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Race*Order	n.s.	n.s.	n.s.	n.s.	n.s.
Verbal	5.14*	5.39*	10.08***	4.26*	n.s.
Faminc	n.s.	n.s.	n.s.	n.s.	n.s.
Within-subject					
Test	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Gender	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Race	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Order	6.23*	n.s.	n.s.	5.94*	4.08*
Test*Gender*Race	n.s.	n.s.	n.s.	n.s.	6.65**
Test*Gender*Order	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Gender*Race*Order	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Verbal	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Faminc	n.s.	n.s.	n.s.	n.s.	n.s.

NOTE: Gender = female or male subjects; Race = African-American or white subjects; Order = the order in which the T-TAT and the M-TAT were presented; Verbal = the Verbal Facility Test score; Faminc = family income; Test = the form of TAT (T-TAT or M-TAT); TOTAL = the Total Word Count; Noun = the Noun Count Totals; Adj = the Adjective Count Totals; Adv = the Adverb Count Totals; Verb = the Verb Count Totals; \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .005$ ; \*\*\*\* $p < .001$ ; # $p$  indicates significance at the adjusted alpha level.

number of nouns. No significant effects were found on the within-subject analysis for this dependent variable. However, a significant Faminc x Gender interaction was found  $F(1,52) = 4.77, p < .05$ .

#### Analysis of Adjective Count Totals

The analysis of the Adjective Count Totals for the between-subjects effects yielded a significant effect for Gender,  $F(1,54) = 7.72, p < .01$ , with Females using more adjectives than Males ( Females  $M = 61.64, SD = 1.07$ ; Males  $M = 51.05, SD = 0.78$ ). The Gender x Race interaction was also significant,  $F(1,54) = 4.73, p < .05$ . In the analysis of simple effects, White Females were more productive than African-American Males,  $F(1,29) = 3.35, p < .001$ ; White Males were more productive than African-American Males,  $F(1,30) = 2.80, p < .01$ ; and African-American Females were more productive than African-American Males,  $F(1,29) = .0001$ , (White Females  $M = 59.35, SD = 0.76$ ; African-American Females  $M = 63.92, SD = 0.80$ ; White Males  $M = 57.03, SD = .76$ , African-American Males  $M = 45.07, SD = 0.74$ ). Additionally, there was a significant effect for the Verbal Facility Test score,  $F(1,54) = 10.08, p < .005$ , indicating greater adjective production for more verbally facile individuals. The within-subject analysis for Adjective Count Totals did not yield any significant results. There was a significant Verbal x Race interaction,  $F(1,52) = 4.12, p < .05$ , and a significant Faminc x Race interaction was found,  $F(1,52) = 5.73, p < .05$ .

#### Analysis of Verb Count Totals

The between-subjects analysis for the Verb Count Total yielded significant results for Gender,  $F(1,54) = 11.40, p < .005$ , with Females using more verbs than males regardless of race or form of TAT ( Females  $M = 147.23, SD = 1.57$ ; Males  $M = 119.63, SD = 1.57$ ). A significant effect was also found for

the Verbal Facility Test score,  $F(1,54) = 4.26, p < .05$ , indicating that noun production increases as verbal facility increases. On the within-subject analysis of the Verb Count Totals, a significant effects for the Test x Order interaction was found,  $F(1,54) = 5.94, p < .05$ . However, when the simple effects analysis was conducted no significant results were found (T-TAT/M-TAT: T-TAT  $M = 139.22, SD = 1.56$ , M-TAT  $M = 130.09, SD = 1.57$ ; M-TAT/T-TAT: M-TAT  $M = 136.66, SD = 1.57$ , T-TAT  $M = 127.75, SD = 1.56$ ).

#### Analysis of Adverb Count Totals

The between-subjects analysis of the Adverb Count Totals did not result in any significant effects. On the within-subject analysis a significant effect was found for the Test x Order interaction,  $F(1,54) = 4.08, p < .05$ . There was greater adverb usage on the M-TAT than the T-TAT for the M-TAT then T-TAT order,  $F(1,29) = 2.29, p < .05$ , (M-TAT  $M = 17.00, SD = .35$ ; T-TAT  $M = 14.84, SD = .35$ ).

A significant Test x Gender x Race interaction was also found for adverb production,  $F(1,54) = 6.65, p < .05$ . On the M-TAT White Females produced significantly more adverbs than African-American Males,  $F(1,54) = 2.04, p = .05$ , ( White Females  $M = 19.73, SD = .47$ ; African-American Males  $M = 13.99, SD = .47$ ). On the T-TAT, African-American Females produced significantly more adverbs than African-American Males,  $F(1,54) = 2.76, p < .05$ , ( African-American Females  $M = 21.21, SD = .40$ ; African-American Males  $M = 12.36, SD = .40$ ). No other significant effects were found for the Adverb Count Totals.

#### Analysis of Type-token Ratios

The ANOVA of the Type-token Ratios for responses to both sessions yielded the results in Table 2.

### Analysis of the Total Type-token Ratio

In the analysis of the Total Type-token Ratio no significant effects were found on the between-subjects or within-subject analysis.

### Analysis of the Noun Type-token Ratio

The Noun Type-token Ratios for the between-subject analysis produced a significant effect for the Gender x Order interaction,  $F(1,54) = 3.97, p < .05$ . Females produced larger Noun Type-token Ratios than Males for the M-TAT then T-TAT order,  $F(1, 29) = 2.58, p < .01$ , ( Females  $M = .84, SD = .003$ ; Males  $M = .80, SD = .025$ ). The within-subject analysis of the Noun Type-token Ratios resulted in no significant effects.

### Analysis of the Adjective Type-token Ratio

The between-subject analysis for the Adjective Type-token Ratios yielded no significant effects for this variable. However, the within-subject analysis yielded a significant Race x Test interaction,  $F(1,54) = 4.45, p < .05$ . When the simple effects were analyzed no significant results were found ( T-TAT: African-American Subjects  $M = .79, SD = .004$ , White Subjects  $M = .79, SD = .004$ ; M-TAT: African-American Subjects  $M = .80, SD = .003$ , White Subjects  $M = .77, SD = .003$ ).

There was a significant interaction for Test x Gender x Order on the Adjective Type-token Ratio,  $F(1,54) = 8.99, p < .005$ . On the M-TAT, Males produced larger Type-token Ratios for the T-TAT then M-TAT order than the M-TAT then T-TAT( Male M-TAT: M-TAT/T-TAT  $M = .76, SD = .05$ ; T-TAT/M-TAT  $M = .81, SD = .05$ ),  $t(32) = 3.03, p < .05$ . Also, Males produced larger Murray Adjective Type-token Ratios for the T-TAT then M-TAT order than females ( Males  $M = .81, SD = .20$ ; Females  $M = .76, SD = .20$ ),  $t(32) = 3.24, p < .05$ . On the Thompson Adjective Type-token Ratio, Males produced larger Adjective

Table 2

Analysis of Covariance  
for the Type-token Ratios

	TTRT	TTRN	TTRAdj	TTRV	TTRAdv
Between subjects					
Gender	n.s.	n.s.	n.s.	n.s.	n.s.
Race	n.s.	n.s.	n.s.	n.s.	n.s.
Order	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Race	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Order	n.s.	3.97*	n.s.	n.s.	n.s.
Race*Order	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Race*Order	n.s.	n.s.	n.s.	4.32*	n.s.
Verbal	n.s.	n.s.	n.s.	n.s.	n.s.
Faminc	n.s.	n.s.	n.s.	n.s.	n.s.
Within subjects					
Test	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Gender	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Race	n.s.	n.s.	4.45*	7.55**	n.s.
Test*Order	n.s.	n.s.	n.s.	4.69*	n.s.
Test*Gender*Race	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Gender*Order	n.s.	n.s.	8.99***	n.s.	n.s.
Test*Gender*Race*Order	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Verbal	n.s.	n.s.	n.s.	n.s.	n.s.
Test*Faminc	n.s.	n.s.	n.s.	n.s.	n.s.

NOTE: Gender = female or male subject; Race = African-American or White

subjects; Order = the order in which the T-TAT and the M-TAT were

administered; Verbal = the Verbal Facility Test score; Faminc = family income;

Test = the form of TAT (T-TAT or M-TAT); TTRT = Total Type-token Ratio; TTRN =

Noun Type-token Ratios; TTRAdj = Adjective Type-token Ratio; TTRAdv =

Adverb Type-token Ratios; TTRV = Verb Type-token Ratios; \*p<.05, \*\*p<.01,

\*\*\*p<.005.

Type-token Ratios when they received the T-TAT then M-TAT order than Females who received the M-TAT then T-TAT order ( Males  $\underline{M} = .80$ ,  $\underline{SD} = .04$ ; Females  $\underline{M} = .75$ ,  $\underline{SD} = .04$ ).

#### Analysis of the Verb Type-token Ratio

The between-subject analysis of the Verb Type-token Ratios resulted in a significant Gender x Race x Order effect,  $\underline{F} (1,54) = 4.32$ ,  $p < .04$ . White Females produced larger Verb Type-token Ratios in response to the M-TAT then T-TAT order as opposed to the T-TAT then M-TAT order ( M-TAT then T-TAT  $\underline{M} = .71$ ,  $\underline{SD} = .22$ ; T-TAT then M-TAT  $\underline{M} = .78$ ,  $\underline{SD} = .22$ ),  $\underline{t} (32) = 3.74$ ,  $p < 0.05$ .

The within-subject analysis of the Verb Type-token Ratios resulted in a significant effect for the Test x Race interaction,  $\underline{F} (1,54) = 7.55$ ,  $p < .01$ , in which African-American subjects produced larger Verb Type-token Ratios in response to the T-TAT ( $\underline{M} = .76$ ,  $\underline{SD} = .02$ ) than the M-TAT ( $\underline{M} = .73$ ,  $\underline{SD} = .02$ ),  $\underline{t} (32) = 2.88$ ,  $p < .05$ . Additionally, a significant Test x Order effect was found,  $\underline{F} (1,54) = 4.69$ ,  $p < .05$ . However, the analysis of simple effects did not yield any significant differences.

#### Analysis of the Adverb Type-token Ratio

The between-subjects and within-subject analysis of the Adverb Type-token Ratios did not yield any significant differences.

#### Analysis of the Initial Session Data

The analysis of variance was also conducted on the data produced during the subject's initial session. This analysis was conducted due to the number of order effects found in the analysis of both of the subjects' session and the analysis involved the examination of each of the dependent variables for main effects and interaction effects for Gender, Race, and Test for the



subjects initial session responses only. Additionally, Verbal Facility Test scores and Family Income were treated as covariates.

#### Analysis of Word Count Variables

The analysis of variance for Total Word Count of the initial session data yielded the results in Table 3. A significant effect for Gender was found on the Total Word Count,  $F(1,54) = 10.49, p < .005$ , with Female subjects writing stories with more words in their initial session than Male subjects (Female  $M = 584.24, SD = 5.54$ ; Male  $M = 481.85, SD = 5.54$ ). Also, a significant effect for the Verbal Facility Test score was found on the Total Word Count,  $F(1,54) = 4.60, p < .05$ , indicating that as verbal facility increases so does the total word production in responses.

On the dependent variable of Noun Count Totals, the analysis did not produce significant results. In the analysis of the Adjective Count Totals a significant effect was found for Gender,  $F(1,54) = 5.72, p < .05$ , with Female subjects producing more adjectives in their responses than Male subjects (Female  $M = 63.41, SD = .81$ ; Male  $M = 52.31, SD = .81$ ). Additionally, a significant effect for the Verbal Facility Test score was found,  $F(1,54) = 6.64, p < .05$ , which indicates that more verbally facile people produced more nouns in their initial session responses.

The Verb Count Total analysis resulted in a significant effect for Gender,  $F(1,54) = 10.26, p < .005$ , with Females using more verbs than Males (Females  $M = 152.06, SD = 1.54$ ; Males  $M = 123.81, SD = 1.54$ ). Also, a significant effect was found for the Verbal Facility Test score,  $F(1,54) = 4.78, p < .05$ , which indicates that verb production in responses increases as verbal facility increases.

Table 3

Analysis of Covariance for the  
Initial Session Data for the Word Counts

	Total	Noun	Adj	Adv	Verb
Gender	10.49***	n.s.	5.72*	10.26***	n.s.
Race	n.s.	n.s.	n.s.	n.s.	n.s.
Test	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Race	n.s.	n.s.	n.s.	n.s.	4.38*
Gender*Test	n.s.	n.s.	n.s.	n.s.	n.s.
Race*Test	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Race*Test	n.s.	n.s.	n.s.	n.s.	n.s.
Verbal	4.60*	n.s.	6.64*	4.78*	n.s.
Faminc	n.s.	n.s.	n.s.	n.s.	n.s.

NOTE: Gender = female or male subjects; Race = African-American or white subjects; Test = form of TAT (T-TAT or M-TAT); Verbal = the Verbal Facility Test Faminc = family income; Total = the Total Word Count; Noun = the Noun Count Totals; Adj = the Adjective Count Totals; Adv = the Adverb Count Totals; Verb = the Verb Count Totals; \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .005$ .

The analysis of the Adverb Count Total yielded a significant effect for the Gender x Race interaction,  $F(1,54) = 4.38$ ,  $p < .05$ . White Males produced significantly more adverbs than African-American Males (White  $M = 21.42$ ,  $SD = .55$ ; African-American  $M = 14.47$ ,  $SD = .55$ ),  $F(1,29) = 2.21$ ,  $p < .05$ . No other significant effects were found Adverb Count Total variable.

Table 4  
Analysis of Covariance of the  
Initial Session Data for the Type-token Ratios

	TTRT	TTRN	TTRAdj	TTRV	TTRAdv
Gender	n.s.	n.s.	n.s.	n.s.	n.s.
Race	n.s.	n.s.	n.s.	n.s.	n.s.
Test	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Race	n.s.	n.s.	n.s.	n.s.	n.s.
Gender*Test	n.s.	n.s.	n.s.	n.s.	n.s.
Race*Test	n.s.	n.s.	5.77*	n.s.	n.s.
Gender*Race*Test	n.s.	n.s.	n.s.	4.71*	n.s.
Verbal	n.s.	n.s.	n.s.	n.s.	n.s.
Faminc	n.s.	n.s.	n.s.	n.s.	n.s.

NOTE: Gender = female or male subjects; Race = African-American or white subjects; Test = form of TAT (T-TAT or M-TAT); Verbal = the Verbal Facility Test score; Faminc = family income; TTRT = the Total Type-token Ratios; TTRN = the Noun Type-token Ratios; TTRAdj = the Adjective Type-token Ratios; TTRAdv = the Adverb Type-token Ratios; TTRV = the Verb Type-token Ratios; \* $p < .05$ .

#### Analysis of Type-token Ratios

The analysis of the Type-token Ratios for the subject's initial session responses yielded the results in Table 4. The results of the analysis of the Total Type-token Ratio and the Noun Type-token Ratio yielded no significant effects. The Adjective Type-token Ratio yielded a significant interaction effect for Race x Test,  $F(1,54) = 5.77$ ,  $p < .05$ , with White subjects producing large Adjective

Type-token Ratios to the T-TAT as opposed to the M-TAT (T-TAT  $M = .82$ ,  $SD = .005$ ; M-TAT  $M = .75$ ,  $SD = .005$ ).

The Verb Type-token Ratio results yielded a significant Gender x Race x Test interaction,  $F(1,54) = 4.71$ ,  $p < .05$ . White Males produced larger Verb Type-token Ratios on the M-TAT than T-TAT order than African-American subjects (White Males  $M = .78$ ,  $SD = .005$ ; African-American Males  $M = .70$ ,  $SD = .005$ ). No significant effects were found on the Adverb Type-token Ratio.

#### Pearson Correlation

Correlations were computed for the word count and type-token dependent variables for the Murray and Thompson responses to see to what degree these are equivalent or parallel forms. The results are shown on Table 5 and Table 6.

The correlations between the Thompson word count variables and the Murray word count variables range from .46, for the Adjectives Word Count Total, to .69, for the Noun Count Total. The correlations of the Thompson type-token ratios to the Murray type-token ratios had a range from .00, for Adverbs, to .53, for Verbs. The Adverb Type-token Ratio correlation was extremely low and this appears to be due to a restriction of range in responses on this variable.

Table 5

Pearson Correlation Matrix for  
the Word Count Variables

---

	MWCT	MNCT	MADJCT	MVCT	MADCT
TWCT	0.65	-	-	-	-
TNCT	-	0.69	-	-	-
TADJCT	-	-	0.46	-	-
TVCT	-	-	-	0.50	-
TADCT	-	-	-	-	0.55

---

NOTE: MWCT = Murray Word Count Total; MNCT = Murray Noun Count Total; MADJCT = Murray Adjective Count Total; MADCT = Murray Adverb Count Total; MVCT = Murray Verb Count Total; TWCT = Thompson word count total; TNCT = Thompson Noun Count Total; TADJCT = Thompson Adjective Count Total; TADCT = Thompson Adverb Count Total; TVCT = Thompson Verb Count Total.

Table 6

Pearson Correlation Matrix for  
the Type-token Ratios

	MTTRT	MTRRN	MTTRADJ	MTTRV	MTTRAD
TTTRT	0.48	-	-	-	-
TTTRN	-	0.47	-	-	-
TTTRADJ	-	-	0.36	-	-
TTTRV	-	-	-	0.53	-
TTTRAD	-	-	-	-	0.00

NOTE: MTTRT = Murray type-token ratio for total number of words;  
MTRRN = Murray type-token ration for the nouns; MTTRADJ = Murray type-  
token ratio for the adjectives; MTTRAD = Murray type-token ratio for the  
adverbs; MTTRV = Murray type-token ratio for the verbs; TTTRT = Thompson  
type-token ratio for the total number of words; TTTRN = Thompson type-token  
ratio for the nouns; TTTRADJ = Thompson type-token ration for adjectives;  
TTTRAD = Thompson type-token ratio for adverbs; TTRV = Thompson type-  
token ratio for verbs.

## DISCUSSION

It was hypothesized that African-American subjects would produce quantitatively more productive protocols in response to the T-TAT than the M-TAT and that white subjects would produce quantitatively more productive protocols in response to the M-TAT than the T-TAT. Only very slight support for this hypothesis was found in this study.

Analysis of the data for both of the subject's session showed support for the hypothesis only on the Verb Type-token Ratio where there was a significant Test x Race interaction. On this interaction, African-American subjects produced larger Verb Type-token Ratios in response to the T-TAT than the M-TAT. However, if alpha inflation is corrected for by using the Bonferroni method this interaction is not significant. No other Test x Race interaction were found, on the nine other dependent variables, to support the hypothesis.

Two other Test x Race interaction were found. One on the Adjective Type-token Ratio on the analysis of both sessions and the other on the Adjective Type-token Ratio for the initial session data. However, in these cases the results were contrary to the hypothesis. The simple effects analysis produced no significant effects for the Adjective Type-token Ratio for both sessions but the pattern was that African-Americans were more productive on the M-TAT than the T-TAT and the converse was true for White subjects. On the initial session Adjective Type-token Ratio White subjects were found to be significantly more productive on the T-TAT than the M-TAT. However, neither of these interactions were significant when alpha inflation was adjusted.

Gender effects were found on the word count variables for both the initial session analysis and the analysis of both sessions. In each analysis, it was found that females were more productive than males for the Total Word Count, Adjective Word Count, and the Verb Word Count. However, none of the type-token ratios resulted in a significant effect for Gender. This indicates that, while females tend to write longer stories than males, females do not use a greater variety of words, per part of speech, for a 100 word segments. Therefore, while female tend to write longer stories than males they do not produce responses with greater vocabulary diversity or vocabulary richness than males. Additionally, when the analysis is adjusted for alpha inflation only the Gender effect on the Total Word Count was significant.

In the overall evaluation of the results, little support was found for the hypothesis that African-American subjects would produce more productive responses to the T-TAT than the M-TAT and that White subjects would produce more productive responses on the M-TAT than the T-TAT. When null results are obtained, the question arises as to the possibility that true differences were not detected by the statistical analysis. This is termed the beta error or falsely failing to reject the null hypothesis. To assess for the beta error, a power analysis was conducted. It yielded a power of .70. This means that beta is .30, a respectable value, indicating a fair degree of confidence that the hypothesis of differences was not falsely rejected.

The absence of significant differences in responses to the M-TAT and the T-TAT when using the liberal alpha level of .05 or the more conservative adjusted alpha level of .003 and the moderate correlations between the Murray variables and the Thompson variables indicates that there is little support for the notion that the M-TAT is more applicable for White subjects than for African-



American subjects in eliciting verbally more productive protocols. Instead, these results offers some support for the acceptability of viewing these tests as equivalent forms which are equally applicable for whites and African-Americans. However, there were some limitations to this study.

One of the limitations of this study is that it only test responses to six TAT cards. Additionally, the cards selected for this study were chosen in part because the T-TAT versions were among the best replications of the M-TAT, with only the race of the figures being altered. In some of the other T-TAT cards which were not used in the study, alterations in expressions and posture have been made. Therefore, the results found in this study may not necessarily apply to the entire T-TAT. Because of the discrepancies in the drawing of the T-TAT picture, it may be necessary to have the T-TAT pictures redrawn where only the race of the figures is altered.

Another limitation of this study is that only a quantitative analysis of verbal production was conducted. This is one means of measuring productivity on the TAT, however, one must keep in mind that this test was designed as a measure of personality and not language production. A more qualitative analysis of identification and production on the T-TAT and the M-TAT is necessary. One means of conducting this type of analysis would be to examine the needs expressed in the response, such as need for Sex, need for Achievement or need for Autonomy. However, one can not simply assume that the similarity of the stimulus picture to the subject will facilitate identification and projection on a more qualitative basis. This issue needs to be examined further. The creation a TAT which depicts the racial diversity of the subject's environment is an alternative which should be examined.

## REFERENCES

- Anastasi, A. (1988). Psychological Testing. New York: Mac Millan Publishing.
- Andolina, C (1980). Syntactic maturity and vocabulary richness of learning disabled children at four age levels. Journal of Learning Disabilities, 13, 373-377.
- Archer, R. P. (1992). Review of the Minnesota Multiphasic Personality Inventory - 2. In J. J. Kramer & J. C. Conoley (Eds.), The Eleventh Mental Measurements Yearbook (pp. 558-562). Lincoln, NE: The University of Nebraska Press.
- Bailey, B. E., & Green, J. (1977). Black Thematic Apperception Test stimuli material. Journal of Personality Assessment, 41, 25-50.
- Bellack, L., & Bellack, S. (1965). The Children's Apperception Test. New York: Larchmont.
- Bellack, L. (1993). The T.A.T., C.A.T., and S.A.T. in Clinical Use (5th ed.) Boston, Allyn & Bacon.
- Bellack, L., & Bellack, S. (1973). The Senior's Apperception Test. New York: Larchmont.
- Breland, H.M. (1979). Population validity and college entrance measures. (College Board Research Monographs No. 8) New York: College Entrance Examination Board.
- Brenner, C. (1973), An Elementary Textbook of Psychoanalysis. New York, Dell Publishing Group Inc.
- Child, I. L., Frank, K. F., & Storm, T. (1950). Self-ratings and the TAT: their relations to each other and to childhood background. Journal of Personality, 25, 96-114.

- Chotlos, J. W. (1944). Studies in language behavior: A statistical and comparative analysis of individual written language samples. Psychological Monographs, 56, 75-111.
- Chowdhury, U. (1960). An Indian modification of the Thematic Apperception Test. Journal of Social Psychology, 51, 245-263.
- Cowan, G., & Goldberg, F. J. (1967). Need achievement as a function of the race and sex of figures of selected TAT cards. Journal of Personality and Social Psychology, 5, 245-249.
- Cowan, G. (1971). Achievement motivation in lower class Negro females as a function of the race and sex of the figures. Representative Research in Social Psychology, 21, 42-46.
- Crablit, N. S. (1977). The verbal environment of the language-impaired child. Journal of Speech and Hearing Disorders, 42, 474-482.
- DuBois, P. (1966). Testing in Ancient China. In A. Anastasi (Ed.) Testing Problems in Perspective. Princeton, N.J.: Educational Testing Service, 1966, pp. 29-36.
- Duran, R.P. (1983). Hispanic's education and background: Predictores of college achievement. New York: College Entrance Examination Board.
- Eron, I. D., & Ritter, A. E. (1951). A comparison of two methods of administration of the Thematic Apperception Test. Journal of Consulting Psychology, 15, 55-61.
- Evans, E. C. (1969). Physiognomics in the Ancient World. Transactions of the American Philosophical Association,
- Freud, S. (1894). The Justification for Detaching from Neurasthenia, a Particular Syndrome: The Anxiety Neurosis. In Collected Papers: Vol.1. (pp.76-106). London: Hogarth Press. (1949).
- Freud, S. (1905). Psychoanalytic Notes on the Autobiographical Account of a Case of Paranoia. In Collected Papers: Vol. 3. (387-470). London: Hogarth Press. (1949).
- Gordon, M.A. (1953). A study of the applicability of the same minimum qualifying scores for technical school to white males, WAF, and Negro males. Lackard Air Force Base, TX: Personnel Research Laboratory.

- Graham, J. R. (1993). MMPI-2: Assessing Personality and Psychopathology - 2nd Edition, Oxford: Oxford University Press.
- Harrison, P. L., Kaufman, A. S., Hichman, J. A., & Kaufman, N. L. (1988). A survey of tests used for adult assessment. Journal of Psychoeducational Assessment, 6, 188-198.
- Hunter, J. E., Schmidt, F. L., & Rauschenberger, J. M. (1984). Methodological, statistical, and ethical issues in the study of bias in psychological tests. In C. E. Renolds & R. T. Brown (Eds.) Perspectives on Bias in Mental Testing. (pp. 41-99). New York: Plenum.
- Hurley, J. R. (1955). The Iowa Picture Interpretation Test: A multiple choice variation of the TAT. Journal of Consulting Psychology, 19, 372-376.
- Jensen, A. R. (1980). Biases in Mental Testing. New York: The Free Press.
- Johnston, R. A. (1957). A methodological analysis of several revised forms of the Iowa Picture Interpretation Test. Journal of Personality, 25, 283-293.
- Kaplan, R. M., & Saccuzo (1989). Psychological Testing: Principles, applications, and issues. (2nd ed.) Pacific Grove, CA: Brooks and Cole.
- Katz, E., Russ, R., Overholser, J. (1993). Sex differences, sex roles, and projection on the TAT: Matching stimulus and examinee gender. Journal of Personality Assessment, 60, 186-191.
- Keppel, G. (1982). Design and Analysis: A Researcher's Handbook (2nd ed.). Englewoods Cliffs, NJ: Prentice-Hall Inc.
- Korchin, S. J., Mitchell, H. E., & Meltzoff, J. (1950). A critical evaluation of the Thompson's TAT. Journal of Projective Techniques, 14, 445-452.
- Lee, S. G. (1953). Manual of a Thematic Apperception Test for African Subjects. Pietermarzburg, South Africa: University of Natal Press.
- Lefkowitz, J. & Fraser, A. W. (1980). Assessment of achievement and power motivation of blacks and whites, using a black and white TAT with black and white administrators. Journal of Applied Psychology, 65, 6, 685-696.
- Libby, W. (1908). The imagination of adolescents. American Journal of Psychology, 19, 249-252.

- Light, B. H. (1955). A further test of the Thompson TAT rationale. Journal of Abnormal and Social Psychology, 51, 148-150.
- Lindsay, J. (1971). Origins of Astrology. New York: Barnes & Noble.
- Lindzey, G., & Heinemann, S. H. (1955). Thematic Apperception Test: Individual and group administration. Journal of Personality, 24, 34-55.
- Lindzey, G., & Silverman, M. (1959). Thematic Apperception Test: Techniques of group administration, sex differences, and the role of verbal productivity. Journal of Personality, 27, 311-323.
- Linn, R.L. (1975). Test bias and the prediction of grades in law school. Journal of Legal Education. 27, 293-323.
- Lundy, A. (1988). Instructional set and Thematic Apperception Test validity. Journal of Personality Assessment, 52, 309-320.
- Maier, M.H. & Fuchs, E.F. (1972). Effects of educational level on the prediction on training success with ACB. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Mann, M. B. (1944). Studies in language behavior: III. The quantitative differentiation of samples of written language. Psychological Monographs, 56, 41-74.
- Manschreck, T. C., Maher, B. A., & Ader, D. N. (1981). Formal thought disorder, the type-token ration, and disturbed voluntary motor movement in schizophrenia. British Journal of Psychiatry, 139, 7-15.
- MacClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). The Achievement Motive. Englewood Cliffs, NJ: Prentice-Hall.
- McClelland, L. (1974). Effects if interview-respondent race interactions on household interview measures of motivation and intelligence. Journal of Personality and Social Psychology, 29, 392-397.
- McReynolds, P. (1975). Advances in Psychological Assessment, Vol.3. San Francisco: Jossey-Bass.
- Megaree, E. (1966). A comparison of the scores of white and Negro male juvenile delinquents on three projective tests. Journal of Projective Techniques, 30, 530-535.

- Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. Archives of Neurology and Psychiatry, 34, 289-294.
- Murray, H. A. (1943). Thematic Apperception Test Manual. Cambridge, MA: Harvard University Press.
- Murray, H. A., & MacKinnon, D. W. (1946). Assessment of OSS Personnel. Journal of Consulting Psychology, 10, 76-81.
- Mussen, P. H. (1953). Differences between the TAT responses of Negro and white boys. Journal of Consulting Psychology, 17, 373-376.
- Quetelet, L. A. (1969) A Treatise of Man. Gainesville, FL: Scholars Facsimiles and Reprints.
- Reynes, R. C., Martindale, C., & Dahl, H. (1984). Lexical differences between working and resistance sessions in psychoanalysis. Journal of Clinical Psychology. 40, 733-737.
- Riess, B. F., Schwartz, E. K., & Cottingham, A. (1950). An experimental critique of assumptions underlying the Negro version of the TAT. Journal of Abnormal and Social Psychology, 45, 700-709.
- Rosen, B. (1958). The achievement syndrome: A psychocultural dimension of social stratification. In J. Atkins (Ed.), Motives in Fantasy and Society, Princeton, N.J.: Van Nostrand.
- Ross, W. (1972). Relation of a Measure of vocabulary skills on the WAIS. Psychological Reports, 30, 415-418.
- Ryan, R. (1985). Thematic Apperception Test. In D. Keyser & R. Sweetland (Eds.), Test Critiques: Vol.2
- Rubin, S. (1964). A comparison of the Thematic Apperception Test stories of two I.Q. groups. Journal of Projective Techniques, 28, 81-85.
- Sarason, B. R., & Sarason, I. C. (1958). The effect of the type of administration and sex of object on emotional tone and outcome ratings of TAT stories. Journal of Projective Techniques, 22, 333-337.
- Sattler, J. M. (1992). Assessment of Children, San Diego, CA: Publisher Inc.

- Schwartz, L. A. (1932). Social situation pictures in the psychiatric interview. American Journal of Orthopsychiatry, 2, 124-132.
- Schwartz, E. K., Riess, B. E., & Cottingham, A. (1951). A further critical evaluation of the Negro version of the TAT. Journal of Projective Techniques, 15, 394-400.
- Shore, C.W., & Marion, R. (1972). Suitability of using common selection test standards for Negro and white airmen. Lackard Air Force Base, TX: Personnel Research Division, Airforce Human Resources Laboratory.
- Stewart, A. J., & Chester, N. L. (1982). Sex differences in human social motives: Achievement, affiliation, and power. In A.
- Stewart (Ed.) Motivation and Society (pp. 172-218). San Francisco, CA: Jossey-Bass.
- Terry, D. (1952). The use of a rating scale of level of response in TAT stories. Journal of Abnormal and Social Psychology, 47, 507-511.
- Thompson, C. E. (1949). Manual for Thematic Apperception Test;Thompson Modification. Cambridge, MA: Harvard University Press.
- Thompson, C. E. (1949). The Thompson modification of the Thematic Apperception Test. Journal of Projective Techniques, 13, 469-478.
- Tomkins, S. S. (1947). The Thematic Apperception Test. New York: Grune & Stratton
- Veroff, J., Wilcox, S., & Atkins, J. (1953). The achievement motive in high school and college-age women. Journal of Abnormal Social Psychology, 48, 108-119.
- Webb, W., & Hilden, A. H. (1953). Verbal and intellectual ability as factors in projective test results. Journal of Projective Techniques, 17, 102-103.
- Worchel, F. T., Aaron, L. L., & Yates, D. F. (1990). Gender bias on the TAT. Journal of Personality Assessment, 55, 595-602.

## APPENDICES

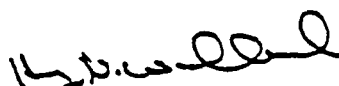


APPENDIX A

Human Research Committee Permission  
to Conduct Research



AUBURN UNIVERSITY AT MONTGOMERY  
Research and Development

TO: Ms. Ana Sullivan  
FROM: Dr. Henry N. Williford, Chairman   
Institutional Review Board for  
Research Involving Human Subjects  
SUBJECT: Approval of Protocol No. H9410.02  
DATE: October 18, 1994

Your protocol entitled "An Investigation of the Rationale Behind the Thompson's Thematic Apperception Test" was approved in an expedited review by the IRB this day, October 18, 1994. You have our best wishes for success with this project.

HW/ss

cc: Dr. Allen Hess, Faculty Advisor

## APPENDIX B

### Statement of Informed Consent

You are being invited to participate in an experiment of imagination. You will be asked to write a story in response to the pictures presented.

The administration of this project will be conducted in two sessions. Each session will last approximately forty minutes. Following the second testing session, you will be administered a word identification test which will take approximately five minutes.

The procedures involve no risk. However, you have the right to terminate your participation in the experiment at any time without penalty. As a result of your participation you will have a better understanding of the work of psychologists in a clinical setting. Also, those subjects who complete both sessions will be given a raffle ticket for a chance to win a fifty or twenty-five dollar bond. The drawing will be held following the completion of the testing. (If you are an introductory psychology student, your instructor may award extra credit for your participation).

Your responses are confidential. The experimenter will evaluate the responses but your identity will not be included in this material.

**YOU ARE MAKING A DECISION ABOUT WHETHER OR NOT YOU WILL PARTICIPATE IN THIS STUDY. YOUR SIGNATURE INDICATES THAT YOU HAVE DECIDED TO PARTICIPATE, HAVING READ THE INFORMATION PROVIDED ABOVE.**

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name

\_\_\_\_\_  
Signature of Witness



APPENDIX D

The Verbal Facility Test

Last 4 digits of Social Security #    \_\_\_    \_\_\_    \_\_\_    \_\_\_

In each row of words below, draw a line under the one word that most nearly means the same as the word on the left that is already underlined. Complete the entire test. If you don't know, guess.

SAMPLE

fight           men           like           battle       task

BEGIN HERE

1.   kill           knock       harbor       lift       slay
2.   labor       sing       help       work       love
3.   trip       tour       verse       pinch       wish
4.   turn       sail       visit       rotate       tie
5.   shine       ring       reach       pull       polish
6.   error       mistake    desire       erase       claim
7.   moist       salient     vapid       damp       tawny
8.   silent       certain    deft       quiet       artistic
9.   scarce       single     rare       recent       peril
10. jolt       time       jam       jar       sink
11. source       height     orient       fancy       origin
12. powerless    impotent   rebellious   piquant     reliant
13. little       muster     flutter       dint       diminutive
14. bill       look       care       back       beak

15.	<u>rot</u>	quaver	decompose	muddy	wither
16.	<u>prize</u>	cherish	flair	catch	bale
17.	<u>detest</u>	contrary	favor	abhor	dilate
18.	<u>clear</u>	dark	round	lucid	heavy
19.	<u>blot</u>	stigma	flame	tray	halo
20.	<u>eminent</u>	captious	illustrious	pertinent	defunct
21.	<u>quarrelsome</u>	haughty	grandiose	contentious	mediocre
22.	<u>tendency</u>	trend	gauge	pang	method
23.	<u>abstract</u>	brief	smile	begin	long
24.	<u>criticize</u>	number	reprehend	oppose	pamper
25.	<u>quality</u>	attribute	inspect	aroma	fragrant
26.	<u>wither</u>	flaunt	wizen	escort	foment
27.	<u>accidental</u>	natural	fortuitous	malaise	fictitious
28.	<u>native</u>	indecorous	normal	indigenous	salient
29.	<u>omen</u>	apparel	cartel	demon	portent
30.	<u>sacrosanct</u>	nasty	wanton	inviolable	voluble
31.	<u>transient</u>	ephemeral	convivial	adhesive	winding
32.	<u>sapid</u>	mordant	natty	lament	palatable
33.	<u>surfeit</u>	glut	harden	depose	ingrain
34.	<u>recreant</u>	iconoclast	hypocrite	apostate	neophyte

## APPENDIX E

### Word Classification

The following rules were used in determining what constituted a word.

1. Each group of letters separated by spaces on both sides from adjacent groups of letters was counted as a word, even though it might be part of a place name, as in Des Moines (two words), and initial, as in James A. Brown (three words), or a neologism coined by a subject.
2. Any number was counted as one word; for example, 125 was tabulated as one word.
3. A hyphenated word was counted as one word.
4. Each time a word was used as a different part of speech it was counted as a different word. For example, mine as a noun and mine as a pronoun were tabulated as two different words.
5. Common nouns and proper nouns having identical spellings were thrown together. For example, the two words Storm Lake were tabulated under the common nouns storm and lake.
6. Contractions were divided into two words, for example, didn't was changed to did not and tabulated as two words.

7. Abbreviations which stood for only one word were written out and tabulated as the complete word. Abbreviations which consisted of more than one unit, as for example M.D. and Ph.D., were tabulated as one word.
8. Misspellings, when it was apparent that they were misspellings and not neologisms, were corrected and tabulated as corrected.

## APPENDIX F

### Part of Speech Classification

The following rules indicate the means by which the part of speech of a word was determined.

**Nouns**-all regularly known common and proper nouns and gerunds which the dictionary recognizes as nouns.

**Verbs**-simple verbs, participles plus auxiliaries, gerunds, and participles unless the dictionary recognizes them as nouns and adjectives, as the case may be.

**Adjectives**-regular classification, and any verb form (i.e. participle) which the dictionary recognizes as as adjective.

**Adverbs**-regular classification.