THE JAMES PERSONALITY INVENTORY

Thesis for the Degree of M. A.

MICHIGAN STATE UNIVERSITY

JAMES J. MORIARTY, JR.

1969

ABSTRACT

THE JAMES PERSONALITY INVENTORY

A MULTIPLE CHOICE TEST OF THE NEED FOR ACHIEVEMENT

By

James J. Moriarty, Jr.

The James Personality Inventory (JPI) was an attempt to develop a multiple choice measure of n Ach. The advantage of a multiple choice test is that it is easier to administer and score than the McClelland TAT measure of n Ach.

An 87 item True/False test (JPI #1) was created from the n Ach literature. It was administered with the TAT to 30 Ss in Research Session #1. The correlation between the TAT and JPI #1 scores was .18 (NS). The internal consistency of the test was high, .77.

Four subscales, JPI #2, #2a, #3, and #4, were empirically generated from this data. These scales were tested on the data of Research Session #1. JPI #2 scores correlated .73 with the TAT scores. The internal consistency of the scale was .70. JPI #2a correlated .64 with the TAT and showed an internal consistency of .65. JPI #3 and JPI #4 did not correlate significantly with the TAT criterion but the internal consistency of the scales was high, .67 and .88, respectively.

Research Session #2 was an attempt to cross-validate the above scales. 94 Ss were administered the TAT and the

87 item JPI. The correlation between the TAT and JPI #1 scores was .19 (NS). The internal consistency of JPI #1 remained high, .76. The correlation between the TAT and JPI #2 scores was significant and somewhat higher, .25 ($p < .02$). However, the internal consistency of this scale dropped to .32. JPI #2a correlated with the TAT .22 ($p < .03$) and had a higher internal consistency, .49. JPI #3 correlated with the TAT .21 ($p < .04$) with an internal consistency of .40. The correlation between the TAT and JPI #4 was .15 (NS). The internal consistency remained high, .76.

The results are consistent with other attempts (Mehrabian, 1968) to create a multiple choice measure of n Ach. The development of an alternate measure of n Ach may not be the fault of the multiple choice testing, but rather the impracticality of using the TAT as the validating criterion. The unreliability of the McClelland TAT renders it an unstable criterion and consequently attenuates the size of validation correlations. A research strategy that employs a clear behavioral criterion of n Ach would prove more fruitful.

Evidence from Mehrabian's and the author's work, as well as internal consistency measures on the TAT, suggest that n Ach is a multidimensional factor. Future attempts to construct a multiple choice test of n Ach should be

developed on the basis of a multidimensional personality model of the high achiever.

Approved James S. Kleiman
Committee Chairman

Date 12/5/69

THE JAMES PERSONALITY INVENTORY

A MULTIPLE CHOICE TEST OF THE NEED FOR ACHIEVEMENT

By

James J. Moriarty, Jr.

A THESIS

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

MASTER OF ARTS

Department of Psychology

1969

In memory of Aunt Mamie,

always a model of achievement for me

ACKNOWLEDGEMENTS

I wish to gratefully acknowledge the professional and moral support that I have received from Professor Uleman who has worked with me on every phase of this research. His open posture informed me that the success or failure of the research was my own, but his insight and experience were always present. An appreciation for his approach has made me a better student and teacher.

My appreciation of Professor Aronoff is as much for his shared views of psychology as it is for his criticisms and suggestions to my research. His depth view of behavior is one that has forced me to reconsider a number of issues in social science. Time will tell where such a perspective will lead me.

Professor Hanley's protean understanding of issues, academic and otherwise, has always provided a critical conscience for me. I sometimes suspect that he understands my motivations better than I. His advice on statistical methodology has been invaluable to the creation of this thesis.

Jerry Gilmore's counsel with computer programming has greatly facilitated the progress of my work as well as given me an appreciation of the age in which I live. Bob Gurney's advice with statistics has proved invaluable and enriching. And finally, Dianne Wilhelm's clerical help and friendship have made the tedious work of the thesis more enjoyable than it would have been alone.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

THE JAMES PERSONALITY INVENTORY

by

James J. Moriarty, Jr.

I - INTRODUCTION

The Problem: Developing a Multiple Choice Test of n Ach

This research was an attempt to develop a multiple choice measure of the need for achievement (n Ach). The James Personality Inventory (JPI) was constructed to be a reliable and easy-to-score test that would identify the same Ss that McClelland's (1953) projective measure of n Ach identifies.

There are several psychometric and pragmatic disadvantages to McClelland's Thematic Apperception Test (TAT). Briefly, the problems with the TAT are: (1) the administration and scoring of the TAT require considerable training; (2) the scoring of the TAT is time consuming; (3) the TAT has low test-retest reliability.

The nature of multiple choice testing obviates problems of scorer training as well as facilitates the speed of test scoring, particularly if computer time is available. The problem of establishing satisfactory test-retest reliability is more complex, but ease of administration and scoring of multiple choice tests facilitates sound test construction.

Strategy of the Research

Essentially two types of experimental design may be employed to select and validate items for a multiple choice personality inventory. The first method is that of motive arousal and subsequent item selection. Items are administered to one group of Ss under neutral conditions, and to another group under motive arousal conditions. Items which differentiate groups are selected for further validation studies. This method of item selection is analogous to the original methodology that McClelland (1953) employed to create the scoring system for the TAT measure of n Ach.

The second method of item selection is that of correlating items against a criterion and selecting those items which correlate with the criterion. This was the experimental design employed in this study because a consistently reliable method of arousing n Ach is not available (Klinger, 1966), and multiple choice items have been traditionally viewed as resistant to situational manipulation.

The basic strategy of the research was to create multiple choice items with face validity from the n Ach literature, and then establish their concurrent validity against a n Ach criterion. Hence if an item correlated with scores on the TAT it was selected for inclusion in the JPI.

An attempt was also made to build a scale of internally consistent items on the assumption that n Ach is essentially a unidimensional personality characteristic. That is, Ss high in n Ach would respond to multiple choice items in a consistent and similar manner.

Significance of the Achievement Motive

The initial work of McClelland and his associates at a Naval base in the 1940's has led to the uncovering of the human motive termed "need for achievement". Research on the achievement motive constitutes a significant body of literature in the field of social psychology. Major research projects have been carried out at Wesleyan University (McClelland, Atkinson, Clark, and Lowell, 1953), the University of Michigan (Atkinson, 1958), Harvard University (McClelland, 1961) as well as various non-academic research centers (e.g., Krumbolz and Farquhar, U.S. Air Force, 1957; Behavior Research Service, 1960).

Researchers have linked the motive to such varied behaviors as academic achievement (e.g., Rosen, 1956; Robinson, 1964), verbal memory tasks (e.g., Karolchuck and Worell, 1956; Sampson, 1963), parent-child relations (e.g., Winterbottom, 1958), and cognitive styles (e.g., French, 1955).

In a provocative clumination of his work on the achievement motive, The Achieving Society (1961), McClelland

applied psychological reasoning and research procedures to an analysis of economic development of civilizations. Since n Ach was considered to be a major psychological motivation of the entrepreneur, it was hypothesized that societies that encouraged the development of such personalities would prosper economically. In effect, McClelland was examining under a psychological cover one of the bases of western civilization, the Protestant work ethnic (cf., Weber's hypothesis on the relation of Protestantism to economic growth).

Despite such widespread research on the high achiever, there is still no method of identifying this personality type which is easy to administer and score. The present TAT scoring system has psychometric and pragmatic deficiencies. The development of an improved measure of n Ach is the subject of this thesis.

Characteristics of the High Achiever

We will draw a composite portrait of the high achiever to give the reader a conception of the personality type that the TAT and JPI are attempting to identify. More detailed reviews of the research and theory of the motive are contained in Brown's Social Psychology (1965, Pp. 423-476) and Byrne's An Introduction to Personality (1966, Pp. 284-338). Heckhausen's The Anatomy of the Achievement Motive (1967) contains the most detailed accounting of research and theory; it provides an indispensible introduction to researchers interested in this field of research.

Definition of n Ach

Need for achievement is defined as the desire to do well or to compete with a standard of excellence. The Achievement Goal is success in competition with some standard of excellence; engagement in competitive activity, other than pure agression, where winning or doing as well or better than someone else is the primary concern; and the subjective satisfaction in succeeding (Atkinson, 1968).

Portrait of the High Achiever

Empirical evidence has shown that people who are high in n Ach work harder at laboratory tasks; learn faster; do better work in high school with I.Q. partialled out; and seem to do their best work when it counts for their record and not when other special incentives are introduced such as pressure from the outside to do well, money prizes, or time off from work (McClelland *et al.*, 1953).

High achievers have realistic aspiration levels, whereas low achievers have relatively low or unrealistically high aspiration levels (Atkinson, 1964). They are resistant to social pressure, choose experts over friends as work partners, tend to be more active in college and community affairs, like risky occupations, perform better under longer odds, and choose moderate risks over either safe or speculative ones (McClelland *et al.*, 1953).

High achievers are better able to delay gratification than low achievers (Mischel, 1961). They are less dogmatic and less neurotic than low achievers. They have a tendency to desire social interaction with high status targets (Mehrabian, 1969). They enjoy traveling and in general are future-oriented. They conceive of time in metaphors of a flying or galloping beast (action) as opposed to a vast peaceful ocean (static) (Knapp and Garbutt, 1958).

High achievers have been indulged less by their parents during their childhood than low achievers (McClelland, 1961). They come from families in which there is stress on early self-reliance and mastery (Winterbottom, 1958). And finally, they are drawn to and do better at entrepreneurial occupations (McClelland, 1961).

The high achiever is likely to be active and successful in business enterprises because business involves risk, personal responsibility (possibilities of credit for success), and measurable results. These are characteristics that attract people with high n Ach. Several studies support the hypothesis that men with high n Ach are predisposed to enter business rather than the professions (Behavior Research Service, 1960; Veroff et al., 1960). The n Ach scores of business managers, representative of the entrepreneurial class, were significantly higher than the n Ach scores of professionals,

representative of men of equal abilities and background. Such findings have led Brown to suggest that whereas McClelland originally attempted to find a generalized measure of achievement motivation, "his measure has turned out to be chiefly a measure of motivation for business or economic achievement."[1]

This large collection of research from various psychological laboratories gives us a composite portrait of a person with high n Ach as someone who is concerned about accomplishment, or who wants to do well at what he undertakes; he is energetic; and he is predisposed toward innovations, toward working at tasks which are not safe and traditional but that involve a moderate degree of risk--perhaps only then he can feel subjective satisfaction from succeeding.

The initial problem of the research was to create a pool of items that would "tap", or identify Ss with, the above described personality characteristics. The details of the creation of the item pool are contained in the "Design and Procedure" chapter.

---

[1]Roger Brown, Social Psychology. New York; The Free Press, 1965. p. 473.

Measuring Need for Achievement

The n Ach TAT

Supported by Freud's original work on dreams, years of psychoanalytic experience, and the clincal success of Murray's Thematic Apperception Test, McClelland decided that an excellent place to look for indicators of the achievement motive was in fantasy (McClelland, et al., 1953). McClelland et al. developed a variation of Murray's TAT based "on the notion from experimental studies of animals that motives could be experimentally aroused and their intensity controlled by manipulating arousal conditions."[1] A four picture series of cards depicting characters involved in potentially achievement oriented situations are administered to a S. For example, one of the n Ach cards pictures a student sitting at his desk with an open book in front of him. Whereas a subject with low n Ach would describe the boy as "day dreaming" about his girl friend or some non-achievement oriented matter, a subject with high n Ach would describe the student as dreaming about how great a physician he will be one day, the great scientific discoveries he will make, etc. The subject is not describing or perceiving what is in the picture but rather projecting his own personality, including his n Ach, into the picture.

---

[1] D.C. McClelland, J.W. Atkinson, R.A. Clark, and E.L. Lowell, The achievement motive. New York: Appleton-Century-Crofts, 1953. p. 3.

the subject's score on n Ach is the simple count of the number of achievement related images in the story he has written.

Smith and Feld (1958) reviewed the published research that used the n Ach system of content analysis. Their survey revealed that interjudge scoring reliabilities of .66, .89, .95, .89, .91, .80, .96, and .82 and score-rescore reliabilities of .95, .93, .95, .88, and .94 have been reported for n Ach. This degree of inter-scorer reliability and score-rescore reliability is satisfactory for research purposes.

Psychometric Disadvantages of the TAT

McClelland's TAT has the psychometric disadvantage of low rest-retest reliability. Product-moment coefficients of equivalence-stability for three to five week intervals have ranged from .22 on college students (Lowell, 1950) to .64 on high school students (Morgan, 1953). Although Haber and Alpert (1958) have found a somewhat satisfactory correlation of .54, with two parallel sets of pictures, the majority of reported reliability tests of the TAT are in the range of .26 (Krumbolz and Farquhar, 1957) and .22 (Birney, 1959). Such low test-retest reliability casts severe doubt on the stability as well as the possible validity of the measure.

At least one researcher (Heckhausen, 1967) has argued that whereas the test is not suitable for

psychodiagnostic purposes, its degree of discrimination is sufficient for experimental research. Consequently a large number of the n Ach studies employ only extreme scorers in their subject pool.

Alternate Measures of n Ach

Possible substitute tests of achievement such as self-report measures (e.g., De Charms, Morrison, Reitman, and McClelland, 1955; the measure of achievement in the Edwards Personal Preference Schedule, Edwards, 1959), or more structured versions of the TAT (e.g., Iowa Picture Interpretation Test, Hurley, 1955; French, 1958) do not correlate with the TAT measure (e.g., DeCharms, Morrison, Reitman, and McClelland, 1955; Williams, 1955; Marlowe, 1959; Atkinson and Litwin, 1960; or Barnette, 1961). Hence for this and other reasons they are rarely employed in studies of achievement.

Heckhausen (1963) has developed a six picture TAT measure of n Ach using "Hope of Success" and "Fear of Failure" as two dimensions of n Ach. McClelland's measure correlates between .40 and .70 with ("Hope of Success" + "Fear of Failure"). Unfortunately, the Heckhausen measure also suffers from problems of administration and scorer training as well as time consuming protocol scoring.

Mehrabian (1968) undertook construction of a multiple choice measure of n Ach simultaneous to the initiation of this research project. He constructed

separate male and female scales of "The Tendency to Achieve." Each scale consists of 34 items which are designed to discriminate high versus low achievers. For example, the difference in time perspective of high versus low achievers was considered. High achievers have a more distant future perspective than low achievers (e.g., Ricks and Epley, 1960). Mehrabian attempted to tap this perspective by the following item, "I think more of the future than of the present and past." (True)

The reliability of the test is good. A ten week test-retest yielded a product-moment correlation coefficient of .78 for the male scale and a coefficient of .72 for the female scale. The TAT correlated .28 ($N=110$, $p<.01$) with the male achievement scale and .17 ($N=111$, $p<.05$) with the female scale. Further investigations of the validity of Mehrabian's scale indicate that it is a useful measure of n Ach. Following an actual course examination, male high achievers who were identified by the Mehrabian scale showed a stronger Zeigarnik effect (i.e., recall a greater percentage of the failed items relative to the percentage of passed items) than male low achievers (Weiner, Johnson, and Mehrabian, 1968). The Mehrabian scale correlated .64 ($N=110$, $p<.01$) with Rotter's (1966) internal-external control scale, indicating that high achievers perceive themselves as having a greater degree of control over

the events which influence their lives (Mehrabian, 1968). These findings are consistent with evidence which differentiates high from low achievers.

Rationale of a Multiple Choice Measure of n Ach

Briefly stated, the problem explored in this thesis is as follows: despite the fact that there is widespread research on the achievement motive, the TAT method of identifying high achievers has some serious disadvantages. The JPI was developed as an alternate test of n Ach because a multiple choice test is easy to administer and score. Since the TAT is the primary means of measuring n Ach, it was necessary to employ it as the criterion of n Ach. The primary objective was to establish the JPI's validity against the criterion of the TAT. Hence scores on the JPI should correlate highly with scores on the TAT.

II - DESIGN AND PROCEDURE OF THE RESEARCH

Strategy of Item Creation and Selection

The strategy of the project was to review n Ach research and theory and create an item pool from which the JPI could be developed. Items were created that attempted to tap the cluster of interrelated characteristics which distinguish high achievers from low achievers (cf., Characteristics of High Achiever, p.4). The initial keying of items was done on the basis of face validity. Hence agreement or disagreement with an item was indicative of a behavioral disposition which has been found to differentiate high versus low achievers. Final keying of items was determined by the empirical performance of the item. If an item distinguished between low and high scorers on the TAT it was retained and keyed in the appropriate direction.

Creation of the Item Pool

300 items were created from an extensive review of the n Ach literature. For example, "I prefer moderate risks (3 out of 6) to either safe (5 out of 6) or speculative (1 out of 6) risks." (Item #65). The item was keyed "True" because n Ach theory suggests that Ss high in n Ach prefer moderate risks to extremely speculative or safe risks (Atkinson, 1964). This item was eliminated from the final scale because it did not help to identify Ss who were consistently scoring high or low on the TAT measure.

The following item was keyed "False," consistent with n Ach theory, and retained throughout the study because it distinguished between Ss who were scoring high and low in the TAT criterion. "It is more important to be socially acceptable than to be a success." (item #68)

The Problem of Item Ambiguity. The problem of item ambiguity raises three issues. (1) Item ambiguity increases the probability that acquiescence sets will influence item response. (2) Item ambiguity disguises the variable that the test is measuring. (3) Item ambiguity expands the opportunity for projection to occur.

Ambiguity in items on the JPI was considered undesirable. This approach was employed in order to avoid acquiescent response sets. "Scores on personality tests will be influenced by acquiescent tendencies only to the degree to which the test contains items which are difficult or ambiguous and thus provide opportunities for acquiescent responses to occur."[1] If an item is somewhat unclear, so that the S is in doubt about the appropriate response, or if he is indifferent about taking the test and is not willing to "figure out the meaning" of the item, response sets become important.

---

[1]Samuel Messick and John Ross, Measurement in Personality and Cognition. New York: John Wiley, 1962. p. 106.

Hence the items were made clear in meaning, brief in wording. This was designed to keep ambiguity to a minimum and reduce the operation of acquiescent response sets.

It appears that the Ss were unaware of the personality variable being measured. Only 4% of the Ss guessed with any accuracy the purpose of the testing sessions. Hence it was not necessary to employ ambiguity in items to disguise the item intent.

Although item ambiguity would provide an increased opportunity for projection to occur, this factor was considered less desirable than avoiding the problem of acquiescent response sets that item ambiguity produces. For the above reasons, ambiguity in items was kept to a minimum.

Reduction of Item Pool

The 300 items were divided arbitrarily into two pools of 150 items. The item pools were administered to two groups of 75 males each. Items were eliminated against three criteria. (1) If 95% or more of the S answered "True" or "False" to the item it was dropped. The item did not discriminate between Ss. (2) Only one item was retained if several items seemed to be asking the same question and had approximately the same index of difficulty.[1] (3) Items were eliminated that Ss

[1]Index of difficulty is the proportion of the total group who got the item wrong.

consistently reported as ambiguous or too difficult to understand. The 86 retained items constituted JPI #1 (cf., Appendix A for JPI #1). According to the à priori method of keying, 44 items were keyed "True"; 43 items were keyed "False."

Research Session #1: Empirical Selection and Keying of Items

Procedure of Research Session #1

Subjects. The Ss were 30 male undergraduate volunteers who were enrolled in Introductory Psychology courses at Michigan State University. Course credit was given to Ss who participated in experiments. Two data gathering sessions, October 3 and 5, 1967, were necessary to obtain the sample. Data was also gathered on 44 females but was not analyzed because the primary purpose of the research was to construct a male test of n Ach.[1] The mean age of the Ss was 18.8, S.D., 1.5. The educational level was 13.7.

Administration of TAT and JPI. A screen projected n Ach TAT, picture series 2, 8, 1, 7 (Atkinson, 1958), was administered under neutral conditions (McClelland, et al., 1953). The story protocols were scored for n Ach.

---

[1]The subject call did not distinguish sex. It has been the experience of this experimenter that it is easiest to gather a sample when sex is not distinguished on subject sign-up sheets.

by a trained rater[1] whose percentage agreement in scoring achievement imagery with expert scorers was in the mid .90's.

Following the administration of the TAT, the 87 item JPI #1 was administered. Ss selected one of four choices to describe how they felt about an individual item.

TRUE:	(1)	This is very true of me (describes me very well).
	(2)	This is often true of me (describes me pretty well).
FALSE:	(3)	This does <u>not</u> describe me very well.
	(4)	This does <u>not</u> describe me at all.

In the item analysis choices (1) and (2) were treated as TRUE. Choices (3) and (4) were counted as FALSE. The four choice selection was employed to make it easier for the Ss to answer the items. Repeated data analysis showed that there was no significant difference in item performance if the response was scored in the four choice categories. Since the collapsed categories proved easier to work with the two choice categories were used in data analysis.

Three questions were printed at the end of the JPI.

Did you enjoy taking this test?
yes <u>60%</u> neutral <u>33%</u> no <u>7%</u>
What do you think was the purpose of the "picture test?"
What do you think was the purpose of the JPI?

---

[1]The trained rater was the author. His scoring proficiency is reported in Appendix B.

Only 3% reported with any accuracy the purpose of either test.[1]

Results of Research Session #1

TAT and JPI #1. The range of TAT scores on the 30 Ss was -4 to 15. The mean was 3.1, S.D., 4.1. Cronbach's (1951) coefficient alpha, an estimate of the internal consistency of the TAT, was .23. All 87 items were used in computing the JPI scores. The items were keyed consistent with n Ach theory. Hereafter this *à priori* keying of items will be referred to as JPI #1. The internal consistency of the scale, as measured by the Kuder-Richardson Formula 20, was .77. The correlation between the TAT and JPI #1 scores was .18(N=30, NS). Table 1 on page 20 reports this and all subsequent data.

Criteria for Item Selection on JPI #2, #2a, #3, & #4

Four scales were developed from the data gathered in Research Session #1. They were constructed on the basis of four sets of criteria for item selection.

---

[1]It appeared that the Ss were unaware of the personality dimension being measured and consequently were not influenced by an artificially induced n Ach set. This may be a problem with Mehrabian's test since it is named The Resultant Achievement Motivation Scale (RAM). That is, merely informing the S that he is taking a test that is concerned with achievement may induce a mental set that will affect his performance. The induced mental set was the basis for McClelland's original research strategy.

Items on JPI #2 were selected if an item showed a biserial correlation of .20 or greater with the TAT scores. Items were selected for JPI #2a, a subscale of JPI #2, if the mean TAT score of the Ss who got the item right was 1.5 higher than the mean TAT score of the Ss who missed the item. Items on JPI #3 were selected from their performance on a contingency table. If 75% of the Ss who scored above the median TAT score responded to an item in a consistent manner, the item was selected. Items on JPI #4 were selected on the basis of item internal consistency.

Construction of JPI #2

The selection criterion for items on JPI #2 was a biserial correlation[1] of ±.20 or greater with the TAT scores. That is, if an item correlated with the criterion measure it was selected. For example, the item analysis showed that item 5 had a biserial correlation of .42 with the TAT scores. Hence it was selected.

[1] Two correlations are commonly used as an index to decide whether to keep or discard an item. The first is the biserial correlation. This is the correlation between the S's performance on an item (right or wrong) and his test score. It assumes that the distribution of test scores is normal and that there is some underlying normal distribution of the right-wrong dichotomy. The point biserial correlation is also a correlation between the S's performance on an item (right or wrong) and test score. It assumes that the test score distribution is normal and that the division on item performance is a natural dichotomy. The biserial correlation was used as the criterion for item selection because the degree of agreement or disagreement with an item was considered to be a continuous normally distributed variable, i.e., Ss varied in the degree an item described them.

Table 1 is a summary of the correlations and test statistics of the two data gathering sessions.

Table 1.

JPI Correlations with the TAT and Summary Test Statistics

Research Session #1 N = 30	JPI #1 87 items	JPI #2 43 items	JPI #2a 27 items	JPI #3 37 items	JPI #4 59 items
Range	43-76	18-39	12-23	16-35	21-54
Mean & S.D.	57.6±8.2	26.4±4.9	18.2±3.4	22.7±4.5	40±8.7
Internal Consistency (Kuder-Rich. 20)	.77	.70	.65	.67	.88
TAT & JPI Correlations	.18	.73**	.64**	.30	.12
Research Session #2 N = 94					
Range	29-73	16-33	8-24	13-33	16-49
Mean & .S.D.	58±8	26.6±3.3	18.2±2.9	22.6±3.4	39.3±6.6
Internal Consistency (Kuder-Rich. 20)	.76	.32	.49	.40	.76
TAT & JPI Correlations	.19	.25*	.22*	.21*	.15

*$p < .05$ **$p < .01$

The critical validating correlations are underlined. Note that JPI #2, #2a, and #3 correlated significantly with the TAT criterion. The internal consistency of JPI #4 is high but the scale is not related to the TAT measure.

Item 74 showed a negative correlation of .81. It was selected but its keying was reversed from the à priori keying of False to True.

43 items were selected according to the biserial correlation criterion. 26 items were keyed True and 17 items were keyed False (cf., Appendix A for items). This scoring key was applied to the JPI data of Research Session #1. The internal consistency of JPI #2 as measured by the Kuder-Richardson Formula 20 was .70. The correlation between the TAT and JPI #2 scores was .73 ($N=30$, $p<.01$).

Since Scale #2 showed good internal consistency and a very strong correlation with the Ss's TAT scores it was decided to test its validity on another sample. These data are reported in Research Session #2.

Construction of JPI #2a

A 27 item subscale, JPI #2a (cf., Appendix A for items), was derived from the 43 item JPI #2. An item was selected if the mean TAT score of the Ss who got the item right was 1.5 higher than the mean TAT score of the Ss who missed the item.

27 items were selected according to this criterion. 16 items were keyed True; 11 items were keyed False. This scoring key was applied to the JPI data of

Research Session #1. The internal consistency of JPI #2a as measured by the Kuder-Richardson Formula 20 was .65. The correlation between the TAT and JPI #2a scores was .64 ($N=30$, $p < .01$).

The good internal consistency and strong correlation with the Ss's TAT scores indicated that JPI #2a should be tested on another sample. These data are reported in Research Session #2.

Construction of JPI #3

JPI #3 was constructed with items from a contingency table. An item was selected if two criteria were met in the item response pattern: (1) 75% of the Ss who scored above the median TAT score loaded in one direction on a particular item; (2) 40% of the Ss who scored below the median TAT score loaded in the opposite direction on the item. The item was keyed in the direction selected by Ss who scored high on the TAT.

For example, on item 78, 86% of the Ss who scored above the median TAT score answered False. 43% of the Ss who scored below the median TAT score answered True. The item was selected and keyed False.

37 items were selected according to this procedure of item analysis. 22 items were keyed True and 15 items were keyed False (cf., Appendix A for items). Scale #3 was applied to the JPI data of Research Session #1. The internal consistency of JPI #3 was .67. The

correlation between the S's TAT scores and their JPI #3 scores was .30 (N=30, NS).

Although JPI #3 did not correlate significantly with the TAT criterion, the internal consistency of the scale warranted testing it on another sample. These data are reported in Research Session #2.

Construction of JPI #4

The selection criterion for items in JPI #4 was item internal consistency. If the biserial correlation of an item was equal to or greater than .20 the item was included. In this instance, the biserial correlation is the correlation between the item's performance and the total JPI #1 test score. A high biserial correlation indicates that there is a tendency for those who got the item right to obtain high scores on the JPI.

If our assumption is correct that n Ach is a unidimensional personality characteristic, and our selection of items is tapping this dimension, we should build an internally consistent test. For this reason, items on JPI #4 were selected on the basis of internal consistency.

59 items were selected according to the internal consistency criterion. 32 items were keyed True and 27 items were keyed False (cf., Appendix A for items).

This scoring key was applied to the JPI data of Research Session #1. The internal consistency of JPI #4 as measured by the Kuder-Richardson Formula 20 was .88. The correlation between the TAT and JPI #4 scores was .12 (N=30, NS).

Although JPI #4 did not correlate with the TAT criterion, the strong internal consistency of the scale made it worthwhile to attempt another validating session with this set of items. These data are reported in Research Session #2.

Rationale of Research Session #1

The purpose of Research Session #1 was to select and key items empirically. It was also to build a scale of items that was internally consistent on the assumption that n Ach is a consistent behavioral predisposition. The problem was to select a wide variety of items that tap the personality syndrome described in the Introduction.

Research Session #1 produced four scales which were internally consistent. JPI #2 and #2a correlated highly, .73 and .64, with the criterion. Research Session #2 was an attempt to cross-validate them. Although JPI #3 and #4 did not correlate with the criterion, their high internal consistency warranted testing them out on another sample.

Research Session #2: Cross-Validation Session

Procedure of Research Session #2

Subjects. The Ss were 94 males drawn from the same population as the Ss in Research Session #1. Two data gathering sessions, November 27 and 29, 1967, were necessary to gather a large sample. Data was also gathered on 123 females but was not analyzed. The mean age of the Ss was 18.6, S.D., 1.8. The educational level was 13.5.

Administration of TAT and JPI. The test administration procedure of Research Session #1 was followed. 57% of the Ss reported that they enjoyed taking the tests; 30% were neutral; 13% did not enjoy the testing. 5% reported accurately that they knew what the test was measuring.

Results of Research Session #2

TAT and JPI #1. The range of TAT scores on the 94 Ss was -4 to 16. The mean was 5.35, S.D., 5.1. Cronbach's coefficient alpha was .26. The internal consistency of the scale was .76. The correlation of the TAT and JPI #1 scores was .19 (N=94, NS). These results are similar to the data of Research Session #1.

TAT and JPI #2. JPI #2, the 43 item scale constructed with items that correlated with the TAT criterion, yielded an internal consistency of only .32.

This is a drop in internal consistency from .70. The TAT and JPI #2 scores correlated .25 ($N=94$, $p < .02$). This was the critical correlation for validating the JPI. There was no significant difference between this correlation and Mehrabian's (1968) validating correlation between the RAM and TAT.

TAT and JPI #2a. JPI #2a, the 27 item subscale of JPI #2, yielded a slightly lower validating correlation with the TAT, .22 ($N=94$, $p < .03$). The internal consistency of the subscale was significantly higher, .49.[1]

The moderate size of these coefficients of internal consistency may indicate that n Ach is not a unidimensional personality characteristic. The significance of this finding is considered in the Discussion.

---

[1] The Spearman and Brown formula (adapted from Guilford, 1954, P. 391) estimates that the internal consistency of this subscale would be .60 if the same kind of items could be obtained to raise the length of the scale to 43 items.

$$r_{nn} = \frac{n \cdot r_{tt}}{1 + (n - 1)\, r_{tt}} \qquad .60 = \frac{\frac{43}{27}\,(.49)}{1 + \left(\frac{43}{27} - \frac{27}{27}\right)(.49)}$$

r_{nn} = estimated internal consistency of the lengthened test.

n = increase in test length

r_{tt} = internal consistency of original test

TAT and JPI #3. JPI #3, the 37 item scale developed from the contingency table, yielded an internal consistency of .40. This is a drop from .67. The TAT and JPI #3 scores correlated .21 (N=94, $p < .04$). Once again the size of the validating correlation was low. Although the validating correlations and internal consistency of JPI #2 and #3 are in the same range, the higher validating correlation of JPI #2 indicates that it is a better testing instrument.

TAT and JPI #4. The internal consistency of JPI #4, the 59 item scale constructed on the basis of item internal consistency, remained high, .76. The validating correlation, however, did not reach significance, .15 (N=94). The scale is obviously a measure of some consistent response pattern, but it does not validate against the TAT criterion of n Ach. Further research is needed to determine what it is measuring.

III - DISCUSSION

In the discussion, three points are considered:

(1) the unreliability of the TAT as a criterion of n Ach;

(2) the strength of the relationship between the TAT and JPI;

(3) n Ach as a multi-dimensional personality characteristic.

Unreliability of the TAT as a Criterion of n Ach

The weak validating correlations of the present research, along with the small TAT validating correlation of Mehrabian's work, suggest to this author the impracticality of developing a multiple choice measure of n Ach with the TAT as the sole criterion for n Ach. The researcher cannot be sure that the low validity correlations that he obtains are due to the lack of validity of his measure or the low reliability of the TAT. As Guilford points out, "We should obtain a very erroneous idea of how well we are doing with a selection test or composite score if the reliability of the criterion measure were only .30, which can bery well happen, and if no correction for attenuation were made. If the uncorrected validity coefficient r_{xy} were .20, the correlation corrected for errors in the criterion would be .36, or almost doubled."[1]

---

[1] J.P. Guilford, Psychometric methods (2nd Ed.) New York: McGraw-Hill, 1954. P. 401.

This is precisely the case with the TAT. Its internal consistency and test-retest reliability coefficients are only in the range of .26 (e.g., Cronbach's alpha = .23, .26; test-retest coefficients = .22, Birney, 1959; .26, Krumbolz and Farquhar, 1957). If we apply the formula that Guilford (1954, P. 401) suggests for oneway correction for attenuation to the JPI data, we find that the validating correlation rises from .25 to .49.

$$r_{xw} = \frac{r_{xy}}{\sqrt{r_{yy}}}, \qquad .49 = \frac{.25}{\sqrt{.26}}$$

Figure 1. Formula for One-Way Correction for Attenuation.

r_{xw} = correlation with correction in Y only (TAT)

r_{xy} = obtained correlation between X & Y (TAT & JPI#2)

r_{yy} = reliability coefficient of Y (TAT)

The correction for attenuation informs us of the relationship between the two variables if the TAT were to have perfect reliability. Such corrected correlations are of little practical value since they cannot be used in prediction. However, correlation coefficients corrected for attenuation are of theoretical importance in the analysis of relationships, in that allowance can be made for variable errors of measurement, i.e., errors of measurement tend to reduce or attenuate the correlation between variables.

The formula for complete correction for attenuation (Guilford, 1954, P. 400) indicates that there may actually be a strong relation between the two tests.

$$r_{vw} = \frac{r_{xy}}{\sqrt{r_{xx}\, r_{yy}}} \quad , \quad .88 = \frac{.25}{\sqrt{.26 \quad .32}}$$

Figure 2. Formula for Complete Correction for Attenuation.

r_{vw} = correlation between true components of X & Y (TAT & JPI #2)

r_{xy} = obtained correlation between X & Y

r_{xx}, r_{yy} = reliability coefficients of X & Y

Once again such corrected correlations are of little practical value, but they give us an idea of what may be the true relationship between variables if the variables had high reliability. Low reliabilities limit the size of any validating correlations.

We must reserve judgement on the strength of JPI #2 because of the unreliability of the TAT as a criterion of n Ach. Because of the limitations of poor TAT reliability, researchers should be careful to employ a different n Ach criterion in future attempts to develop alternate measures of n Ach.

Relationship Between the TAT and JPI

Although there is a statistically significant relationship between the TAT and JPI #2, the relationship is too small, on the evidence presented here, to consider JPI #2 as a valid substitute measure of n Ach. Further validating experiments are necessary to establish the validity of the JPI. In view of Mehrabian's work (1968, 1968a, 1969), such validity studies may prove fruitful.

Mehrabian undertook construction of a multiple choice measure of n Ach simultaneous to the initiation of this research project. His approach was similar, that is, creation of items from the n Ach literature, and partial validation against the TAT criterion. His results were markedly similar. Mehrabian's Resultant Achievement Motivation Scale (RAM) correlated with the TAT .28 ($N=110$, $p < .01$), while JPI #2 correlated with the TAT .25 ($N=94$, $p < .02$).

In addition to using the TAT as a validating criterion, Mehrabian employed various laboratory problem solving tasks as n Ach criteria. For example, Weiner, Johnson and Mehrabian (1968) used the RAM scale in a study in which a greater Zeigarnik effect was hypothesized for high achievers than for low achievers. Ss who scored high on the TAT-TAQ index (an individual's

z-score on the TAT minus his z-score on the Test Anxiety Questionnaire, Mandler and Sarason, 1952) showed a Ziegarnik effect in 15 out of 20 cases, whereas Ss who scored high on the RAM showed a Ziegarnik in 17 out of 21 cases. The difference between the proportion of high versus low achieving Ss who recalled a greater percentage of incompleted than completed questions approaches significance ($p < .10$) when the TAT-TAQ is used as a measure of achievement and is significant ($p < .05$) when the RAM is used.

Law (1968) investigated problem solving behaviors of high and low achievers categorized according to short versions of the RAM. He used match-stick problems similar to those originally used by Wertheimer (1945). Groups of three to four male subjects were presented with 11 match-stick problems and allowed 15 minutes to solve them, after they had seen a sample solution and were told, "We would like to know how many you can solve in 15 minutes." The product-moment correlation of achievement scores and number of problems attempted was .38 ($N=57$, $p < .01$). Thus, high achievers tend to persevere and attempt more problems and also solve more problems than low achievers. Furthermore, for low achievers Law found a .41 correlation ($N=35$, $p < .05$) between the degree of preference for returning to participate in a similar experiment and the subjects'

subjective experience of success in solving the problems. However, for high achievers, he found no such significant correlation ($r = .17$, $N=35$, $p<.10$).

Shapira (1968) hypothesized that the difference in the performance of high versus low achievers is less in a group achievement situation than in an individual achievement situation. Her dependent variable was performance on the Digit Symbol substitution subscale of the Wechsler Adult Intelligence Scale (Wechsler, 1955) Performance scores were the number of correct substitutions in one minute. In the individual achievement condition, an emphasis was placed on scores obtained by each individual on the Digit Symbol test. Ss were told that the test carried much weight in the determination of intelligence level. In the group achievement condition, an emphasis was placed on scores obtained by the entire group. Ss did not write their names on the tests and were told that the experimenter was interested in the intelligence level of UCLA students whom they happened to represent.

The Ss were classified according to the RAM as high (low) achievers for above (below) median scores. Shapira found that the only significant effect was due to the high versus low achiever condition which indicates that irrespective of instructions, high achievers perform better than low achievers.

In addition to the findings reported thus far, there is also some correlational data (Mehrabian, 1968) relating the RAM to other scales. With an N of 110, the RAM correlated with Mandler and Sarason's (1952) Test Anxiety Questionnaire -.16 ($p < .05$); .29 with the TAT-TAQ ($p < .01$); .64 with Rotter's (1966) internal-external control scale ($p < .01$); and -.18 with the Crowne and Marlowe (1960) social desirability scale ($p < .05$).

Mehrabian (1969) reports further correlational data with a revised version of the RAM which consists of 26 items (the revised scale correlated .94 with the 34 item RAM). The revised RAM correlated .20 ($p < .05$) with the Bass (1967) Task Orientation Scale; .62 ($p < .05$) with the Jackson (1967) Achievement Scale; -.24 ($p < .05$) with the Liverant (1958) Social Love and Affection Scale; .23 ($p < .05$) with the Guilford and Zimmerman (1949) Sociability Scale; -.40 ($p < .05$) with the Eysenck and Eysenck (1963) Neuroticism Scale; -.25 ($p < .05$) with Rokeach's (1960) Dogmatism Scale; .56 with the Cattell and Eber (1957) Shy-Venturesome Scale, with the high achievers scoring on the more venturesome end. And finally, the RAM correlated .24 ($p < .05$) with the Crowne and Marlowe (1960) social desirability scale. This correlation is in the opposite direction from an earlier study (Mehrabian, 1968).

With the exception of the contradictory findings with the Crowne-Marlowe scale, these correlations are consistent with n Ach research and theory.

n Ach as a Multi-Dimensional Factor

One of the questions raised by this research was whether or not n Ach should be considered as a unidimensional or multidimensional personality factor.

An analysis of the TAT data of Research Sessions #1 and #2 indicated that n Ach is a multidimensional factor. Cronbach's (1951) coefficient alpha on the TAT was only .23 (N=30) and .26 (N=94) on the samples tested. This is an analysis of variance approach to estimating internal consistency that is analogous to the Kuder-Richardson Formula 20. The low internal consistency of the TAT indicates that several factors are operating which contribute to an S's overall TAT score.

Mehrabian started with the assumption that n Ach was a heterogeneous personality dimension. This provided the rationale underlying the construction of the RAM. "Two factor analyses of the male scale yield 11 factors and the first factor accounts for no more than 15% of the total variance. The results of the factor analyses therefore indicate that the male scale is heterogeneous in content."[1] If we accept the RAM as

---

[1]Albert Mehrabian, An analysis of personality theories. Englewood Cliffs, N.J.: Prentice-Hall, 1968a. P. 114.

as a valid measure of n Ach, n Ach must be considered as a multidimensional factor.

The low coefficient alpha of the TAT, the results of Mehrabian's factor analysis, and the low internal consistency of JPI #2 are evidence that n Ach is a multidimensional personality factor. Future attempts to construct multiple choice measures of n Ach should consider this heterogeneous conception of n Ach.

BIBLIOGRAPHY

BIBLIOGRAPHY

Atkinson, John W. Motives in fantasy, action, and society. Princeton, N.J.: Van Nostrand, 1958.

Atkinson, J.W., and Litwin, G.H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. Journal of Abnormal and Social Psychology. 1960, 60, 52-63.

Atkinson, John W. An introduction to motivation. Princeton N.J.: Van Nostrand, 1964.

Barnette, W.L. A structured and semi-structured achievement measure applied to a college sample. Educational and Psychological Measurement, 1961, 21, 647-656.

Bass, B.M. Social behavior and the orientation inventory: a review. Psychological Bulletin, 1967, 78, 260-292.

Behavioral Research Service. Motive patterns of managers and specialists. New York: General Electric Co., 1960.

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

Brown, Roger. Social psychology. New York: The Free Press, 1965.

Byrne, Donn. An introduction to personality. Englewood N.J.: Prentice-Hall, 1966.

Cattell, R.B., and Eber, H.W. Sixteen personality factor questionnaire. Champaign, Illinois: The Institute for Personality and Ability Testing, 1957.

Cronbach, L.J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Crowne, D.P. and Marlowe, D.C. A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 1960, 24, 349-354.

De Charms, R., Morrison, W., Reitman, W., and McClelland, D.C. Behavioral correlates of directly and indirectly measured achievement motivation. In D.C. McClelland (Ed.), Studies in motivation. New York: Appleton, 1955. Pp. 414-423.

Edwards, A.L. Edwards personal preference schedule. New York: Psychological Corporation, 1954.

French, Elizabeth G. Interrelation among some measures of rigidity under stress and nonstress conditions. Journal of Abnormal and Social Psychology, 1955, 51, 114-118.

French, Elizabeth G. Development of a measure of complex motivation. In J.W. Atkinson, 1958. Pp. 242-248.

Guilford, J.P. and Zimmerman, W.S. The guilford-zimmerman temperament survey: manual of instructions and interpretations. Beverly Hills, California: Sheridan Supply Co., 1949.

Guilford, J.P. Psychometric methods. (2nd Ed.) New York: McGraw-Hill, 1954.

Haber, R.N., and Alpert, R. The role of situation and picture cues in projective measurement of the achievement motive. 1958. In J.W. Atkinson, 1958, Pp. 644-663.

Hays, William L. Statistics for psychologists. New York: Holt, Rinehart & Winston, 1963.

Heckhausen, Heinz. Hoffnung und furcht in der leistungsmotivation. Meisenheim/Glan: Hain, 1963. Reported in Heckhausen, 1967.

Heckhausen, Heinz. The anatomy of achievement motivation, New York: Academic Press, 1967.

Hurley, J.R. The iowa picture interpretation test: a multiple choice variation of the TAT. Journal of Consulting Psychologists. 1955, 19, 372-376.

Jackson, D.N. Manual for the personality research form. London, Canada: Department of Psychology, the University of Western Ontario, 1967.

Karolchuck, Patricia, and Worell, L., Achievement motivation and learning. Journal of Abnormal and Social Psychology. 1956, 66, No. 4, 291-308.

Klinger, Eric. Need achievement as a motivational construct. Psychological Bulletin. 1966, 66, No. 4, 291-308.

Knapp, R.H., and Garbutt, J.T. Time imagery and the achievement motive. Journal of Personality, 1958, 26, 426-434.

Krumboltz, J.D., and Farquhar, W.W. Reliability and validity of the n achievement test. Journal of consulting psychologists. 1957, 21, 226-228.

Law, Robert. Problem solving behaviors of high and low achievers. Unpublished work, U.C.L.A. Reported in A.M. Mehrabian, 1968.

Liverant, S. The use of rotter's social learning theory in developing a personality inventory. Psychological Monographs, 1958, 72, (2, Whole No. 455).

Lowell, E.L. A methodological study of projectively measured achievement motivation. Unpublished master's thesis, Wesleyan University, 1950.

McClelland, D.C., Atkinson, J.W., Clark, R.A., and Lowell, E.L. The achievement motive. New York: Appleton-Century-Crofts, 1953.

McClelland, D.C. The achieving society. New York: Free Press, 1961.

McClelland, D.C. Achievement drive and economic growth. In D.C. McClelland (Ed.), The roots of consciousness. Princeton, N.J.: Van Nostrand, 1964. Pp. 16-45. Reprinted from Harvard Business Review, 40, No. 4, July 1962, Pp. 99-112.

Mandler, G., and Sarason, S.B. A study of anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 47, 166-173.

Marlowe, D. Relationships among direct and indirect measures of the achievement motive and overt behavior. Journal of Consulting Psychologists. 1959, 23, 329-332.

Mehrabian, A.M. Male and female scales of the tendency to achieve. Educational and Psychological Measurement. 1968, 28, No. 2, 493-502.

Mehrabian, Albert M. An analysis of personality theories. Englewood Cliffs, N.J.: Prentice-Hall, 1968a.

Mehrabian, A.M. Measures of achieving tendencies. Educational and Psychological Measurement, 1969 in press.

Messick, Samuel, and Ross, John. Measurement in personality and cognition. New York: John Wiley, 1962.

Mischel, W. Delay of gratification, need for achievement, acquiescence in another culture. Journal of Abnormal Social Psychology, 1961, 62, 543-552.

Morgan, H.H. Measuring achievement motivation with picture interpretation. Journal of Consulting Psychologists, 1953, 17, 289-292.

Ricks, D., and Epley, D. Foresight and hindsight in the TAT. Paper read at Eastern Psychological Association, N.Y., April, 1960.

Robinson, W.P. The achievement motive, academic success, and intelligence test score. British Journal of Social Clinical Psychology, 1964, 4, 98-103.

Rokeach, M. The open and closed mind. New York: Basic Books, 1960.

Rosen, B.C. The achievement syndrome. American Sociological Review, 1956, 21, 203-211.

Rotter, J.B. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80 (1, Whole No. 609).

Sampson, E.E. Achievement in conflict. Journal of Personality, 1963, 31, 510-516.

Shaffer, Laurance F. Preparing doctoral dissertations in psychology. (5th Ed.) New York: Teacher's College Press, 1967.

Shapira, A. The effect of shared responsibility on the problem performance of high and low achievers. Unpublished manuscript, U.C.L.A., 1967. Reported in A.M. Mehrabian, 1968.

Smith, Charles P., and Feld, Sheila. An evaluation of the objectivity of the method of content analysis. 1958. In J.W. Atkinson, 1958. Pp. 234-241.

Smith, Charles P., and Feld, Sheila, How to learn the method of content analysis for n achievement, n affliation, and n power. 1958. In J.W. Atkinson, 1958a. Pp 685-818.

Tyler, Leona E. Tests and measurements. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

Veroff, J., Atkinson, J.W., Feld, Sheila, and Gurin, G. The use of thematic apperception to assess motivation in a nationwide interview study. Psychological Monographs, 1960, 74, No. 12 (Whole No. 499).

Weiner, B., Johnson, P.B., and Mehrabian, A.M. Achievement motivation and the recall of exam questions. Journal of Educational Psychology. 1968, 59, Pp. 181-185.

Wechsler, D. Wechsler adult intelligence scale. New York: Psychological Corporation, 1955.

Wertheimer, M. Productive thinking. New York: Harper, 1945.

Williams, J.E. Mode of failure, interference tendencies, and achievement imagery. Journal of Abnormal and Social Psychology, 1955, 51, 573-580.

Winterbottom, Marina R. The relation of need for achievement to learning experiences in independence and mastery. 1958. In J.W. Atkinson, 1958. Pp. 453-478.

APPENDICES

APPENDIX A

COMPARISON OF ITEM KEYS

Appendix A reports individual item keying for purposes of keying comparison.

JPI #1 = 87 item a priori scale; 44 True & 43 False.
JPI #2 = 43 item scale; 26 True & 17 False.
JPI #2a = 27 item subscale of JPI #2; 16 True & 11 False.
JPI #3 = 37 item scale; 22 True & 15 False.
JPI #4 = 59 item scale; 32 True & 27 False.

Key: #1	#2	#2a	#3	#4	Item
t					1. I believe that it is necessary for Man to master his physical and social environments.
f	f			f	2. I like to be whimsical.
f					3. All I want out of life in the way of a career is a secure, not too difficult job, with enough pay to afford a nice car and eventually a home of my own.
t					4. My parents expected me to make friends on my own at an early age.
t	t	t			5. Even when I like the "underdog" in a contest I generally do not bet on him.
t			t	t	6. I generally show improvement in task performance rather than performance at the same or slower rate.
f					7. My parents did not stress independent accomplishment when I was a child.
t	t		f		8. I have a better memory for tasks that I have left incompleted than for tasks that I have completed.

Key: #1	#2	#2a	#3	#4		Item
t				t	9.	My parents consider me to be a highly motivated person.
f					10.	I prefer bright reds to subdued blues in my color choices.
t				t	11.	I enjoy traveling more than my friends do.
t			t	t	12.	In choosing work partners I prefer successful strangers to unsuccessful friends.
t			t		13.	At times I feel that I am unconsciously trying to please my mother.
f	f		t		14.	I do <u>not</u> consider myself independent in thought.
f				f	15.	I like to think of time as "a large revolving wheel." (a metaphor)
t	t	t			16.	I am a member of an upward-mobile family.
f					17.	I am <u>not</u> concerned over accomplishment.
t				t	18.	If my parents told me to stop seeing a friend of my own sex, I would see that friend anyway.
t	t	t	t	t	19.	I get the greatest satisfaction out of playing horseshoes when I stand a moderate distance from the ring rather than far away.
t	t	t		t	20.	Competition and success should not be viewed as the "great criminals" of the American Society as a number of critics have claimed.
f	f			f	21.	Parents should <u>not</u> expect a lot of their children.
f	f		f	f	22.	I am <u>not</u> energetic.

Key: #1	#2	#2a	#3	#4		Item
t	t	t	t	t	23.	I think that everyone should become a professional man if he is capable.
f				f	24.	I am <u>not</u> more productive in my work than the average person.
t	t		t	t	25.	I believe that a man should always compete with a "standard of excellence."
f	t				26.	I like to think of time as "a vast expanse of sky." (a metaphor)
f	f			f	27.	I do <u>not</u> like to volunteer for psychological experiments.
f				f	28.	The prestige of my job makes little difference to me.
t				t	29.	My mother developed independence and feelings of trying for self-mastery in me as a youngster.
f				f	30.	When choosing to perform a task I tend to choose friends instead of experts.
t			t	t	31.	I am a future-oriented person.
t	t	t	f	t	32.	The values that I hold are primarily individualistic.
t			t	t	33.	When I have left a task incompleted I have a great desire to go back and finish it.
f			f		34.	My parents put many restrictions on my behavior as a youngster.
t	t	t	f		35.	I enjoy organizing groups and committees.

Key: #1	#2	#2a	#3	#4		Item
f	f		f	f	36.	Parents should <u>not</u> expect a lot of their children.
f					37.	I like to perform routine tasks.
f	f	f	f	f	38.	Even when one gets married his main loyalty still belongs to his mother and father.
f				f	39.	Planning only makes a person unhappy since your plans hardly ever work out anyway.
t			t	t	40.	I am oriented toward performing tasks well for their own sake rather than to please others.
t	t		t	t	41.	I make more demands on myself than others make on me.
f	f	f		f	42.	I do <u>not</u> like to engage in competitive activity.
f			f	f	43.	I have a domineering father.
t	t	t	t	t	44.	I like to volunteer for psychological experiments even when credit is <u>not</u> being given for it.
t			t	t	45.	I usually work at tasks just as hard whether or not I am going to receive a reward.
f			f	f	46.	I do <u>not</u> consider myself independent in action.
f	f	f	f		47.	I think that when subjects take psychological questionnaires they should try hard to please the psychologist.
f					48.	I do <u>not</u> have a fear of failure.
t	t	t		t	49.	I like to think of time as "bird in flight." (a metaphor)

Key: #1	#2	#2a	#3	#4		Item
f				f	50.	I do <u>not</u> work harder than most people.
t				t	51.	Much of my effort has been to compete and strive for success.
f	f	f			52.	I do <u>not</u> like a good argument.
t					53.	Parents should reward their children with hugs and kisses.
t	t	t	t	t	54.	I like jobs involving responsibility for initiating decisions.
f	f	f		f	55.	I would rather get $10 right now than have to wait a month to get $30 later.
t				f	56.	I tend to value money not so much for its own sake but rather because it is a concrete measure of my success.
t				t	57.	My mother stressed independence training in our family.
t	t	t	t	t	58.	I am confident and optimistic about what I can accomplish.
f				f	59.	I am a past-oriented person.
t	t	t		t	60.	My life has an activistic orientation.
f	f	f		t	61.	When the time comes for a boy to take a job, he should stay near his parents even if it means giving up a good job.
f			t	t	62.	When I have left a task incompleted which I know will be easy to finish I have great desire to go back and complete it.

Key: #1	#2	#2a	#3	#4		Item
t				t	63.	I am a "man of action" as opposed to a "man of tranquility."
f			f	f	64.	If I have invested a great deal of time and energy in a project and find that I have been making a fundamental error, I generally do <u>not</u> take the time to go back to the beginning and start over again.
t					65.	I prefer moderate risks (3 out of 6) to either safe (5 out of 6) or speculative (1 out of 6) risks.
t	t	t			66.	I feel that I conform less than do some of my peers.
f	f	f		f	67.	I believe that independence is <u>not</u> that important a virtue.
f	f	f	f		68.	It is more important to be socially acceptable than to be a success.
f	f	f		f	69.	I do <u>not</u> consider myself an active person.
t					70.	I like to take calculated risks.
f				f	71.	I do <u>not</u> like to innovate.
t	t	t		t	72.	I work hardest when I know that my personal efforts will make a difference in the outcome of a project.
t					73.	Parents should teach their children to be self-reliant.
f	f		t	f	74.	I do <u>not</u> like hard work.
f	f		f		75.	I do <u>not</u> like to gamble.
t	t	t	t	t	76.	I consider myself to be a person of originality.
t			t	t	77.	I like to engage in undertakings where there is a challenge.

Key: #1	#2	#2a	#3	#4		Item
f	f	f	f	f	78.	My father tends to be dogmatic.
t	f	f		t	79.	As a child my mother often rewarded me with "hugs and kisses" when I tried my best.
f	t		t	f	80.	Usually I will <u>not</u> volunteer for a difficult task.
f			f	f	81.	I prefer extreme risks (1 out of 6) to moderate risks (3 out of 6).
f	f		f	f	82.	I am uncomfortable when I know exactly how well I am performing on a task.
t	t			t	83.	My father developed independence and feelings of trying for self-mastery in me as a youngster.
t	t	t	t	t	84.	I like to think of time as "marching feet." (a metaphor)
f			f	f	85.	I do <u>not</u> like to perform somewhat difficult tasks.
t	t				86.	Others consider me to be high in need for achievement.
f					87.	I consider myself to be high in need for achievement.

APPENDIX B

RELIABILITY OF TAT SCORING BY MORIARTY

Appendix B reports the reliability of TAT protocol scoring by Moriarty. Moriarty learned the method of n Ach scoring by studying the manual "How to Learn the Method of Content Analysis for n Achievement, n Affliation, and n Power" (Smith and Feld, 1958). Smith and Feld report that "approximately three out of four persons who complete the training procedure outlined here can be expected to be ready to score protocols for research purposes with very little additional practice. Inter-judge reliability in the .90's is, of course, desirable, but a number of useful studies have been conducted with scoring reliabilities in the .80's."[1] Moriarty's percentage agreement in scoring achievement imagery against the expert coding was in the mid 90's. This is more than adequate for research scoring.[2]

Two criteria must be met to become a reliable scorer. (1) There must be a high percentage agreement on the prescence of motive-related imagery between the

---

[1]Charles P Smith and Sheila Feld, How to learn the method of content analysis for n achievement, n affliation, and n power. 1958. In J.W. Atkinson, 1958. P. 691.

[2]"Most important is to be sure you can code reliably. Check against expert coding in book. More than adequate means r = .90 or better." Personal correspondence with Atkinson.

novice and the expert. This is "the single most important scoring decision to be made."[1] (2) There must be a high rank order correlation on total score assigned to each story between the novice scorer and the experts.

Moriarty achieved a mean percentage agreement in scoring achievement imagery with the experts of .915; his median agreement was .96. The norms presented in the manual between novice scorers and experts for practice story sets B-D were a mean of .856 and a median of .84. Moriarty's mean rank order correlation with the experts on total score of the sample stories was .87; his median rank order correlation was .91. The norms presented in the manual between novice scorers and experts for the rank order correlations on total scores were a mean of .78 and a median of .785. Thus Moriarty's scoring was well above that of the novice scorers in the manual and correlated more than adequately with the expert scorers.

Table 2 (adapted from Atkinson, 1958, P. 238) reports the percentage agreement in scoring achievement imagery and the rank order correlations between n Ach scores obtained by novice scorers and experts. It also reports Moriarty's percentage agreement and rank order correlations with the experts on the sample scoring stories.

---

[1] Charles P. Smith and Sheila Feld, An evaluation of the objectivity of the method of content analysis. 1958. In J.W. Atkinson, 1958. Pp. 234-241.

Table 2.

Percentage Agreement in Scoring Achievement Imagery and Rank Order Correlations between n Ach Scores Obtained by Novice Coders/Experts and Moriarty/Experts

Subject	Reliability Index	Story Set B	C	D	E	F	Mean	Median
Median of Novice Scorers	%	.82	.91	.84	*	*	.856	.84
	Rho	.785	.805	.755	*	*	.78	.785
Moriarty	%	.96	1.00	.88	.744	.96	.915	.96
	Rho	.97	.91	.79	.72	.95	.87	.91

* = data not available

% = percentage agreement on the presence of motive-related imagery between novice scorers and experts

Rho = rank order correlations on total score assigned each story between novice scorers and experts