DOCTORAL DISSERTATION SERIES
Publication No.: 11,246

AUTHOR: Robert Charles Birney, Ph. D., 1955
University of Michigan

TITLE: STUDIES ON THE ROLE OF PICTURE CUES IN PROJECTIVE MEASUREMENT OF ACHIEVEMENT MOTIVATION

University Microfilms, Ann Arbor, Michigan

STUDIES ON THE ROLE OF PICTURE CUES IN PROJECTIVE MEASUREMENT OF ACHIEVEMENT MOTIVATION

by

Robert C. Birney

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1955

Committee in charge:

Associate Professor Edward L. Walker, Chairman
Assistant Professor John W. Atkinson
Professor Clyde H. Coombs
Assistant Professor Edward J. Furst
Associate Professor Daniel R. Miller

STUDIES ON THE ROLE OF PICTURE CUES IN PROJECTIVE MEASUREMENT OF ACHIEVEMENT MOTIVATION

by Robert C. Birney

The general purpose of this study was to determine more accurately the role of the picture cues in the projective measurement of the Achievement motive. First it was proposed to use ratings of the pictures on the scoring categories of the n Achievement system as an independent definition of the picture cues. Then these rated values were compared with the obtained n Achievement scores in an effort to see how well cognitive estimates of the pictures would predict motivational data.

One hundred and fifty college students rated an average of four pictures each from a total sample of sixty pictures. Each picture was rated by an average of ten persons, both male and female, and from these ratings median values were obtained for each picture on each of the respective scoring categories. These medians were then combined into one distribution and rank order positions then assigned to the pictures. Eighteen pictures were then selected to provide three sets of six pictures. These sets were made parallel in ratings of strength, affect, and content.

The three sets of pictures were then made part of a two hour testing session in which a total of 73 college males performed on a Zeigarnik task, wrote stories to the pictures, recalled the Zeigarnik task items, then took speed tests in Addition and in Word Coding. All three tasks were known to have provided performance scores which in past studies showed a positive relationship to the n Achievement score. Control procedures made it possible to treat differences in n Achievement score distributions between pictures as due to pictures per se.

It was found that the raters could reliably predict both the achievement-related affect scores, and the likelihood of scorable achievement themes. However

they could not anticipate the details of the achievement stories.

In the second part of the study an effort was made to compare pictures in terms of the discriminability and the directional predictiveness of their scores. It was found that the correlations of picture scores with performance scores on Additions and Word Coding yielded distributions symmetrical about zero. Pictures closest to the theoretically most discriminating position of eliciting stories from 50 per cent of the subjects were found to yield scores with the highest correlations with performance, regardless of direction of relationship of score to performance.

Efforts to show that picture differences in direction of prediction were a function of content and affect cues were unsuccessful. When pictures were given definitions in terms of ratings of strength or affect these same hypotheses about picture differences again failed.

The failure of individual pictures to predict performance more reliably was attributed to an unforeseeable combination of circumstances which tended to depress differences in performance as a function of motivation. The failure of the hypotheses about the role of picture cues in effecting the directional prediction of n Achievement scores was taken to mean that the operations designed to test them had not proved adequate.

The results of this study are taken to indicate that the development of an operation to give picture cues a definition independent of the scores they elicit is feasible. The finding that many pictures gave scores negatively related to performance raises many interesting questions about the factors which determine the n Achievement score. A discussion of these factors and their interaction is presented along with suggestions for future research on this problem.

ACKNOWLEDGEMENTS

I wish to acknowledge the kind support of my thesis committee members; and to extend special thanks to Dr. Edward L. Walker for his sustained interest, invaluable advice, and patient instruction.

Sincere thanks are also due Dr. David McClelland of Wesleyan University, who first made me aware of the problem; to Dr. John W. Atkinson for his unfailing encouragement of my interest, and finally to my wife, Margaret Ann, who in the most realistic sense made this work possible.

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
LIST OF FIGURES	v
Chapter	
I. THE EFFECT OF PICTURES ON THE PROJECTIVE MEASUREMENT OF HUMAN MOTIVATION	1
History of the Achievement Motive Research	2
The n Achievement Score	3
A Review of the Literature of Picture Studies	4
II. DEFINITION OF PICTURE CUES	10
Theory of the n Achievement Score	12
Hypotheses	16
Procedure for Rating Pictures	20
Administration of Rating Tasks	23
Preliminary Analysis of Rating Data	23
Selection and Use of Pictures	25
Procedure for Testing Session	26
Statement of the Hypotheses	28
Results	30
III. PICTURE EFFECTS ON THE VALIDITY OF n ACHIEVEMENT SCORES	37
The Performance Measures	37
Tests of Hypotheses	39
Description and Treatment of Zeigarnik Task Data	47
Additional Analysis	49
Summary of Results	54
IV. DISCUSSION	56
Suggestions for Future Research	60
Summary	62
APPENDICES	65
BIBLIOGRAPHY	83

LIST OF TABLES

Table		Page
1.	The Latin Square Design	31
2.	Mean Performance Scores on Criteria Tasks	33
3.	Pictures Ranked by Consistency of Score Correlations with Criteria Scores	51
4.	Frequency and Percentage of Each Type of Imagery for Each Picture	53
5.	Median Rank Values for Pictures on Rated Categories .	75
6.	Rank Orders of Pictures Used in Correlational Analysis (obtained scores)	76
7.	Values of Tau_b for n Achievement Scores of Each Picture with Performance Scores on Additions and Word Coding Tasks: Zeigarnik Index Values for Each Picture	77
8.	Distribution of w Coefficient	78

LIST OF FIGURES

Figures Page

1. Set of Pictures Shown to Group 1 44

2. Set of Pictures Shown to Group 2 45

3. Set of Pictures Shown to Group 3 46

CHAPTER I

THE EFFECT OF PICTURES ON THE PROJECTIVE MEASUREMENT OF HUMAN MOTIVATION

Progress in psychology, as in all science, depends on the accumulation of reliable facts. Thus far studies of human motivation have been handicapped by a lack of reliable measuring operations, and thus a lack of repeatable or verifiable facts. One of Freud's great contributions to psychology has been the demonstration that a rich source of material, relevant to an individual's basic motivation, lies in the individual's fantasy products such as dreams, free associations, and imaginatively produced stories. Morgan and Murray (18) initiated a method of measuring important human motives by the analysis of imaginatively produced stories. This method, known as the Thematic Apperception Test, was given new form by McClelland and his associates over the past several years.

One of the chief features of fantasy behavior is its free response character, that is, the individual himself sets the "problem" and gives the answer. Under such circumstances differences between persons appear, and the observor's capacity to predict behavior from a knowledge of the stimulus is at a minimum. In addition the fantasy products are usually full of emotional, affective content; an aspect of experience which always plays an important role in the deepest motives of human life. Thus since fantasy products are unique to the

individual and affective in character they are well suited to become the data of operations designed to measure motivation.

History of the Achievement Motive Research

In the Thematic Apperception Test the subject is asked to tell imaginative stories about pictures of an ambiguous character. McClelland and his associates were interested in developing a system of analysis of protocols of such stories which had high reliability of agreement between scorers; and which required only a reasonable amount of training and practice. This effort reached its fullest early development with a single motive which they judged to be important in this culture, i.e., the achievement motive. They have carried out a long series of studies designed to obtain a valid measure of this motive. Their hope has been to develop within this framework a general method applicable to the measurement of a variety of human motives. The success of this effort may be seen in their summary of this research in The Achievement Motive (17).

The experiments of McClelland and his associates have dealt mainly with the effects of motive-arousing instructions and experiences on the achievement score (n Achievement score) taken from stories told by the subjects. They have also demonstrated the validity of the n Achievement scores by showing that persons with high scores work hardest in achievement situations. The pictures used in these studies were selected by the experimenters on a priori grounds as having some relevance to culturally defined achievement situations. Pictures were then validated in terms of their degree of agreement in ordering subjects in a manner similar to that given on the total scores on all pictures

taken together. No effort was made to determine how suggestive of achievement the pictures were, and no selection procedures have been devised to permit selection of pictures yielding valid n Achievement scores.

The n Achievement Score[1]

In scoring for achievement references the initial task is to determine whether or not a character in the story is concerned with " . . . success in competition with some standard of excellence (17, p. 111)." Operationally this usually means the story contains an evaluative reference to some performance, such as doing better, wanting to be best, etc. Such content is scored as Imagery, and then the scorer goes back and looks for the remaining categories of behavior which make up the scoring system. These additional categories reflect the basic learning theory paradigm of Need—Instrumental Activity—Goal. The sum total of such scores across all pictures is the subject's n Achievement score.

Thus far only two studies have had picture differences central to their purposes. Using appropriate controls Atkinson (17) had 32 college students write stories to eight pictures. He was able to order the pictures in two parallel forms of three pictures each. The remaining two pictures showed little agreement with the others. He also found that the n Achievement score

> . . . was significantly higher in response to pictures of adults in achievement-related situations than in response to pictures of adults in unrelated or unstructured situations. Pictures containing only child characters were found to be relatively invalid.

---

[1]See The Achievement Motive (17) for the complete scoring system.

Jacobs (15) has made the first attack on the problem of obtaining an independent definition of the stimulus characteristics of the pictures. He had subjects rate the pictures for concern with achievement and concluded that there was no relationship between a subject's rating of a picture and that subject's n Achievement score obtained when writing a story to the picture. Some of his pictures were not rated as high in achievement cues and a comparison of the order of pictures on strength by Jacobs' raters shows the same order of pictures as Atkinson obtained when he compared mean n Achievement scores per picture. This result would indicate that the scores are unreliable to a substantial degree but that a positive relationship between ratings of pictures and score can exist.

These two studies have been the first steps in an effort to learn more about the role of picture cues in the determination of the stories. The effects of pictures on projective materials have been dealt with by researchers mainly interested in the standard Thematic Apperception Test.

A Review of the Literature of Picture Studies

The Thematic Apperception Test, (TAT), first published in 1936, has been the object of considerable research since that time. In addition a variety of other picture tests have appeared from time to time. The purpose of constructing new pictures has usually been to study a more specifically determined area of the subject's experience in his culture.

The TAT itself has had wide use as a diagnostic tool by clinical

psychologists, but it has rarely been used in experimental studies of personality variables because it does not provide an objective and reliable scoring system. Rather the clinician is free to interpret the material obtained according to his best judgment of its meaning. Fortunately for persons interested in TAT research, and this term covers all picture research, Holt began publishing an annual review of the TAT literature in 1948 (14). The classification system he employed reflects the trends in this research. Initially the TAT was put to work by clinicians as a device for eliciting fantasy material. With the general increase in more objective studies in the clinical area the number of validity and reliability studies has also increased. Thus we find many efforts to show that blind protocol analysis reveals material independently obtained from some other source. However the important feature of Holt's bibliographies for the present study is that the research has centered on response-response relationships, i.e., TAT scores related to other test scores. Efforts to experimentally validate the TAT by establishing stimulus-response relationships are notably absent.

The major evidence concerning the stimulus characteristics of the pictures can be found in studies to determine the normative nature of the responses. Norms imply some commonness of response due, in part, to the stimulus properties of the pictures. Normative data on the TAT has been published by Rosenzweig (24) and Eron (9). These studies show that the majority of the 20 pictures elicit stories of hostility, affiliative loss, and anxiety. Three pictures, the boy-violin, No. 1, the boy-surgical mural (8 BM), and the silhouetted figure (14) commonly

elicit stories about achievement and success. However normative data is so rare that Bellak (2) can say

> No experimentally derived performance standards for 'normal' or even for various diagnostic groups have as yet been fully established, although much work is currently being carried on to provide evidence of modal performance for members of various groups (p. 205).

Very few studies have been done on dimensions of the TAT score other than the theme of the stories. However it should be recognized at the outset that Murray does not feel that this type of research should be emphasized. Rather he feels the TAT should be used to perfect the interpreter whom Murray calls psychology's forgotten instrument.

Eron et al (10), and Wittenborn and Eron (31) have established that the TAT pictures differ in their elicitation of "emotional tone," most being negatively toned despite wide picture differences. In another study Eron (11) found that cards with figures in social situations provided most themes of aggression, hostility, conflict, and occasional aspiration, while cards without clear social significance provided stories of a descriptive, impersonal, and symbolic nature.

To some extent the demonstration that the TAT cards do have distinctive and unique suggestive character is merely a proof of the obvious. Certainly such results must validate the unpublished notions of Murray's co-workers who selected the pictures by trying out several hundred cards until they found a set that gave good responses. The basic issue is one of determining how to predict the modal responses to pictures so that pictures can be devised for specific problems in important areas of personality research.

The construction of new picture tests has provided some research on picture differences. Wekstein (30) believed that children give their

most revealing responses when being as fanciful as possible. Therefore he created pictures of extreme symbolic and bizarre appearance and asked the children to tell fairy tales. He reported that he obtained very elaborate tales of diagnostic value.

Two researchers have changed the appearance of the figures to correspond to that of the subjects on whom the tests were to be used. Thompson (26) changed the TAT figures to Negroes and reported an immediate increase in meaningful material. Subsequent work on this "identification" variable with Negroes has proved equivocal, both success and failure with such efforts being reported. Henry (13) adapted the TAT situations for use with American Indians by creating pictures with Indian figures in situations unique to their culture. He also reports success in obtaining useful protocols. Weisskopf (27) made the most drastic changes in pictures by superimposing actual pictures of the subjects being tested, and also in another study by altering the physique of the figures to match the subjects (28), and in neither case found any change in the amount of projective responses.

Weisskopf (27) studied the problem of picture ambiguity by using pictures under low and high illumination, and found no differences in protocols. She also used pictures with complete vs incomplete lines, and found the complete pictures to be far superior. From this she concluded that physical ambiguity of a picture has to do with the perceptual completeness of the picture. Thus a fragmented picture presents the subject with the task of organizing it into something meaningful, while the complete picture permits him to interpret the familiar figures. Weisskopf is among the first to give operational meaning to

the term "ambiguous."

To the writer's knowledge only Symonds (25) has published a detailed report on the selection of pictures for a new picture test. In selecting pictures for adolescents he first set up on a priori grounds criteria by which to judge the usefulness of imaginative products. These stated that the stories should have within them a variety of themes, a minimum of description of the picture, and little commonality across subjects. Stories written to the pictures were then rated by judges on these criteria, and the pictures given a rating on a five point scale of "goodness" as a function of the stories they elicited. Next he set out six dimensions on which to rate pictures. These were (a) emotionality, (b) amount of detail, (c) age relevance to subjects, (d) on-going action, (e) familial involvement, and (f) common-placeness of setting. The same judges then rated the pictures on these dimensions and correlations were run between the ratings of stories and ratings of pictures.

Symonds reports that pictures rated high on the "goodness" scale showed high ratings on three picture dimensions, namely amount of detail, age relevance, and common-placeness of setting. These results are based on a sample of 81 pictures.

This review of the literature permits some general conclusions. First it appears that pictures do differ in their suggestiveness of themes and emotions. Also the nature of the stories may vary with both the physical and the social cues in the picture. There is some agreement between investigators on these points. None of the literature thus far has dealt with the basic problem of establishing a method for

the selection of pictures so as to obtain stories with predictable fantasy material in them. It is in this area that the work of this study will begin.

CHAPTER II

DEFINITION OF PICTURE CUES

The decision to determine the cue properties of the pictures places us squarely in the position of facing an ancient and formidable problem in psychology—How do we define the stimulus? Traditionally the word stimulus referred to impingement of energy upon the receptors of an organism. Gradually, however, the word lost its rigor until we find it indiscriminately used to refer to objects external to the organism, receptor patterns, and to central states in the nervous system. This loss of precision has occurred in part because a common system of technical language has not been formulated. However, Brunswick (5) has suggested that the term proximal stimulus be used for receptor patterns, and distal stimulus used for sources outside the organism. Thus our question, reformulated, becomes—How do we define the distal stimulus properties of pictures?

There is a difference of opinion about this issue. Postman (22) insists that true stimulus-response laws can be written only when the distal stimulus has been given a physicalistic definition. However, Chapman (7) and later Boring (3) have argued that physical definitions have in common with social definitions the basic characteristic of consensual agreement between judges. The chief difference being that physical definitions require agreement by experts about reliable indices of objective measuring techniques, while social definitions usually lack

the objective techniques and have to rely more heavily on experts whose agreement arises from common training. The result may be considerably less consensus. It is the position of the writer that the consensual criterion is best for defining the social character of a distal stimulus so long as the selection of the judges is made on explicitly defensible grounds, e.g., the judges are drawn from the same population as the subjects to whom the stimulus is presented. The physical description of a picture in terms of its suggestiveness for achievement, for example, is impossible to attain.

Brunswick has had a pioneering interest in problems of the sort proposed for this investigation. One of his major points has been the necessity of providing for the subject a "representative" sample of stimulus objects so that the obtained relationships would have ecological validity and generality (5). Sampling, he has held, should apply to both subjects and problems. Similarly Bruner (4, p. 145) quotes McClelland as suggesting ". . . a phenomenological census to discover what things and attributes in the environment people look for and attend to in guiding their behavior." It has been from such suggestions and thought that the procedures for obtaining a definition of the stimulus characteristics of the picture have been worked through. Brunswick has also proposed that attributes such as "achievement suggestiveness" be called <u>covert</u> distal value as opposed to the <u>overt</u> measures of physical dimensions, leaving us to define picture values with the term covert distal stimulus value. Since this is an awkward compound term, we will refer henceforth to the covert distal stimulus value of the picture as the picture cues.

We have now provided a background against which to define the purpose of this thesis. It was our purpose to inquire further into the question of the effects of picture cues on the stories written to the pictures. A method of defining the picture cues in terms of picture ratings was devised and an effort made to test the method's usefulness. Then the pictures were defined both in terms of ratings and in terms of achievement imagery scores, respectively, and hypotheses about various picture cue dimensions were put to test. Finally those pictures which differed most reliably in the usefulness of scores they elicited were closely analyzed for new clues about the important picture cues.

Theory of the n Achievement Score

A story written to a picture is determined by three broad classes of variables:

1. Situational cues such as instructions, experimenter, physical setting, etc.
2. Picture cues.
3. Personality variables such as values, beliefs, and motives.

Achievement scores taken from stories have been systematically related to changes in the situation (17, p. 194); and individual differences in scores have been systematically related to performance in achievement situations, to values, and to beliefs (17). Considerably less is known about the effects of picture differences on n Achievement score distributions. As one considers the problems of this research a very pertinent question may come to mind. Why not ask the subjects if they have a high achievement motive? G. Allport (1) recently discussed

some research which implies that such a procedure might save us a considerable amount of trouble.

There are several answers to this question:

1. First, we know that self-ratings on achievement do not predict the n Achievement score, or performance in an achievement situation (17, p. 218).

2. Second, answers to direct questions may be no easier to interpret than stories simply because people differ widely in their understanding of the meaning of simply phrased questions.

3. Third, direct verbal expression may be impossible while other forms of expression such as occur in the various arts, of which story telling is one, may reveal quite clearly what the subject wishes to "say."

We are obviously using the third alternative when we obtain imaginative stories. The assumption here is that in filling in the skeletonized version of larger themes, which is the effect of our guide questions provided to the subject, the subject makes use of a much less directly articulate, but none the less readable, mode of expression.

We have made no programatic efforts to learn more about the effects of other relevant cues in the testing situation. There is some evidence that the n Achievement measure has wide applicability across subjects, being used successfully on Navaho school children, American high school boys and girls, and mature scientists (17, p. 184). However, the effects of such variables as the administrator's personality, non-academic settings, and important sociological variables have never been seriously studied.

An adequate conceptualization of the n Achievement score demands a theory of response determination which encompasses statements and assumptions about central processes customarily talked about under

perception and motivation.

The assignment of a score to a statement written in response to a picture is basically language analysis. When we obtain verbal data from a subject in response to some physical stimulus we can make several inferences about the relation of that stimulus to the response. With pictures as stimuli we generally find that verbal reports for different subjects differ widely. This means we cannot predict what will be said from our knowledge of the stimulus. Therefore if we are trying to establish stimulus-response laws we must hypothesize some intervening processes in the subjects which can then mediate their variability. The intervening variables to be used in this study are symbol, hypothesis, and motive. The last two are defined as subclasses of expectancies.

The terms are defined as follows:

1. A symbol is the conscious concommitant of the perception of the stimulus, such as thought or sensation.

2. An hypothesis is the expectancy aroused by stimulus pattern A about the likelihood of stimulus pattern B.

3. A motive is the expectancy that the affect level accompanying stimulus pattern A will be followed by a new and changed affect level for stimulus pattern B.

The majority of behavior theories in the past have found it necessary to infer a capacity in the organism for preparatory alertness to stimuli not yet in the field or environment. The observed fact of such behavior led the theorist to provide for it with a term designed to emphasize the forward looking character of the behavior. "Set," "anticipatory response," "expectancy," and "hypothesis" have all been proposed and are still in current use. The terms hypothesis and motive as defined above are being used for the same purpose.

We begin with the assertion that the expectancies we are dealing with arise from the learning experience of the subject. The exact nature of the laws of this learning are in some dispute. It is postulated here that the behavior sequence of stimulus ---> action ---> subsequent stimulus results in cognitive expectancy learning as well as motor learning. With this rudimentary notion set forth we can now turn to the development of the processes relevant to the n Achievement score.

In the course of the development of children in our culture they are frequently put through a sequence which involves competition with a standard of excellence, personal action toward success, and feelings of success or failure. Due to differences in ability, parental expectation, and a host of other influences they gradually build up feelings of relative success or failure. We then would expect persons with experience of success or failure to differ both in responses made in achievement situations, and in the motivational expectancies cued off in the presence of achievement situation stimuli. This means that if we present pictures of such situations to subjects the relevant expectancies are cued off; and stories written about these pictures contain goal statements, instrumentalities, and affective references which permit us to infer what the emotional and directional characteristics of the subject's expectancies may be.

The further assumptions involved in the use of story analysis can best be described with an example. Suppose a subject writes that "The boy wants to become a famous inventor." The scoring of this sentence is based on the assumptions that:

1. The subject is using words which parallel his thoughts about the picture.

2. The words mean roughly the same thing to both subject and experimenter.

3. The thoughts (symbols) reflect the expectancies cued off by the picture, and the situation in which the testing is taking place.

We find in this sentence an achievement goal, "inventor," referred to with positive affect, "famous." Thus we would score for achievement Imagery. By summing up such tallies we can roughly order the subjects on the degree to which achievement-related expectancies dominated their thoughts while writing the stories. The assumption that these thoughts are mediators of motivation is tested by predicting that persons with a high score will display more achievement motivation, i. e., work harder, in actual achievement situations than will persons with a low n Achievement score. The empirical findings to date confirm this prediction.

Hypotheses

An hypothesis was defined as an expectancy with no affective sign, either positive or negative. Rather the person has simply learned that event A usually heralds the event B. Thus the knowledge that the school room presents ample opportunity for competition and achievement can be learned through observation by a child who neither competes or achieves any more than is necessary to be left alone. Such expectancies can be learned with little or no direct action on the part of the organism. This means that there need be little affective quality to the expectancy, and the knowledge that such expectancies existed in a subject would tell us nothing about his probable performance in an achievement situation. However the knowledge that many subjects had

the same hypotheses about a stimulus could tell us what the definition of the stimulus should be for those subjects, i.e., a schoolroom picture has achievement stimulus properties.

In order to get such knowledge we need an operation which permits a maximum of determination of the response by the stimulus, and a minimum of determination by the subject. That is, we want the subjects to attend to the same aspect of the stimulus and respond in terms of it, not decide for themselves how to respond to it. We have such an operation when we ask subjects to identify, rank order, or otherwise compare stimuli. These methods have in common the fact that the experimenter provides the subject the very information which may be elicited in a free response situation. Thus if we tell the subject to rank the pictures for appropriateness to achievement we are asking the subject to tell us to what extent the hypothesis "if-boy-in-classroom-then-achievement-thoughts-by-him" applies to the picture. And his judgment does not give us information about his affective and goal oriented thoughts in such situations so as to permit us to make inferences about his motivation from his ranking data. Recognition of the achievement character of the situation tells us nothing of his motives.[1]

Implicit in the selection of pictures in past achievement motive studies has been the assumption that the pictures were in some way relevant to past achievement experience of the subjects. The most reliable way to assure ourselves of such relevance is to have the judgments made by persons similar to the subjects who will be asked to write stories.

---

[1]The question of the degree to which recognition of the cues by judges is conscious is not relevant to our use of these data.

Therefore we proposed to have a large sample of college students rate the pictures in order to provide us with a definition of the stimulus characteristics of the pictures.

However we faced a problem in trying to determine what kind of a rating task will give us the data we need. The answer seemed to be that we should give the raters as explicitly as possible the very attributes which go to define the n Achievement scoring system. Fortunately these attributes can be stated in a language which is meaningful to our subjects (a circumstance which would not occur if we used a system of latent content analysis).

In addition by making the attributes explicit we have ended the free response to pictures and forced the subjects to adopt the same frame of reference. This has the effect, if successful, of giving us response distributions which have considerably less variance than those obtained from free response situations. High agreement between raters then is taken as the chief criterion by which the consensual character of the definition of the picture cues can be determined. To the extent that high agreement between raters is demonstrated, individual differences in experience as a source of variance will have been reduced to a minimum. And such disagreement between raters as does exist may be either due to lack of common experience or more dynamic personality variables.

From the above argument we then concluded that if our subjects rated pictures for achievement relevance we would obtain a cognitive estimate of the stimulus character of the pictures. But if our subjects wrote stories about the pictures and made achievement references we

would obtain an estimate of the achievement motivation of the subjects. We will now try to make this argument more precise.

Let us begin by making two assumptions about the relationship between expectancies and stimuli, and between expectancies themselves. Our first assumption is taken directly from Postman (23). Let us assume:

1. An expectancy and its confirming information are related in a compensatory fashion such that the stronger the expectancy the less information needed to confirm it.

2. Two distinct expectancies, present simultaneously, contribute to the response pattern in an independent fashion. Interaction between the expectancies is confined to a raising of the amount of information required for confirmation of either expectancy.

where

expectancy is a motive or hypothesis related to achievement and measured by (a) the n Achievement score, or (b) the rank position of a picture on achievement relatedness.

information is the achievement cue in the picture.

First consider the rating task. In it the subject is provided with a number of alternative themes relevant to the picture such as Influence, Affiliation, Sex, etc., and asked to rank them in order of appropriateness to the picture. From the second assumption it follows that with many competing hypotheses only those given strong informational support will get immediate confirmation. Thus the information, or picture cues in combination with the ranking instruction, has a major role in determining the response. And if the subjects have had roughly the same experience with the stimulus there will be little disagreement between persons.

However, when we consider the story writing, free response, situation we find a wholly different state of affairs. Here the subjects,

after being exposed to certain motive-arousing conditions, differ widely in the amount of achievement related expectancies aroused. And when subjects are presented with picture cues, the primary determinant of the scorable achievement response lies in the strength of the expectancy brought to the picture by the subject. Following the arousal experience we would not expect the subjects to have a large repetoire of roughly equal competing expectancies. Rather there would be a few strong expectancies which may or may not be achievement related. Thus if a power expectancy has been aroused a power story will be written, and it will not be scored for achievement. The result of all this is a very wide range of achievement scores across subjects. And these differences prove diagnostic of actual effort in performance situations.

The relationship of ratings of pictures to story scores should now be clear. To the extent that the raters' judgments of pictures do relate to n Achievement scores we can attribute the score determination to the picture cues. And to the extent that scores appear that are not anticipated by the raters we can attribute the scores to motivational determination.

Procedure for Rating Pictures

The procedure for rating pictures used in this study was a modification of a procedure originally used by Jacobs (15). His procedure required that the subject first rank 14 "concerns" for appropriateness to the pictures. These "concerns" were statements of needs such as hunger, affiliation, power, etc. Then the subject was asked to rank 12 pictures for their appropriateness to concern for achievement.

He found that raters showed good agreement in their rankings. And when the pictures were ranked using the median ranking of the achievement concern from the first task the order correlated highly (rho +.95) with the rank order of pictures from the second task. He also reported there was no relationship between the n Achievement scores of his subjects and their rankings of the pictures.

The same basic procedure was used in the present study. Fifteen need related descriptions of the situation in the picture were placed in a rank order by the subjects. Next the subjects were asked to consider each picture as related to achievement and then to rank order ten descriptions of achievement-related action taken directly from the n Achievement scoring system. Then they ordered three degrees of affective tone, positive, neutral, and negative. Finally they were faced with ten pictures and asked to rank order them for appropriateness to an achievement situation; and lastly they repeated the same ordering of pictures on achievement-related affective tone.

These procedures provided two ways of obtaining rank orders of pictures on appropriateness to achievement situations. Thus we obtained for each picture the median value of the achievement statement from the list of 15 need statements. Then the pictures were ordered by these median values. This use of the ratings involves the following assumptions:

1. All subjects use the same, or nearly the same, frame of reference such that the ratings of achievement can be meaningfully compared.
2. All subjects interpret the definition given the motives in approximately the same fashion.
3. The fifteen motives are independent dimensions on which to judge a picture.

The second ordering of the pictures on achievement is obtained directly. Here the assumptions involved are:

1. The definition given achievement has approximately the same meaning for all subjects.

2. All pictures are being judged independently.

And a final assumption common to both procedures is that:

> the subjects are comparable as members of the same sample so that the rating values from different groups can be combined into a single order.

Now it is quite possible that the ordering of pictures obtained from the two rating tasks could disagree. Thus a picture could be very unsuggestive of anything, but to the extent that it has any cues in it they are relevant to achievement situations. Such a picture would have the achievement motive ranked high on the first task, but the picture itself would rank low on achievement when compared with other more suggestive pictures. In order for the two rating tasks to give identical results the pictures must be roughly equivalent in terms of their ranges of motive cues, i.e., each strong on some dimension, weak on others, but all equally suggestive of something. Fortunately this condition is subject to test.

Finally there is the problem of which ordering of pictures to use in comparing ratings with stories. We want the rating procedure which is most comparable to the story writing situation so that the two situations differ only in the crucial aspect of rating on a provided dimension versus free response in writing. In writing stories the subject looks at each picture for 20 seconds and then is given four and one-half minutes to write his story. At no time does he see all the pictures simultaneously. Thus the first task of ranking 15 motives to

a single picture is most like the story writing situation.

Administration of Rating Tasks

Five pictures were mounted on a large card and placed directly in front of five or more persons. Each person was given a booklet with instructions and pages for his rankings. The experimenter then read the instructions page to the subjects (see Appendix A). Each booklet contained the numbers of the pictures the subject was to consider, and these were ordered differently across subjects so as to minimize possible serial position effects. The subject then considered each picture in turn as he made his rankings. After doing four pictures the subjects were stopped, and the cards containing the pictures were combined to provide ten pictures in front of the subjects. They then turned to the last page where the between picture rankings on achievement strength and achievement-related affect were made. The raters were 150 college students in the introductory psychology course at the University of Michigan. Both male and female subjects were used.

Preliminary Analysis of Rating Data

First we considered the question of whether or not sex differences appear in the ratings. To get the most reliable estimate we used the between picture rank order for ten pictures which had been done by thirteen females and six males. The rank order correlation between sexes for the ten pictures was a rho +.85. Testing for significance by converting Kendall's tau (16) we find $P < .001$. Using Kendall's coefficient of concordance it was possible to determine the significance

level of interjudge agreement, and to compute an average rho value for the interjudge correlation. For these judges the coefficient of concordance was significant at the $P < .001$ level. And the average interjudge correlation converted to a rho of +.50.

Next we compared the rank order of the pictures obtained by the two types of rating tasks. Fifty-four pictures were rated and in order to obtain a complete ordering of all fifty-four pictures a considerable amount of combining of picture values across groups of subjects had to be done. That is, each subject rated just four pictures on the first task, and ten on the second task. To get a rank order of all the pictures required combining rank values across groups. This, of course, would introduce between-groups variability. Never the less we found the product moment correlation for the two orders of pictures to be +.61, which is significant at the $P < .001$ level. Using just 19 pictures, which introduces considerably less between-groups variance we got a product moment correlation of +.82.

Finally we had to estimate how well the obtained average of ten raters to each picture gave rankings comparable to those obtained from a larger number. Accordingly we took the previously mentioned sample of nineteen subjects, six males and thirteen females, and drew several samples of ten each. The amount of interjudge agreement for the small samples was found to be comparable to that of the total sample.[1]

On the basis of the preliminary analysis, rating data used in this study are based on (a) ratings of males and females, (b) a minimum

---

[1] See Table 8, Appendix B.

of ten raters per picture, and (c) the first ranking task. Data for a total of 54 pictures was obtained.

Selection and Use of Pictures

Since we wished to study the degree to which ratings of achievement cues in pictures can predict obtained achievement imagery, we must now consider how the n Achievement scores were obtained.

Because of practical limitations upon the number of pictures a subject can respond to in one sitting, three batteries of six pictures each were chosen from the total sample of pictures. Each subject was to write stories to six pictures. The three sets of pictures were arranged so as to be parallel on strength and affective sign according to the raters' judgments. In addition they were chosen to represent three social content areas for use in the second part of the study.

The six pictures of each set were arranged in a Latin Square design according to the procedure recommended by Fisher (12). The Latin Square requirement dictated one change in the standard administration procedure of the n Achievement measure. In the past the practice has been to flash the pictures on a large screen. Here, however, it was necessary to test as many persons as possible under the same working conditions. Therefore the pictures were run off on an ozalid process, and presented to the subjects in booklet form. This administration permits the simultaneous use of as many pictures in as many orders as the experimenter desires. The danger of such a procedure lies in the additional "testing" situation cues which such a booklet may present. Before the subjects arrived at the auditorium in which the testing took place

the booklets were placed in the seats by blocks of 18 so as to assure equal numbers of subjects across each sequence of pictures. Upon arrival the subjects were asked to be seated from front to back, presumably giving us a random assignment of subjects to booklets.

Procedure for Testing Session

In addition to the n Achievement measure the booklets contained three performance measures. The thinking behind their use and the hypotheses about them are presented in Chapter III. The three performance measures used were a Zeigarnik task used by Atkinson (17), consisting of of twenty paper and pencil tasks, an Additions task composed of ten pages of thirty-six columns of three two-digit numbers, and a Word Coding task consisting of seven pages of four letter words, forty to a page, to be coded according to the numbers beside them (see Appendix A).

The arrangement of the tasks in the booklet was Zeigarnik task, n Achievement measure, recall sheet for Zeigarnik task, Additions, and Word Coding. All the tasks were timed with a stop watch, and all but the n Achievement measure were presented as speed tests. The entire administration of the booklet took one hour and thirty minutes.

Turing to the actual procedure followed in the testing session, the initial remarks to the subjects (male college students) were as follows:

> Open the booklet and look at the first sheet. On it put the information asked for. Your name is being taken for attendance purposes only, and the code number you give will be used to label your data. If you do not know your I. Q. then estimate it, bearing in mind that the estimated mean for adult population of the country is 100. (A normal curve was drawn on the board by way of illustration.)
> When you finish wait until I tell you to turn the page. The

steps in this booklet must all be run with a stop watch, so wait for instructions for every page.

All right, turn the page and read the instructions as I read them to you. (E. then read instructions. There were one or two questions on minor matters.)

The Zeigarnik test was then administered. Following the task the experimenter continued:

Now relax a moment and prepare to read the instructions for the next task with me. (E. then read instructions, emphasizing the point about not looking back at the picture, or turning ahead to the next one. Every effort was made to point out that this was not a speed test, but rather was timed to assure standard administration.)

The n Achievement measure was administered without incident, and then the subjects were told:

You have reached the half way mark but before we take a short rest there is one more thing to do. On the blank sheet of paper now in front of you, list, as they come to your memory, the tasks you did on the first test. Identify them any way you wish so long as we will recognize what you mean. (The subjects were given four minutes for this task.)

A short, three minute, rest period followed with the subjects remaining in their seats. Then the experimenter continued:

Now turn to the task marked Additions, and read the instructions with me.

Following the administration of the Additions task the subjects moved directly to the Word Coding task, and once again read the instructions with the experimenter. Before beginning the Word Coding task the experimenter remarked "I know you are getting tired now, but these tests work best when people are slowing down, so do your best." Then at the fifth trial he remarked "Two more trials to go!"

Each criterion task had a short time limit on each page, and the time was carefully kept with a stop watch by the experimenter who told the subjects when to proceed to the next page. There were far more

columns of addition on each page than anyone could do, but the page of coded words could be done in the time allotted, and a few subjects did it.

The first testing session was given to 62 male students. It was held in a large auditorium, with subjects spaced out in every other seat as is customary during final examinations. Two proctors assisted the experimenter. The second session was given two weeks later in a smaller classroom, with eleven college males and the experimenter present. The subjective impression of the experimenter was that both sessions were marked by high achievement involvement on the part of most subjects. There was very little restlessness although the tasks were long and tedious, and no expressed hostility. Some indication of the interest of the subjects may be had from the fact that the two hour session was followed by an explanation of the experimenter's purposes and measures. Only one man chose to leave before this explanation. All the subjects were either outright volunteers, or had chosen this experiment from several so as to fill a course requirement.

Statement of the Hypotheses

With this introduction to the problem let us now proceed to the statement of the hypotheses about the relationship of picture ratings to picture stories.

Since the actual scoring categories of the n Achievement system were empirically developed and validated the raters were asked to rate the pictures with reference to the actual scoring categories. The categories were transcribed into simple statements and the subjects asked

to order them for appropriateness to the picture. In the second rating task the subjects were told to consider several pictures with regard to suggestiveness of achievement themes, and to rank order the pictures. For a copy of the actual rating booklet see Appendix A. These rating operations provided the definition of the achievement cues in the pictures. We then attempted to predict from these data the amount and kind of achievement imagery which appears in stories written to the pictures.

The general hypothesis from which these statements derive is that picture cues do determine, in part, the amount and kind of achievement imagery. The specific hypotheses then read as follows:

Hypothesis 1:

> There will be a positive correlation between the rank order of pictures by judges' ratings of achievement strength and the rank order of pictures by the percentage of stories scored for n Achievement.

Hypothesis 2:

> There will be a positive correlation between the rank order of pictures by judges' ratings of achievement related affect and the rank order of pictures by the proportion of positive affect scores.

Hypotheses 3:

> There will be positive correlations between the rank order of pictures by judges' ratings of the sub-categories of the achievement scoring system and the respective rank orders of pictures by the percentage of such scores obtained from stories scored for achievement.
>
> Note: This hypothesis must be restricted to the upper two-thirds of the strength distribution of pictures since the low third of pictures on strength will provide too few sub-categories.

Results

Reliability of Scoring

The n Achievement scores for each subject, the pictures used, and the order of pictures appear in Appendix C.

The experimenter's level of consistency and reliability in the use of the n Achievement scoring system was established by his scoring stories of 27 subjects used in a previous study. These stories had been scored by an experienced judge[1] and the rank order correlation between the two n Achievement score distributions proved to be +.92 (rho).

Latin Square Data

The Latin Square design on each of the three sets of pictures had been used to attempt to control for any systematic positional or sequential differences in n Achievement score that might occur. The assignment of subjects was thought to be random across the sets of pictures, and we assume no systematic subject differences between groups. These procedures then permit us to deal with differences in n Achievement score distributions between pictures as being solely due to effects of the pictures themselves. Let us then examine the relative performances of the three groups on both the achievement measure and the criterion measures to determine whether there are significant differences in motivation between groups.

Table 1 shows the Latin Square data in terms of mean n Achievement score with a constant of two added in each individual score, for each position-by-sequence cell in the design. Along the base we find the mean

---

[1]Professor J. W. Atkinson.

TABLE 1

THE LATIN SQUARE DESIGN

(Pictures are designated by (no.). N for each cell is indicated at the left.)

Group	Sequence	N	Ordinal Position 1	2	3	4	5	6	Total	Mean
3	1	4	(53) 2.75	(34) 4.00	(2) 2.25	(69) 1.00	(12) 1.75	(64) 1.75	13.50	2.25
	2	4	(2) 1.75	(53) 2.00	(12) 2.75	(34) 3.25	(64) 1.75	(69) 1.00	12.50	2.08
	3	5	(69) 1.00	(12) 2.20	(53) 3.60	(64) 3.60	(2) 4.20	(34) 3.80	18.90	3.07
	4	5	(64) 1.80	(2) 1.00	(34) 5.60	(53) 2.60	(69) 1.20	(12) 3.80	16.00	2.67
	5	5	(34) 3.20	(64) 4.00	(69) 1.20	(12) 1.00	(53) 3.80	(2) 2.20	15.40	2.57
	6	4	(12) 4.25	(69) 1.00	(64) 2.50	(2) 2.75	(34) 2.25	(53) 1.25	14.00	2.33
	Total	27	14.75	14.20	17.90	14.20	14.95	13.80	89.80	
	Mean		2.46	2.37	2.95	2.37	2.49	2.30		2.49
1	1	4	(66) 4.75	(17) 1.00	(54) 1.00	(24) 3.75	(21) 2.50	(41) 1.50	14.50	2.41
	2	4	(41) 2.50	(21) 1.50	(66) 6.75	(17) 3.75	(54) 1.25	(24) 5.50	21.25	3.54
	3	3	(24) 3.67	(66) 4.33	(41) 3.67	(54) 1.00	(17) 1.00	(21) 2.00	15.67	2.61
	4	4	(17) 2.00	(54) 2.00	(24) 2.25	(21) 2.50	(41) 1.50	(66) 4.75	15.00	2.50
	5	3	(21) 1.00	(24) 3.67	(17) 2.00	(41) 2.67	(66) 6.00	(54) 1.67	17.01	2.83
	6	5	(54) 1.20	(41) 1.80	(21) 1.80	(66) 4.60	(24) 3.80	(17) 2.40	15.60	2.60
	Total	23	15.12	14.30	17.47	18.27	16.05	17.82	99.03	
	Mean		2.52	2.38	2.91	3.04	2.67	2.97		2.75
2	1	4	(50) 1.00	(32) 3.75	(7) 4.25	(14) 2.75	(4) 4.50	(19) 1.25	17.50	2.91
	2	4	(32) 4.75	(14) 4.25	(50) 2.25	(7) 4.75	(19) 1.75	(4) 2.75	20.50	3.41
	3	3	(7) 1.33	(19) 2.33	(4) 1.00	(50) 1.00	(32) 3.00	(14) 3.00	11.66	1.94
	4	5	(19) 2.40	(7) 2.40	(32) 3.80	(4) 2.00	(14) 3.80	(50) 1.00	15.40	2.57
	5	3	(4) 2.67	(50) 1.00	(14) 3.00	(19) 1.33	(7) 4.00	(32) 5.67	17.67	2.94
	6	3	(14) 5.33	(4) 3.67	(19) 1.00	(32) 6.33	(50) 1.00	(7) 4.67	22.00	3.67
	Total	22	17.48	17.40	15.30	18.16	18.25	18.34	104.73	
	Mean		2.91	2.90	2.55	3.03	3.01	3.06		2.91
Pictures			(50) 1.21	(32) 4.55	(7) 3.57	(14) 5.69	(4) 2.77	(19) 1.68	(66) 5.19	(17) 2.03 (54) 1.35
			(24) 3.77	(21) 1.88	(41) 2.27	(53) 2.67	(34) 3.68	(2) 2.36	(69) 1.06	(12) 2.63 (64) 2.57

n Achievement score per picture, also with the constant added. Inspection of Table 1 shows that sizeable picture differences do exist within all three sets of pictures. No systematic trends appear in either the ordinal or sequence means, and a test of the difference between the most disparate group means shows no reliable difference.

Since we are using the performance measures as indices of motivation we will also present the group differences in performance in Table 2. We see that there are no reliable differences on the Word Coding task, the only task on which a sizeable difference between group means appears. The difference between Groups 2 and 3 fails to reach an acceptable level of significance ($t = 1.33$, 46 df). In the light of these data the analysis will proceed with the assumption that our control procedures have succeeded.

Test of Hypotheses

Hypothesis 1, as stated formally on page 29, predicts a positive correlation between the rated strength of a picture and the actual amount of achievement themes written to it. Since the number of persons in each group differed we computed the percentage of persons writing achievement themes to each picture. This percentage was used to construct a rank order of the pictures on actual strength. Then the median of the judges' rating of each picture on n Achievement was computed for each picture, and the pictures ordered according to the median values. These two distributions were correlated using Kendall's tau (16), a rank order technique. Our interest in framing this hypothesis was in determining how well ratings of picture cues could substitute for actual score definitions of the picture cue values. Therefore the tau formula which

TABLE 2

MEAN PERFORMANCE SCORES ON CRITERIA TASKS

Zeigarnik Task

Group	Mean Recall Scores* Completed	Incompleted	Zeigarnik Index
1	5.25	5.24	-.01
2	4.91	5.02	+.12
3	4.79	4.82	+.03

Additions Task

Group	Mean Total Performance	Sigma
1	136.6	33.9
2	141.6	33.0
3	135.1	32.0

Word Coding Task

Group**	Mean Total Performance	Sigma
1	168.4	47.7
2	158.7	37.3
3	173.7	45.9

*Arc sin transformation of percentage tasks recalled, in turn transformed to z scores, with a constant of 5 added to give positive values.

**A test of the difference between Groups 2 and 3 yields t = 1.33, 46 df. A t value of 2.02 is significant at the .05 level with 46 df.

makes no allowance for ties was used, giving us an index of reproduction of one order by the other, rather than an index of mutual agreement between orders.

The value for tau_a in Hypothesis 1 is

$t_a = +.39$ $P < .03$, one tailed, $n = 18$

Hypothesis 1 is confirmed.

The size of this relationship is sufficiently small to suggest that the use of ratings alone to predict the likelihood of obtaining achievement themes from pictures would lead to considerable error. However the position that cognitive ratings should not be able to predict projective material is not sustained by this result. Even projective material has some systematic relationship to the stimuli before the subject.

Hypothesis 2, stated formally on page 29, predicts a positive correlation between judges' ratings of the positive or negative quality of the cues of achievement related affect, and the actual n Achievement affect scores. Because the weaker pictures elicited so few stories scored for n Achievement, providing an unstable estimate of the proportion of positive affect signs to the total scored, this correlation was run on the upper two thirds of the strength distribution, giving an n of 12. The same ranking procedures were used here as in the first hypothesis, but it must be noted that the sign index reflects the proportion of signs in the population of <u>scored</u> stories.

The value of tau_a in Hypothesis 2 is

$t_a = +.59$ $P < .001$, one tailed, $n = 12$

Hypothesis 2 is confirmed.

Thus it would appear that the raters were even more accurate in their estimates of the quality of the achievement related affect cues in the pictures.

The third hypothesis, stated formally on page 29, says that raters should also be able to predict the various amounts of achievement related statements of need, instrumental activity, and the like which go to make up the n Achievement scoring system. It is a fact that some of these categories are habitually rare. Four categories[1] appeared too rarely to permit a test of their predictability. Once again, for the reasons given previously, the upper two thirds of the pictures on the strength distribution were used, n = 12.

The values for the tau_a's were

Instrumental Activity[2] positive outcome	$t_a = -.08$	
Instrumental Activity negative outcome	$t_a = +.14$	
Need	$t_a = +.03$	
Goal Anticipation positive	$t_a = +.11$	
Goal State, positive	$t_a = +.50$	$P < .01$, one tailed
Goal State, negative	$t_a = +.24$	$P < .15$, one tailed

The only categories predicted by the raters are those which provided most of the scores used in the test of Hypothesis 2. It would appear that the raters cannot anticipate the exact instrumentalities and achievement related statements used in the stories as well as they can judge the more global aspects of the theme.

---

[1]The categories were Goal Anticipation, Negative; Personal Block; Environment Block; and Nurtrant Press.

[2]For detailed definitions of these categories see The Achievement Motive, pp. 110 (17).

In general Hypothesis 3 is infirmed.

This concludes the results reported for the efforts to determine how well raters could predict the actual scores made when subjects wrote stories about the pictures. The median values of the ratings appear in Table 5 of Appendix B. The discussion of these results will appear in Chapter IV.

CHAPTER III

PICTURE EFFECTS ON THE VALIDITY OF n ACHIEVEMENT SCORES[1]

The second approach to determining picture effects on projective imagery is to test the n Achievement score distributions in terms of their ability to predict performance in an achievement situation. There are a variety of ways to do this, and the particular hypotheses set forth here simply represent an effort to draw on three very broad classes of possible relationships between n Achievement scores and their picture sources. We can view picture stimulus characteristics in terms of their psychometric qualities, their social meaning, or their particular achievement-stimulus character. In this chapter an hypothesis from each of these areas is set forth with a brief statement of its origin, and the results pertaining to it.

The Performance Measures

The most reliable performance measures predicted by the n Achievement score to date have been the Additions and Scrambled Word tasks of Lowell (17). Lowell used the Additions task to tap persistence and speed of performance, arguing that simple addition was overlearned for college students. The Scrambled Word task was thought to require some immediate learning, and the high n Achievement group showed

---

[1] A tabulation of raw performance scores on the criterion tests appears in Appendix C.

the most rapid improvement in performance.

In the present study a similar Additions task was used,[1] and what appeared to be an improved learning task, called a Word Coding task.[2] In addition to these two tasks the third performance measure used was the Zeigarnik task used by Atkinson (17). This consisted of 20 paper and pencil tests, half of which were interrupted before completion. Atkinson had found that under achievement oriented conditions there appeared to be a positive relationship between his Zeigarnik index, which expresses the degree to which recall of incompleted tasks is favored over completed tasks, and the n Achievement score.

Processing of the data revealed that one subject had not completed the last two tasks, and five subjects had failed to give complete data on some one of the tasks. Therefore the total N of persons with data on all three tasks was sixty-seven. Using these data the rank order correlation, tau_b, between the three measures was computed. This formula yields a coefficient of agreement.

The values of tau_b for the criteria are shown below.

Correlation between Criteria (N = 67)

Tasks	tau_b	One tailed level of significance
Word Coding—Additions	+.30	$P < .001$
Word Coding—Zeigarnik	+.08	$P < .17$
Zeigarnik—Additions	+.12	$P < .09$

[1]Due to an editorial error in Lowell's report three two-digit numbers were used instead of two three-digit numbers.

[2]This task was shown to be related to n Achievement scores for college students in two studies by the Conformity Project at the University of Michigan.

These correlations indicate some consistency between the last two measures of the test session, Additions and Word Coding.

Tests of Hypotheses

Since strength of cue, operationally defined as the likelihood of eliciting achievement imagery, had been measured it is now possible to determine whether or not the predictive efficiency of the n Achievement scores could be shown to vary with the strength of the pictures. In mental test theory it has been shown that the most discriminating items are those with the greatest variance, validity being held constant. Since we have pictures which are analogous to test items in the sense that some elicit stories from everyone, some from a moderate number, and some from few or no one, we are in a position to see whether pictures behave like items. We would expect that pictures of medium strength would create the greatest spread of subjects and give the most reliable discrimination on whatever the scores measure. And we must define the term medium here as that level of strength which elicits stories from 50 per cent of the subjects.

In order to test this hypothesis we must order the pictures in terms of the predictive capacity of their scores. The analysis procedure used consisted of computing the tau_b values for the n Achievement score distribution of each picture with the total score on each of the two tasks, Additions and Word Coding. Then we ordered the tau_b values from a high positive value, which indicates positive prediction of performance on the criterion task, down through zero to negative values which indicate negative prediction of performance. This distribution

was then correlated with the distribution of the pictures on strength from median strength down through high and low values (see p. 41 for details); and here we used both the rated and obtained score definitions of strength.

When we computed the tau's for the pictures as just described, we found the distribution ran from +.17 to -.24 for the Word Coding task, and from +.21 to -.20 for the Additions task.[1] Both distributions are symmetrical about zero. It was expected, of course, that this distribution would have its midpoint at a positive value at least approaching significance. These distributions, rather unexpectedly, resembled random distributions of coefficients. It now remains to be seen whether we can account for the obtained order of pictures with independent predictor variables. The problem to be faced is whether we actually have some pictures yielding positive predictions while others yield negative predictions, or whether we have a collection of pictures yielding unreliable scores.

The general hypothesis that medium strength pictures would be the best discriminators makes no assumption about the direction of the predictions. Therefore we can still determine whether those pictures whose scores yield the highest tau values tend to be pictures of medium strength.

Hypothesis 4 is stated as follows:

> Pictures of medium strength will yield scores with the greatest relationships to performance.

In an absolute sense a medium strength picture is one which elicits stories from 50 per cent of the subjects. To test our hypothesis

---

[1] See Table 7, Appendix B.

we ordered the pictures by computing the differences between a 50 per cent value on strength, and the computed value on strength. The smallest difference was given a rank of one, and the largest difference the highest rank. Thus a very strong picture and a very weak picture equidistant from the 50 per cent point would be tied in rank order position. This was done for both definitions of strength (ratings and scores). The orders so obtained may be seen in Table 6 of Appendix B. These orders were then correlated with the rank order of pictures in terms of the absolute size of the correlation of scores with performance, regardless of direction. Thus the pictures with scores yielding taus of +.20 and -.20 are tied for the highest rank.

We have three tasks and two definitions of strength giving us six tau values. The exact treatment of the Zeigarnik task data will be reported for Hypothesis 6. We simply report here the ability of medium strength pictures to predict the index.

The tau_b values obtained are as follows:

Strength (rated)	tau_b	
Additions task (N = 18)	-.12	
Word Coding task (N = 18)	-.26	P <.14, two tailed
Zeigarnik task (N = 15)	+.06	

We see here that the hypothesis that medium strength pictures will predict criterion performance better than strong or weak pictures is not confirmed from the rating definition of strength.

Using strength defined by obtained scores the values of tau_b are as follows:

	tau_b	
Additions task (N = 18)	+.33	P < .03, one tailed
Word Coding task (N = 18)	+.18	P < .14, one tailed
Zeigarnik task (N = 15)	+.20	P < .16, one tailed

Here it appears that those pictures eliciting stories from roughly 30 per cent to 60 per cent of the subjects yield more discriminating scores than those whose percentages fall outside this range. For the Additions task we can say the hypothesis is confirmed, and the fact that the other two tasks fall in the predicted direction is taken as support for not rejecting the hypothesis.

In summary then, the hypothesis is infirmed with respect to a definition of picture cues by ratings, but confirmed with respect to scores derived from written stories.

Hypothesis 5 is drawn from a consideration of the social content of the picture as judged by the experimenter. The pictures used in the study were selected from three content classes, school situations, white-collar situations, and blue-collar situations. This was done to permit the test of an hypothesis which follows rather clearly from certain basic assumptions about the n Achievement score.

The work of Atkinson (17, p. 264) suggested that the closer the n Achievement testing situation was in cue value to the performance situation, the more discriminating the n Achievement scores become. Thus under relaxed achievement conditions the high n Achievement group actually completed fewer Zeigarnik tasks than the low n Achievement group. However when the situation was made an achievement oriented one this relationship was reversed. From this finding it appears that the n Achievement score should be obtained in the presence of cues as similar to the criterion situation as possible.

Our testing session used only the achievement oriented instructions. However the cues in the pictures do vary in the degree to which

they correspond to the nature of the testing situation, and it is possible to determine whether cues in the pictures operate in the same fashion as cues in the environment.

Figures 1-3 show the three sets of pictures used and gives their code number and content designations. Since it has already been demonstrated that strength of picture, defined by obtained scores, is important, a fair test of the content hypothesis would demand that the three content areas be parallel in obtained strength. Unfortunately this is not the case. School pictures tend to be either high or low in strength. This tends to work against the hypothesis.

Hypothesis 5 is stated as follows:

> School pictures will yield n Achievement scores that relate more positively to the three criteria than will scores obtained from the Blue-Collar pictures. White-Collar pictures will fall in an intermediate position.

In order to test this hypothesis the three content sets of pictures were ordered in a trichotomous fashion (School-White collar-Blue collar, three rank values with six ties each), and this order correlated with the order of pictures in terms of the tau values expressing the relationship of the n Achievement scores to the criteria.

The tau_b values were found to be as follows:

Task	tau_b (N = 18)
Additions	+.04
Word Coding	-.17
Zeigarnik	-.08

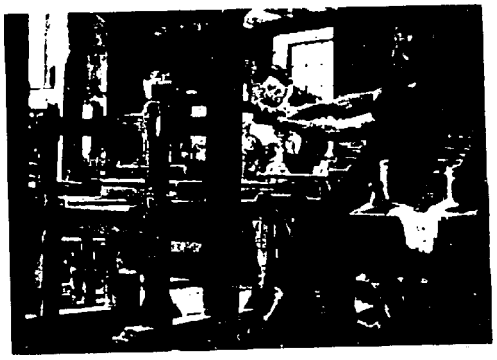
Hypothesis 5 is infirmed. However in the light of the failure to obtain parallel groups on strength it cannot be said the hypothesis received an adequate test.



Pic. 41, Teaching Fellows



Pic. 21, Large Class



Pic. 24, Colonial Printers



Pic. 17, Office Boy



Pic. 66, Pasteur



Pic. 54, Post Office

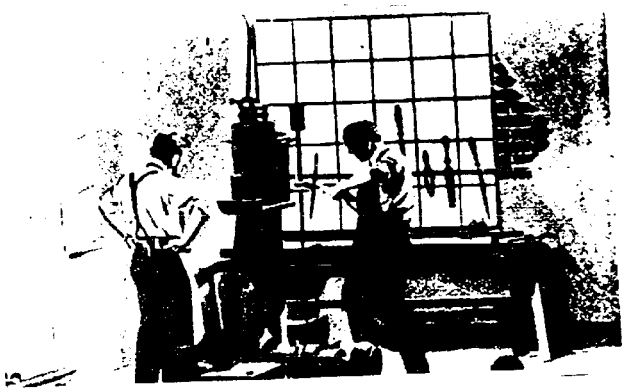
Fig. 1. Set of Pictures Shown to Group 1.



Pic. 32, Potter



Pic. 19, Four Men



Pic. 14, Inventors



Pic. 50, Small Schoolboy



Pic. 7, Boy-Notebook



Pic. 4, Father-Son

Fig. 2. Set of Pictures Shown to Group 2.



Pic. 69, Man-Phone



Pic. 34, Law Students



Pic. 12, Boys-Shop



Pic. 2, Boy-Classroom



Pic. 53, Elderly Workman



Pic. 64, Salesman

Fig. 3. Set of Pictures Shown to Group 3.

Description and Treatment of Zeigarnik Task Data

The last hypothesis to be tested has to do with the Zeigarnik data previously reported by Atkinson (17, p. 264). A number of studies have suggested that a curvilinear relationship between the n Achievement score and performance may be expected in achievement oriented situations. Persons getting scores in the middle range have acted in such a fashion as to suggest that they are motivated more by a fear of failure than by hope of success (17). One of these specific findings reported occurred in Atkinson's work with the Zeigarnik effect.

The Zeigarnik task used by Atkinson, and in this study also, consisted of 20 paper and pencil items to be done as quickly as possible by the subjects. Actually ten items were very easy, and ten were very long or even insoluble. This was designed to give a condition in which half the tasks were completed and half left incompleted. The Zeigarnik index is an expression of the tendency to recall a greater percentage of incompleted tasks than completed tasks.

Atkinson reports that as the situational cues became more achievement oriented

> . . . the relative preponderance of completions over incompletions recalled increases significantly for this middle group . . . . between Relaxed and Achievement orientation (17, p. 270).

Thus the positive relationship between n Achievement score and the Zeigarnik index is <u>reduced</u> by this behavior on the part of the middle third of the n Achievement distribution. And these subjects are said to have expectancies of failure. Now Moulton (19) was able to show that these subjects also tended to write negative achievement imagery,

and many escape oriented stories to Picture 2, the boy in the classroom. Thus if we could argue that certain cues in pictures would tend to elicit from these subjects escape stories, which do not get scored under the n Achievement scoring system, we would expect the relationship between n Achievement and the Zeigarnik index to be greater.

It seems reasonable to argue that the pictures most likely to elicit non-scorable stories of escape and despair from persons with strong expectancies of failure would be those with strong negative affect cues. These cues would tend to confirm their thoughts of failure. It is from this assumption then that we derive the following hypothesis.

Hypothesis 6:

> Pictures with negative achievement-related affect cues (fear of failure cues) will predict the Zeigarnik index better than will pictures with positive achievement-related affect cues (hope of success cues).

Once again a fair test of this hypothesis requires that negative and positive cued pictures be roughly parallel in strength. The rated values were chosen so as to assure that this was the case, but the obtained values of strength show that the medium strength pictures tend to be of positive sign value. If we compute the correlation between the order of pictures on affect sign, plus to minus, and the order on strength (obtained scores), we find that

$$tau_b = +.26 \quad P < .30, \text{ two tailed } (N = 12)$$

So that there is once again a condition which tends to work against the hypothesis.

The Zeigarnik index used in testing this hypothesis is designed to compare relative preference for recall of Incompleted tasks by the

high and low halves of the n Achievement distribution for each picture. Thus the formula is

$$Z = (IR_{hi} - CR_{hi}) - (IR_{lo} - CR_{lo})$$

where

IR = mean number of Incompleted tasks recalled.

CR = mean number of Completed tasks recalled.

hi = upper half of the n Achievement score distribution.

lo = lower half of the n Achievement score distribution.

We test the hypothesis by computing the tau_b values for both rated and obtained score definitions of achievement-related affect. Thus the rank order of the pictures of the Zeigarnik index values was first correlated with the order of pictures given by the raters' judgments of achievement related affect, and then with the order given by the proportion of signs from stories written to the pictures. The values of tau_b are as follows:

Source	tau_b (N = 12)	
Ratings	-.34	
Scores	-.30	$P < .11$, two tailed

Hypothesis 6 is infirmed. And, indeed, there seems to be some tendency for positive cue pictures to provide scores giving the most positive relationship to the Zeigarnik effect. This is consistent with the previous finding that the positive cued pictures tended to be of medium strength, which in turn had been shown to give the best discrimination between subjects.

Additional Analysis

In the light of the above results the question of whether

picture cue differences effect the usefulness of the n Achievement scores for predictive purposes remains open. It is clear that picture cue definitions in terms of the raters' judgments of strength and sign, though correlated with obtained n Achievement scores, provide little or no support for the hypotheses tested here. And we must still face the question of whether the correlations of n Achievement scores with performance represent a random distribution. Up to this point only the finding that the medium strength pictures gave the best predictions of performance, positive and negative, suggests that we are not dealing with a random distribution of correlation coefficients.

The results of two more analyses may be brought to bear on this question. It is felt that they give some additional support for the notion that the correlations through low and symmetrical about zero, are meaningful.

First it seems reasonable to expect that random correlations would not show consistency across criteria. That is, we may ask how the picture, whose scores predict Additions performance, fares when asked to predict Word Coding or the Zeigarnik index. Table 3 presents the pictures ordered roughly in terms of their consistency of rank across criteria. We see that pictures 66, 34, and 24 (see Figures 1 and 3). yield scores which predict most negatively to all three criteria. (Pictures 54, 69, 50 yield so few scores that the correlations for all three criteria are zero, and they have been assigned the mid rank positions for zero in the distributions.) We see further that four pictures, 14, 17, 21, and 12 give the most positive predictions to Word Coding and Additions, which are significantly correlated.

TABLE 3

PICTURES RANKED BY CONSISTENCY OF SCORE CORRELATIONS WITH CRITERIA SCORES

Picture	Rank on Word Coding	Rank on Additions	Rank on Zeigarnik Index
14	3	1.5	10
17	5	1.5	11
7	4	13	3
12	2	4	14
21	7	3	16
19	18	5	4
4	6	15	2
32	1	17.5	6
64	8	14	1
54	10	9	7
69	10	9	7
50	10	9	7
53	12	7	12
42	13	6	18
2	14	12	5
66	15	16	15
34	17	11	13
24	16	17.5	17

Thus we find that seven of the fifteen pictures for which correlations could be computed showed consistency of position despite the fact that the absolute size of the correlations is not large (see Table 7, Appendix B).

Another analysis was suggested upon rereading the stories written to the pictures. It seemed to the experimenter that the pictures whose scores predicted negatively to performance had themes of an historical or emotional type which sounded the same from one subject to another. Therefore three of the types of imagery suggested in the scoring system plus a fourth emphasizing historically oriented

achievement themes were set out.

The definitions of the four types are as follows:

1. "One of the characters in the story is engaged in some competitive activity . . . . (17, p. 111)."
2. "One of the characters in the story is involved in attainment of a long-term achievement goal (17, p. 113)."
3. Some one in the story is "affectively aroused over goal attainment (17, p. 112)."

The fourth type imagery is not recognized in the scoring system. However achievement stories about historical figures are not rare. Therefore the fourth type of imagery was defined as:

4. Someone in the story is mentioned as a well known figure in history.

Each story was then read and classified into one of the four types of imagery. In the case of multiple possibilities, which occurred occasionally, the most dominant theme was favored. The classification was done twice, with a ten day interval between, and the per cent of agreement for assigning the stories to types was .91.

Treatment of this data consisted of tallying the frequency of each type of imagery for each picture, converting the totals to percentages of the total imagery scored, and permuting the four columns of types to give the most consistency across pictures. Table 4 shows that pictures did differ in the types of imagery they elicited. It was decided to test for a relationship between type of imagery and predictive relationship to the criteria by combining types 3 and 4, and computing their percentage of the total amount of imagery. This

TABLE 4

FREQUENCY AND PERCENTAGE OF EACH TYPE OF IMAGERY FOR EACH PICTURE

Pictures	Imagery Types							
	1		2		3		4	
	f	%	f	%	f	%	f	%
14	3	.30	5	.50	0	-	0	.20
17	5	.84	1	.16	0	-	0	-
7	5	.50	4	.40	1	.10	0	-
12	2	.33	4	.67	0	-	0	-
21	0	-	2	.50	2	.50	0	-
19	0	-	1	.33	2	.67	0	-
4	5	.625	0	-	3	.375	0	-
32	4	.33	12	.67	0	-	0	-
41	0	-	2	.50	2	.50	0	-
64	3	.375	3	.375	2	.25	0	-
53	3	.25	8	.67	1	.08	0	-
2	1	.20	0	-	4	.80	0	-
66	4	.225	4	.225	1	.05	9	.50
34	10	.72	2	.14	2	.14	0	-
24	0	-	3	.23	0	-	10	.77

gives a single percentage figure with which to order the pictures. This order was then correlated with the previous orders of pictures, for each criterion, which expresses how well the picture scores yield the task performance.

The tau_b values obtained are as follows:

Task	Imagery (N = 15) $\frac{3+4}{\text{total imagery}}$	Level of Significance
Additions	-.21	$P < .42$
Word Coding	-.54	$P < .0006$
Zeigarnik	-.13	

These results are interpreted to indicate that the pictures whose

scores predicted negatively to performance tended to give the greatest percentages of Imagery types 3 and 4, "affective concern over an achievement goal," and "well known historical figure." Thus it appears that invalid pictures have something in common which is not present in the pictures whose scores tend to predict in a positive direction. And this too, is taken to mean that the distribution of correlation coefficients is meaningful and not random.

This completes the analysis of the data in the second portion of this study. In general it seems clear that scores from individual pictures do not predict performance well, and that the obtained score characteristics of the pictures did not provide us with three parallel sets of pictures. Possible reasons for these results will be discussed in Chapter IV.

Summary of Results

1. The hypothesis that medium strength pictures yield scores which gave the best predictions to performance tended to be sustained across criteria, with the correlation for Additions reaching the $P < .03$ level of significance.

2. The hypothesis that pictures of school situation would yield the most predictive scores was not sustained.

3. The hypothesis that pictures with negative affect would yield the most predictive scores on the Zeigarnik task was not sustained.

In addition to the test of these hypotheses it was found that correlation coefficients between picture scores and performance, though low and symmetrical about zero, did show some consistency across the criteria. And the negative and positive predictors tended to yield different types of imagery.

The failure of the obtained scores to match the parallelism of

the three sets of pictures given by the ratings meant that the last two hypotheses could not be tested with strength held constant. This, plus the possible presence of a negative form of n Achievement score, rendered a rigorous test of the last two hypotheses impossible.

CHAPTER IV

DISCUSSION

McClelland and associates discuss the effect of picture cues on written imagery as one of three sources of protocol determination, the other two being personality variables and situational cues. And the n Achievement score can reflect any or all of these factors in a variety of combinations. The point is also made in their discussion that the predictive capacity of the n Achievement score may vary considerably as a function of the source of the score's determination. With this last point in mind we can now turn to the interpretation of our findings.

It was proposed in this study that pictures could be given their stimulus value definitions by using a cognitive measuring operation, and it was shown how a rating task qualified as such. It was then argued that we would be able to obtain an estimate of the "confirming" qualities of picture cues. That is, we would be defining picture cues in terms of expectancies for which they would be most relevant. Under such conditions we expected that the raters' judgments would predict what the subjects would write in the stories. We found that this was the case for achievement related affect, and achievement strength. However the details of the achievement themes could not be predicted by these rating data.

The prediction of affect was much more accurate than the

prediction of strength. Since researchers on the TAT have reported fairly stable norms with respect to affect this is consistent with previous findings. A possible explanation for the accuracy of the raters may be that the affective tone of a picture is fairly clear, and generalizable across most themes written to the picture. That is, any story written to the picture would tend to reflect its general affective tone.

The prediction for strength was not as accurate, and here we must face a complicated problem. Some of the misses of the raters would be cases of underestimation of the strength of the picture. Thus a picture rated low on achievement may be rated higher on power, or influence, etc. Picture 4, the father-son pictures, is a good example of this. The achievement stories written here may be interwoven with themes of power and influence, and these themes are rated highest. Thus an index of picture strength derived from a combination of ratings whose intercorrelations are known might give us more reliable estimates of strength.

However a more fundamental issue remains. We dealt with pictures in terms of what they confirm. We made no allowance for their own arousal capacity. That is, it is conceivable that a person with a strong affiliative expectancy may face a picture with so strong an achievement cue in it that he writes an achievement story. Now we would not expect that a score determined by the picture would predict performance because it is an indication of an hypothesis instead of a motive. It merely demonstrates that the subject can recognize an achievement situation when he sees one, and it makes his task very similar to that

of the rater. This line of argument suggests that very high correlations between raters and writers may only be obtained at the expense of invalid scores.

Our obtained correlation between the two estimates of strength is sufficiently small (tau_a = +.39) to suggest that the n Achievement scores may be motivationally determined, and do not merely reflect the ability of our subjects to recognize achievement cues. Thus we can treat differences in mean n Achievement scores between pictures as, at least in part, differences in confirming capacity, and not just as differences in arousal capacity. Now when we tested our hypotheses concerning the effects of picture cues on the utility of n Achievement scores we found that if we defined picture cues by ratings the hypotheses were infirmed. Whereas in the case of the hypothesis that medium strength pictures would provide the most discriminating scores, the hypothesis was confirmed when picture cues were defined by obtained scores. However the hypotheses concerning the role of content, and of negative cues, also failed despite definition of picture cues by obtained score. We must then consider the possible reasons for this null result.

First of all such relationships can only be tested when the other two determinants of the n Achievement score are held constant within the testing situation. And there is evidence that we did this, see page 32. Secondly, the subjects' level of motivation should be optimal for producing the widest possible individual differences between subjects. And it is here that we may have failed. Research reported after this study was done indicates that our condition may have reduced the performance differences between the high and low n Achievement groups.

Burdick (6) found that volunteer subjects tended to have higher n Achievement scores than non-volunteers. We used a majority of volunteers while Atkinson and Lowell each used entire college classes made available by their instructors. Wendt (31) found that a paced situation, in which timing devices were used, tended to depress the n Achievement score differences of the high and low n Achievement groups over those differences obtained in a more leisurely, unpaced task. This is attributed to the fact that even persons with low motivation can be made to work hard if the situation is made pressing enough. And of course one last fact that is important in the interpretation of our performance data is that we are trying to predict the performance from scores from a single picture, while all the other studies used scores from three to six pictures. There can be no doubt that we are dealing with less reliable scores.

All of the above facts indicate that we are working with less reliable performance differences than those obtained in previous studies. We tried to duplicate Lowell's conditions, and wanted to heighten the involvement so as to arouse fear of failure in some subjects. But in doing so we of necessity introduced situational cues wthat worked against the findings. This illustrates very well the difficulties of research of this type, and indicates the nature of the controls required. Never the less we have argued that the distribution of correlation coefficients indicating the pictures' ability to predict performance is not random. This contention is based on (a) the consistency of half of the pictures across criteria, and (b) the finding that the positive and negative pictures differ reliably in the types of imagery they

elicit. The interpretation of this fact has been postponed until now.

The pictures characterized by negative affect and historical themes are, for the most part, among our strongest pictures. It seems improbable that one could support the argument that some types of content are more valid than others. It is our contention that the reasons for the failure of imagery types 3 and 4 to predict positively to performance is that they were strongly suggested by the pictures. Stories of Pasteur, Guttenberg, and Franklin represent cognitive recognition of the picture cues in pictures 66 and 24. Picture 34, the Law Students, elicits imagery of types 1 and 2, long term achievement, but still gives negative to zero prediction of performance. Perhaps this is because the stories tell of students studying to be lawyers and looking ahead to success, when that is exactly what the picture shows!

This argument leads to a description of pictures yielding positively predicting scores which is not unlike that of Symonds. The best pictures are of medium strength, tended to be underestimated by raters, and yield themes not clearly dictated by the picture cues. From these considerations it is now clear than an adequate test of hypotheses such as we had here, concerning content and affect cues, requires that we deal with an adequate sample of positive predicting pictures of medium strength. These we did not have and rather than reject the hypotheses it seems fairer to say they did not receive a fair test.

Suggestions for Future Research

With this initial demonstration that raters can reliably

predict n Achievement scores we can now move from the strategy of using many pictures and begin to learn more about a few pictures. Instead of using many raters and having to combine their ratings across groups, it would now be possible to eliminate between groups variance in ratings by working with a sample of pictures small enough to permit every rater to rate each picture. Then it should be possible to learn under what conditions, relaxed, task, or involved, the scores from these pictures shows the greatest predictive capacity. If the utility of the rating task as a source of picture cue definition can be improved with the use of better rating techniques, we may well be on the way to a better understanding of the role of picture cues in projective measurement.

The finding that it is possible to have negative predicting pictures raises some interesting possibilities. Should subsequent research show that some pictures, or batteries, yield highly reliable scores which relate negatively to performance we may have entered into a new stage of research with the n Achievement motive. The finding that low performers get high n Achievement scores to certain types of pictures might lead to some interesting research in the role of defense mechanisms, or at the very least tell us more about or low scorers.

Finally there are two additional facets of this study which deserve mention. First, the addition of four or five new pictures to the present battery could give us enough pictures to permit the construction of several parallel forms of the n Achievement measure, an indispensible requirement for longitudinal research on the motive. Second, the use of the booklet adminstration of the n Achievement measure not only

provides additional control, but also permits the study of intra-test differences in the meaning of the n Achievement scores. Thus we might study whether scores made early in the testing are better estimates of motivation than those made after exposure to two or three pictures.

In the last analysis the success or failure of any scientific method of studying human behavior rests on the degree of precision and reliability the method can attain. The success of the hypotheses for the use of ratings is taken to mean that a start has been made on a difficult problem in the use of projective materials. The failure of the hypotheses about picture differences is taken to mean that the problem's requirements of precise control have not yet been met. It is from such combinations of success and failure that new facts and new problems emerge.

Summary

The general purpose of this study was to determine more accurately the role of the picture cues in the projective measurement of the Achievement motive. First it was proposed to use ratings of the pictures on the scoring categories of the n Achievement system as an independent definition of the picture cues. Then these rated values were compared with the obtained n Achievement scores in an effort to see how well cognitive estimates of the pictures would predict motivational data.

One hundred and fifty college students rated an average of four pictures each from a total sample of sixty pictures. Each picture was rated by an average of ten persons, both male and female, and from these ratings median values were obtained for each picture on each of

the respective scoring categories. These medians were then combined into one distribution and rank order positions then assigned to the pictures. Eighteen pictures were then selected to provide three sets of six pictures. These sets were made parallel in ratings of strength, affect, and content.

The three sets of pictures were then made part of a two hour testing session in which a total of 73 college males performed on a Zeigarnik task, wrote stories to the pictures, recalled the Zeigarnik task items, then took speed tests in Addition and in Word Coding. All three tasks were known to have provided performance scores which in past studies showed a positive relationship to the n Achievement score.

Scores from the stories written in this session were then used to compare the picture scores with the ratings of the pictures obtained previously. A Latin Square presentation of the pictures made it possible to treat differences in n Achievement score distributions between pictures as due to pictures per se.

It was found that the raters could reliably predict both the achievement-related affect scores, and the likelihood of scorable achievement themes. However they could not anticipate the details of the achievement stories.

In the second part of the study an effort was made to compare pictures in terms of the discriminability and the directional predictiveness of their scores. It was found that the correlations of picture scores with performance scores on Additions and Word Coding yielded distributions symmetrical about zero. Pictures closest to the theoretically most discriminating position of eliciting stories from 50 per cent

of the subjects were found to yield scores with the highest correlations with performance, regardless of direction of relationship of score to performance.

Efforts to show that picture differences in direction of prediction were a function of content and affect cues were unsuccessful. When pictures were given definitions in terms of ratings of strength or affect these same hypotheses about picture differences again failed.

The failure of individual pictures to predict performance more reliably was attributed to an unforeseeable combination of circumstances which tended to depress differences in performance as a function of motivation. It was argued that the distribution of score correlations with performance was not random however, on the grounds that the relative predictive capacity of the pictures tended to remain stable for the three criteria, and that negative and positive pictures differed in the types of imagery they elicited. The failure of the hypotheses about the role of picture cues in effecting the directional prediction of n Achievement scores was taken to mean that the operations designed to test them had not proved adequate.

The results of this study are taken to indicate that the development of an operation to give picture cues a definition independent of the scores they elicit is feasible. The finding that many pictures gave scores negatively related to performance raises many interesting questions about the factors which determine the n Achievement score. A discussion of these factors and their interaction is presented along with suggestions for future research on this problem.

APPENDIX A

RATING QUESTIONNAIRE AND CRITERION MEASURES

Instructions.

You will be shown a number of pictures and asked to examine them carefully. One or more persons will be portrayed in each picture. You are to decide what situation the picture portrays, then you will be asked to judge these pictures in a variety of ways.

Two methods will be used. One will be the simple rank order method in which you will be asked to rank the statements or pictures on some attribute. The other method will be the pick 5 method in which you are asked to pick the five pictures or statements that most clearly show some attribute, and the five that show the attribute least. The first five are designated by a "1," the second five by a "2" and the last five by a "3." Then after that is done you will go back and rank all of them, ordering the first five from 1 to 5, the second five from 6 to 10, and the last five from 11 to 15.

The choices would appear as below.

Group	Rank
2	6
1	2
1	1
3	11
2	8
3	13
1	3
2	7
1	5
3	14
1	4
2	10
3	12
3	15
2	9

Any changes in instructions will be given before the new task. Work at a normal rate, and do not ponder too long over any one choice.

Task 1. Read the definitions of the situations given below, and then, drawing on your knowledge of how people behave and feel in situations similar to those portrayed in the pictures, you are to:

1. First use the "pick 5" method.
2. Rank the pictures as described in the instructions.

A rank of 1 is "most appropriate," a rank of 15 is "least appropriate." Read the complete list of situations before you begin.

Picture Group	Rank	The situation portrayed deals with:
____	____	influence, interest in controlling the opinions or behavior of others.
____	____	attention, interest in securing the attention, praise or interest of others.
____	____	authority, interest in legal or social control over actions of others.
____	____	avoiding rejection, interest in avoiding social rejection by others.
____	____	achievement, interest in doing well on a unique task of personal importance, or on a long-term career basis.
____	____	physical safety, interest in restoration or maintenance of physical well being.
____	____	guilt feelings, interest in reducing guilt over wrong doing.
____	____	friendship, interest in the restoration of maintenance of warm, interpersonal relations.
____	____	knowledge, interest in knowing about oneself and the world in which one lives.
____	____	sexual gratification, interest in being loved and esteemed by one of the opposite sex.
____	____	avoiding failure, interest in avoiding poor performance, either on a unique task or in a long-term career.
____	____	relaxation, interest in rest and relaxation from one's duties.
____	____	submission to authority, interest in submitting to and complying with the wishes of someone in authority.
____	____	reputation, interest in the preservation or restoration of one's reputation.
____	____	health, interest in the maintenance and restoration of good health.

Task 2.

Still considering picture ____ look at it as portraying

achievement, interest in doing well, in a unique task of personal importance or in a long term career.

Now rank the behaviors and experiences listed below for likelihood of occurrence in the pictured situation. If, in other words, we assume the chief interest to be achievement which of the following descriptions of behavior and feelings is most appropriate. Let a rank of "1" be most likely, and a rank of "10" be least likely. Read the entire list before you begin.

Rank

_____ wanting to achieve
_____ striving successfully to achieve
_____ striving unsuccessfully to achieve
_____ being helped to achieve
_____ anticipating success
_____ anticipating frustration
_____ facing environmental obstacles to achievement
_____ facing personal shortcomings
_____ experiencing the satisfaction accompanying success
_____ experiencing the dissatisfaction of failure

Task 3.

Still regarding the picture as portraying an achievement situation now consider the three emotional possibilities listed below. Rank the picture for achievement-related emotional tone, using a rank of "1" for most appropriate, and a rank of "3" for least appropriate.

_____ success, happiness, satisfaction, elation, optimism
_____ neutral
_____ unsuccessful, unhappy, dissatisfaction, dejected

Task 3.

Return once more to the pick 5 procedure. However this time your task is to consider ten pictures, compare them, and first group them into five most appropriate and five least appropriate, then rank them from 1 to 10 on their appropriateness to a situation which portrays:

achievement, i.e., interest in excellence of performance on a unique task or on a long-term career.

Simply use the pick and rank method on the ten pictures compared for appropriateness to the achievement situation.

Group Pic. No.	Rank Group	Rank
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____

II. Still considering the pictures for appropriateness to achievement now use the pick and rank method to indicate the emotional tone of the pictures. Rank 1 being most successful and happy, rank 10 being least.

Group Pic. No.	Rank Group	Rank
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____
____	____	____

Additions Task Instructions

This task consists of ten sheets of simple addition problems. It is your task to add the columns as quickly as possible. You will be given one minute per sheet. When I say "Begin," start working. At the end of each minute I will tell you to go on to the next page. If you finish before the time is up wait until the signal to turn is given. If you have not finished proceed to the next page when the signal is given.

Are there any questions?

49	23	69	44	19	55	73	11	23	83	10	12
91	80	45	10	49	72	42	62	52	17	37	54
45	33	98	48	12	85	48	13	37	49	99	42

95	83	91	26	49	33	36	97	63	64	59	80
89	99	43	31	36	31	91	37	38	18	26	93
29	59	62	47	47	12	86	75	24	81	89	54

29	66	40	62	19	92	19	20	30	76	16	46
40	98	99	17	54	30	73	21	28	37	14	97
56	63	80	90	60	34	83	73	83	84	98	83

Word Coding Test

This is also a speed test. You are given a list of forty four-letter words to each sheet. Above the list the columns of letters have each been given a number. Beside each word you will find the same numbers in scrambled order. It is your task to write the word in scrambled fashion according to the code. Here is an example:

1 2 3 4

B A L L 2 3 1 4 A L B L

Let me repeat. It is your task to code the words as quickly as possible. You will be given two and one half minutes to do each list of forty words. At the end of the time I will tell you to go on to the next page. Wait for the signal if you are finished, and move on if you are not. Any questions?

		1234								
1.	AMID	3	2	1	4	_	_	_	_	1.
2.	COZY	1	2	4	3	_	_	_	_	2.
3.	SAKE	2	1	3	4	_	_	_	_	3.
4.	PAWN	1	3	2	4	_	_	_	_	4.
5.	ONCE	4	3	1	2	_	_	_	_	5.
6.	FORK	2	1	4	3	_	_	_	_	6.
7.	OPAL	3	1	2	4	_	_	_	_	7.
8.	PROW	2	1	4	3	_	_	_	_	8.
9.	REIN	1	3	4	2	_	_	_	_	9.
10.	SIGH	2	4	1	3	_	_	_	_	10.
11.	TICK	1	3	2	4	_	_	_	_	11.
12.	HILT	2	4	3	1	_	_	_	_	12.
13.	THUG	1	4	2	3	_	_	_	_	13.
14.	MAST	2	4	1	3	_	_	_	_	14.
15.	PIPE	3	1	2	4	_	_	_	_	15.
16.	MILK	2	1	4	3	_	_	_	_	16.
17.	PINK	3	1	4	2	_	_	_	_	17.
18.	CLAD	2	1	4	3	_	_	_	_	18.
19.	QUAD	1	4	3	2	_	_	_	_	19.
20.	GIST	2	1	4	3	_	_	_	_	20.
21.	SLAB	1	3	4	2	_	_	_	_	21.
22.	HUES	3	2	1	4	_	_	_	_	22.
23.	WRAP	1	4	3	2	_	_	_	_	23.
24.	FARM	2	3	1	4	_	_	_	_	24.
25.	DOCK	1	4	3	2	_	_	_	_	25.
26.	PIES	4	1	3	2	_	_	_	_	26.
27.	LOAM	1	4	2	3	_	_	_	_	27.
28.	HOCK	2	4	3	1	_	_	_	_	28.
29.	PLAT	1	4	2	3	_	_	_	_	29.
30.	AIMS	4	1	3	2	_	_	_	_	30.
31.	BOLT	2	1	4	3	_	_	_	_	31.
32.	RAMP	3	1	2	4	_	_	_	_	32.
33.	DATA	2	4	3	1	_	_	_	_	33.
34.	RINK	1	4	2	3	_	_	_	_	34.
35.	NORM	2	1	4	3	_	_	_	_	35.
36.	YOLK	3	1	4	2	_	_	_	_	36.
37.	VERB	3	4	2	1	_	_	_	_	37.
38.	DIRE	2	1	4	3	_	_	_	_	38.
39.	CUES	1	2	4	3	_	_	_	_	39.
40.	DOES	4	2	1	3	_	_	_	_	40.

APPENDIX B

TABULATION AND TREATMENT OF DATA

TABLE 5

MEDIAN RANK VALUES FOR PICTURES
ON RATED CATEGORIES

Picture	Categories							
	Ai	Sign	N	I+	I-	Ga+	G+	G-
14	1.20	7.25	1.60	7.50	8.00	5.00	8.62	10.00
17	2.50	7.83	2.00	4.75	8.00	4.50	6.75	6.50
7	2.28	6.40	1.08	2.40	8.25	3.00	4.30	9.10
12	4.00	6.00	2.00	6.50	7.67	5.00	7.50	9.60
21	4.00	- *	2.25	4.00	8.10	3.83	7.00	9.83
19	4.50	-	6.00	2.50	8.17	4.50	6.75	6.50
4	6.50	7.25	2.20	8.50	6.75	6.00	8.30	7.00
32	1.21	7.30	1.90	2.50	9.00	2.90	4.80	9.70
64	3.00	1.10	1.80	3.30	8.62	1.67	3.60	9.80
54	1.50	-	-	-	-	-	-	-
69	3.30	-	-	-	-	-	-	-
50	2.25	-	-	-	-	-	-	-
53	2.00	4.20	1.40	2.40	9.00	2.90	2.30	9.80
41	6.75	-	1.50	2.25	8.83	3.50	4.50	6.50
2	1.75	7.67	1.25	6.75	5.50	6.20	8.50	7.50
66	1.37	7.64	1.25	4.50	7.25	2.75	5.00	9.67
34	1.17	8.10	1.10	3.16	8.50	4.50	7.17	9.50
24	1.41	7.00	2.50	1.50	8.17	3.25	5.75	9.00

Possible Median Ranges:

Ai	1 to 15
Sign	1 to 10
Others	1 to 10

*Missing values were not used in the analysis due to lack of comparable obtained scores.

TABLE 6

RANK ORDERS OF PICTURES USED IN CORRELATIONAL ANALYSIS (OBTAINED SCORES)

	Strength			
% Imagery	Mid vs Hi-Lo (Obt. Mid pt.)	Mid vs Hi-Lo (50% pt.)	Sign + total	Imagery 3 + 4 total
66	64	7 -	12	2
32	17	14 -	64	24
34	12 -	53	53	19
24	4 -	24	7	66
7 -*	2	4	4	41
14 -	21 -	64	24	21
53	41 -	34	34	4
4	53	17	32	64
64	14 -	12	12	14
17	7 -	2	14	34
12	19	21 -	66	7
2	24	41 -	17	53
21 -	34	32	2	32
41 -	32	19		12
19	66	66		17
50 -	54 -	54 -		
54 -	50 -	50 -		
69	69	69		

*Adjacent dashes indicate ties.

TABLE 7

VALUES OF TAU_b FOR n ACHIEVEMENT SCORES OF EACH PICTURE WITH PERFORMANCE SCORES ON ADDITIONS AND WORD CODING TASKS: ZEIGARNIK INDEX VALUES FOR EACH PICTURE

	Word Coding		Additions		Zeigarnik	
Picture	Rank	tau_b	Rank	tau_b	Rank	Index
32	1	+.17	17.5	-.20	7	- .09
12	2	+.16	4	+.08	11	- .63
14	3	+.13	1.5	+.21	6	- .01
7	4	+.08	13	-.15	3	+ .61
17	5	+.06	1.5	+.21	8	- .13
4	6.5	+.04	15	-.17	2	+ .70
21		+.04	3	+.16	13	- .79
64	9.5	.00	14	-.16	1	+1.00
54	9.5	.00	9	.00		
69	9.5	.00	9	.00		
50	9.5	.00	9	.00		
53	12	-.03	7	+.01	4	+ .51
41	13	-.04	6	+.06	15	-1.55
2	14	-.11	12	-.03	5	+ .50
66	15	-.12	16	-.19	12	- .78
24	16	-.17	17.5	-.20	14	-1.22
34	17	-.22	11	-.02	10	- .59
19	18	-.24	5	+.07	9	- .50

TABLE 8

DISTRIBUTION OF w COEFFICIENT
(Samples of ten rankings taken from a total of twenty rankings of ten pictures)

w N = 10 m = 10	Parameters
.59 .57 .56 .56 .51 .51 .45 .41 .40 .40 .40	Mean .497 Sigma .074 Limits, + or - 3 Sigmas from Mean, .719 to .275 w value for ten rankings of ten elements must be .21 for conversion to a X^2 significant at the $P < .01$ level.

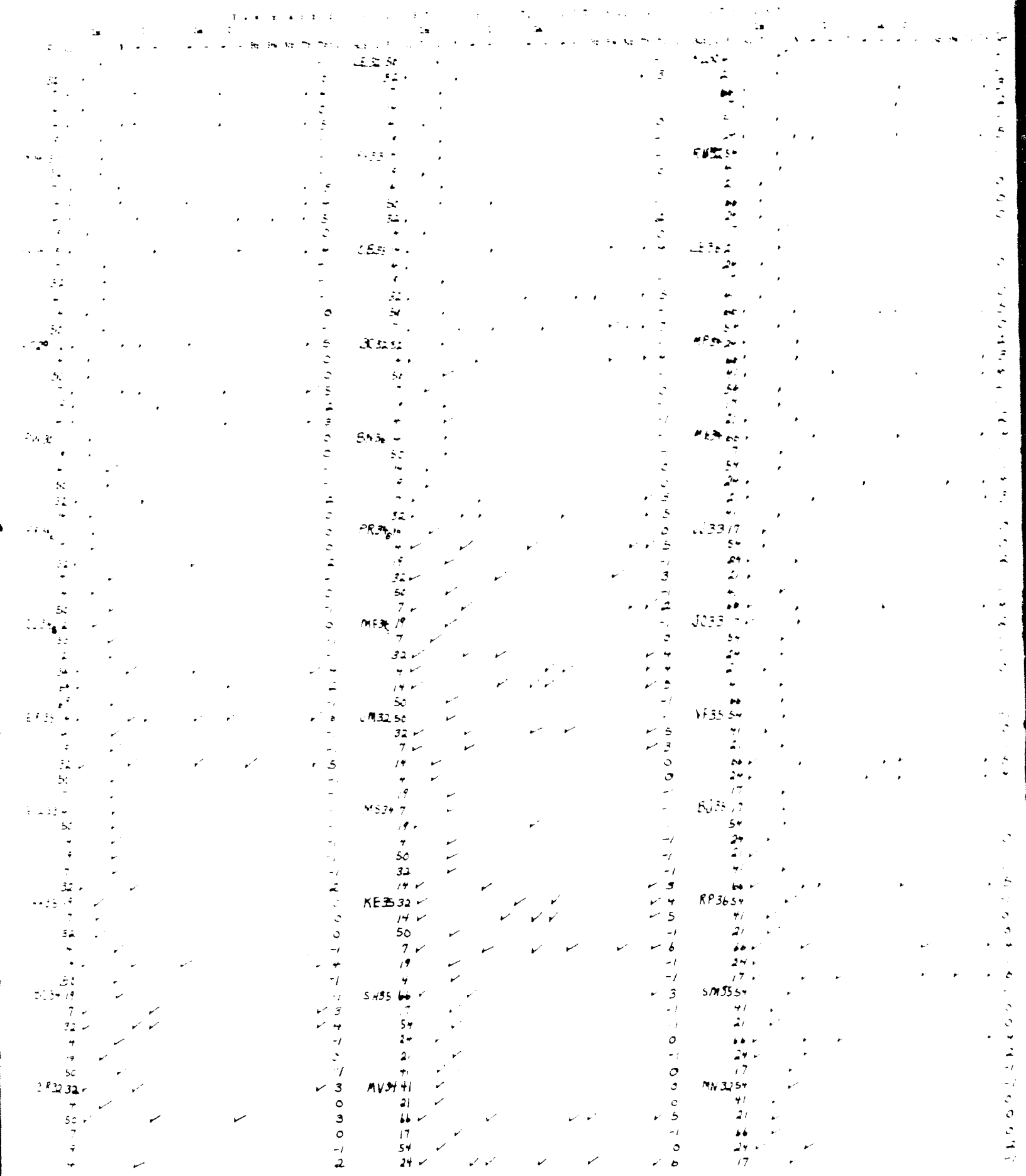
APPENDIX C

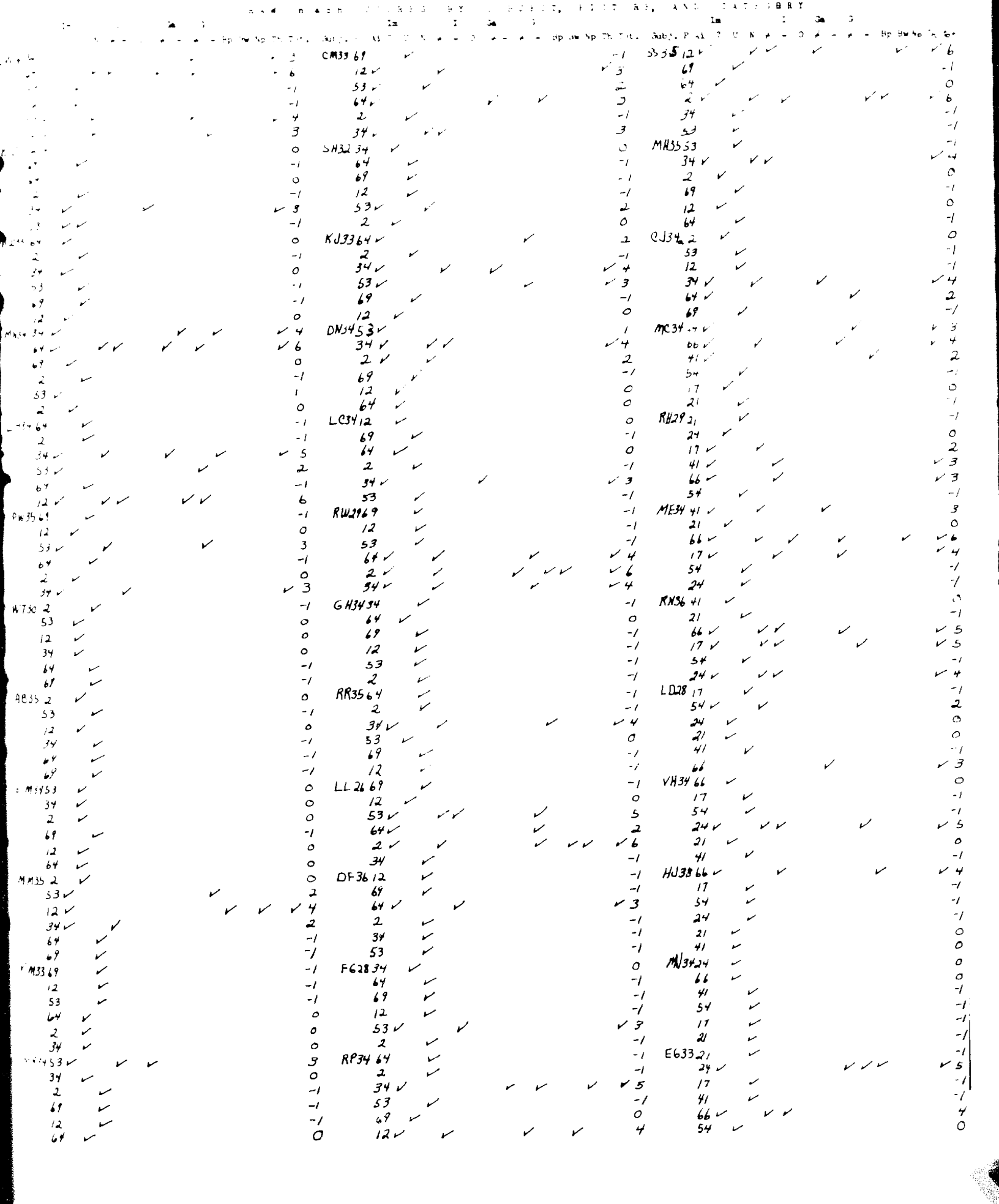
RAW DATA OF CRITERION TASKS AND n ACHIEVEMENT SCORING

RAW PERFORMANCE SCORES ON ZEIGARNIK, ADDITIONS, AND WORD CODING TASKS.

	Items		Recalled			Transformed* Recall Scores		Additions Task Trials										Word Coding task Trials.							
	Com.	Incom.	Com.	Inc.	Tot.	Com.	Incom.	1	2	3	4	5	6	7	8	9	10	Mean	1	2	3	4	5	6	7	
MC35	7	13	2	5	7	3.30	4.08	10	7	11	9	8	6	9	9	7	7	8.3	16	21	15	14	17	16	16	16.4
EM34	11	9	5	4	9	4.36	4.49	11	11	14	11	11	10	13	12	12	13	11.8	22	21	23	22	22	22	20	21.7
[illegible]	7	11	4	6	10	4.30	5.24	16	19	19	17	15	16	18	17	19	16	17.2	25	25	24	24	21	21	20	22.7
[illegible]	12	8	9	4	13	6.35	4.90	14	11	12	11	13	11	10	11	13	11	11.7	16	19	19	16	15	17	24	18.0
[illegible]	10	10	6	6	12	5.32	5.58	17	16	18	15	18	16	17	18	15	16	16.6	5	15	16	16	15	17	16	14.3
[illegible]	10	10	7	4	11	5.99	4.22	15	14	14	12	12	8	12	13	13	14	12.7	12	20	25	26	23	26	25	22.4
[illegible]	11	9	3	6	9	3.16	6.87	12	15	15	16	18	14	15	19	13	19	15.6	24	24	23	25	24	23	22	23.6
EK35	9	11	6	3	9	5.78	3.29	11	11	13	11	11	12	10	11	11	10	11.1	14	17	19	17	17	18	16	16.9
EW33	9	11	5	4	9	5.06	3.94	12	13	14	12	14	11	13	9	11	10	11.9	19	18	18	20	18	20	17	18.6
HH35	8	12	3	8	11	4.49	5.27	5	5	6	4	7	5	7	7	—	6	5.2	13	14	13	13	13	14	12	13.1
BC34	9	11	5	6	11	5.65	4.53	13	12	15	13	14	11	12	12	11	12	12.5	15	20	25	23	25	25	26	22.7
DR32	4	11	3	9	12	4.17	6.32	14	12	15	16	13	13	15	14	16	11	13.9	21	21	24	21	25	21	25	22.6
LE32	9	11	5	6	11	5.64	4.53	12	10	13	11	13	9	12	11	12	12	11.5	18	20	18	20	17	19	17	18.4
PF33	11	9	4	6	10	4.37	5.27	17	20	20	18	19	19	20	19	20	19	19.1	20	27	31	32	26	29	25	28.3
CB35	9	11	4	8	12	4.88	5.66	12	10	16	13	14	14	14	16	13	12	13.4	18	25	29	27	25	23	21	24.0
BC32	11	9	3	8	11	3.76	6.92	20	23	29	20	23	20	21	21	21	22	21.4	30	30	33	33	33	33	27	31.9
EN36	9	11	3	7	10	4.16	5.08	11	10	10	11	12	11	9	9	12	9	10.4	17	17	19	20	19	20	17	18.4
PR34	10	10	3	7	10	3.96	5.96	13	12	13	10	13	12	13	13	14	13	12.6	21	21	25	28	25	26	27	24.7
MF35	11	9	8	6	14	6.76	5.27	17	23	22	19	21	21	24	20	23	17	20.7	13	14	17	26	35	35	37	25.3
JM32	9	11	5	9	14	5.64	6.32	10	12	15	12	16	14	15	12	14	14	13.4	31	27	28	32	37	39	34	32.6
MS34	9	11	4	3	7	5.12	2.78	16	18	14	15	14	13	17	16	15	14	15.2	15	16	12	16	16	18	11	14.9
KE35	10	10	6	7	13	5.89	5.46	16	15	15	14	17	14	15	17	16	15	15.4	28	27	29	28	26	31	24	27.6
MV34	10	10	5	6	11	5.26	4.83	13	12	14	14	15	12	14	13	16	13	13.6	21	24	23	21	20	25	23	22.4
[illegible]	10	10	5	4	9	5.26	3.62	16	16	16	15	17	15	16	15	14	16	15.6	30	28	27	30	29	30	31	29.5
RW32	12	8	6	8	14	5.26	8.96	19	16	17	17	18	15	16	18	17	17	17.0	29	36	40	40	40	40	40	37.8
LE36	9	11	8	7	15	8.87	5.08	16	18	19	15	20	15	17	18	17	18	17.3	33	34	40	40	39	40	40	38.0
MP34	10	10	5	5	10	5.26	4.22	13	14	15	13	15	13	15	16	15	16	14.5	13	21	21	18	23	29	27	21.7
MW34	10	10	7	5	12	6.90	4.22	12	13	15	14	16	14	13	13	15	14	13.9	22	24	23	25	24	25	21	23.4
JJ33	10	10	6	5	11	5.89	4.22	13	12	17	13	13	11	13	12	13	10	12.7	18	14	19	19	20	16	15	17.3
JC33	11	9	3	6	9	3.76	5.27	20	20	21	17	18	18	18	18	18	18	18.6	27	34	31	39	36	36	32	33.6
VF35	9	11	3	6	9	4.17	4.53	10	10	14	7	11	9	12	9	11	10	10.3	14	13	15	15	17	16	13	14.7
BJ35	9	11	3	6	9	4.05	4.53	9	8	10	7	12	7	9	6	8	9	8.5	21	26	27	29	31	25	28	26.7
RP36	12	8	6	4	10	4.68	4.90	No other tasks performed																		
SM35	9	11	5	4	9	5.06	3.29	14	12	15	14	14	12	10	15	13	14	13.3	17	17	18	18	21	21	17	18.4
MN31	10	10	7	4	11	5.99	4.22	18	15	15	15	15	14	17	17	15	14	15.5	24	24	23	23	23	23	23	23.3
MC34	7	13	3	8	11	4.23	5.64	7	6	7	6	7	6	9	9	8	7	7.2	12	9	11	9	8	13	15	11.0
RH29	7	11	3	7	10	3.58	5.85	15	17	15	14	12	13	14	13	15	14	14.2	26	28	32	31	36	32	28	30.4
ME34	11	9	6	2	8	6.00	2.89	15	19	20	15	21	17	18	19	20	20	18.4	21	26	27	28	26	28	23	25.6
RN36	12	8	8	6	14	5.78	6.66	14	12	14	11	12	12	—	16	10	14	11.5	21	20	22	22	20	22	20	21.0
LD28	9	11	7	7	14	6.58	5.85	19	21	21	21	20	20	—	24	20	18	18.4	23	23	22	18	15	21	38	22.8
VH34	12	8	9	5	14	6.35	5.85	20	18	18	17	16	17	19	18	19	17	17.9	31	12	27	25	25	24	27	24.9
HJ33	10	10	3	7	10	3.37	6.28	15	16	17	14	16	16	16	16	16	14	15.6	21	24	23	26	27	26	26	24.7
MJ34	11	9	5	3	8	4.36	3.73	10	11	8	10	11	10	11	10	9	10	10.0	18	17	18	21	21	24	17	19.4
EG33	12	8	7	5	12	5.19	5.78	12	9	10	9	12	9	12	11	9	9	10.2	26	24	27	27	24	28	26	26.0
SH35	9	11	5	4	9	5.06	3.94	10	8	8	7	10	8	9	11	8	10	8.9	15	16	14	16	16	17	12	15.1
LW36	12	8	8	6	14	5.78	6.66	18	21	21	18	18	19	15	18	21	18	18.7	23	24	26	27	25	27	24	25.1
KL35	12	8	6	5	11	4.68	5.78	17	18	18	17	18	16	19	18	19	18	17.8	29	28	32	30	31	30	28	29.7
RW35	13	7	8	5	13	6.05	6.35	20	18	21	21	23	20	17	22	24	19	20.5	21	24	26	28	22	23	18	23.1
MW34	10	10	6	5	11	5.32	4.90	10	10	10	9	9	10	11	11	10	11	10.1	16	15	18	17	18	16	14	16.3
LH34	8	12	6	5	11	6.35	4.35	13	11	15	11	16	13	15	16	14	14	13.8	23	23	23	25	24	27	24	24.1
RW35	9	11	4	6	10	4.36	5.24	13	14	12	13	13	11	14	12	14	12	12.8	22	21	22	21	21	23	21	21.6
WT30	9	11	6	3	9	5.78	3.29	14	12	13	10	12	11	13	13	11	12	12.1	20	36	28	30	31	33	32	29.1
RB35	9	11	5	5	10	5.06	4.56	11	10	10	9	12	7	11	7	7	7	9.1	11	14	13	4	13	14	11	11.4
EM34	10	10	6	4	10	5.32	4.22	11	9	12	9	12	7	10	12	10	8	10.0	25	32	28	36	33	39	31	32.0
MM35	13	7	5	4	9	3.91	5.37	16	19	18	14	16	14	16	14	15	15	15.7	38	37	39	40	40	40	40	39.1
JM33	14	6	6	4	10	4.23	6.07	14	15	15	16	14	13	15	12	15	14	14.3	25	28	25	24	29	26	27	25.6
[illegible]	8	12	2	8	10	3.02	6.07	12	12	14	11	12	12	14	12	13	12	13.4	16	20	21	20	21	21	20	19.8
RM33	12	8	9	4	13	6.35	4.90	11	10	11	9	12	9	10	11	12	9	10.4	23	24	24	23	24	24	21	23.3
SH32	13	7	6	4	10	5.61	4.65	13	13	14	14	15	14	15	15	16	15	19.4	23	21	28	28	33	30	25	28.0
KJ33	9	11	2	7	9	3.39	5.08	13	11	16	13	13	11	12	13	11	12	12.5	26	38	34	28	29	35	38	32.5
DN34	12	8	7	5	12	5.76	5.02	11	9	12	10	12	11	12	13	12	11	11.2	22	20	20	18	22	21	17	20.0
LC34	11	9	6	4	10	5.57	3.86	14	17	17	15	15	15	15	14	16	16	15.4	26	33	39	38	37	40	39	36.0
RW29	10	10	3	6	9	3.96	4.83	10	11	13	10	10	9	10	13	13	10	10.9	21	21	23	22	24	22	21	22.4
GH34	11	9	5	5	10	4.94	4.59	9	12	14	13	15	14	15	16	14	13	13.5	23	23	23	21	26	36	37	21.0
RR35	10	10	6	7	13	5.89	5.46	21	21	22	20	23	18	21	20	21	19	20.6	25	22	26	26	23	26	23	24.4
CL26	11	9	4	7	11	4.37	6.02	12	11	14	12	16	11	14	11	11	11	12.3	17	18	20	22	25	21	20	20.4
DF36	8	12	2	9	11	3.61	5.80	9	7	8	10	11	10	13	11	11	12	10.2	35	28	31	34	33	37	35	38.9
FG28	10	10	5	5	10	5.26	4.22	7	12	14	12	13	11	12	12	11	12	11.6	11	15	14	14	15	13	17	14.1
RP34	9	11	4	5	9	4.88	3.93	8	8	12	9	9	7	11	10	10	9	9.3	14	16	14	16	16	16	12	14.9
SS35	11	9	4	6	10	4.37	5.27	19	18	18	17	18	17	18	17	18	17	17.9	24	25	29	37	34	39	39	32.4
MH35	9	11	3	4	7	4.17	3.37	13	12	12	15	15	13	14	13	15	14	13.6	26	24	25	26	27	26	26	25.6
CJ34	11	9	6	5	11	5.57	4.59	17	19	19	17	19	16	—	20	16	18	16.1	23	24	24	21	23	23	23	23.4

* = score of Arc Sin value of percentage of completed or imcompleted items recalled.





BIBLIOGRAPHY

1. Allport, G. The trend in motivational theory. Amer. J. Orthopsychiat., 1953, 23, 107-119.

2. Bellak, L. The Thematic Apperception Test in clinical use. in Abt, L. E., and Bellak, L., Projective psychology., New York: Alfred Knopf, 1950.

3. Boring, E. Visual perception as invariance. Psych. Rev., 1952, 59, 141-147.

4. Bruner, J. Personality dynamics and the process of perceiving. in Blake, R. E., and Ramsey, G. V., Perception, an approach to personality. New York: The Ronald Press Co., 1951.

5. Brunswick, E. Systematic and representative design of psychological experiments. Univ. Calif. Press., 1949.

6. Burdick, H. Personal communication, 1954.

7. Chapman, D. Personal communication, 1952.

8. Coombs, C. A theory of psychological scaling. Engineering Res. Bull. 34., May 1952. ERI, Univ. Mich.

9. Eron, L. D. A normative study of the TAT. Psych. Mono., 1950, 2, 64.

10. Eron, L. D., Terry, D., and Callahan, R. The use of rating scales for emotional tone of TAT stories. J. consult.Psychol., 1950, 14, 473-478

11. Eron, L. D. Frequencies of themes and identifications in the stories of patients and non-hospitalized college students. J. consult. Psychol., 1948, 12, 387-395.

12. Fisher, R. A. The design of experiments. New York: Hafner Pub. Co., 1951.

13. Henry, W. E. The TAT Technique in the study of culture-personality relations. Genetic Psych. Mono., 1947, 35, 3-135.

14. Holt, R. The TAT Newsletter. in J. Proj. Tech., 1950, 14, 80-100.

15. Jacobs, B. A method for investigating the cue characteristics of pictures used in projective measurement of motives. Honors thesis. Univ. Mich., 1953.

16. Kendall, M. G. Rank correlation methods. London: Chas. Griffin and Co., Ltd., 1948.

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

18. Morgan, C. D., and Murray, H. A. A method for investigating fantasies. Arch. Neur. and Psychiat., 1935, 34, 289-306.

19. Moulton, R. Personal communication, 1954.

20. Murray, H. A. The effect of fear upon estimates of maliciousness of other personalities. J. Psychol., 1933, 4, 310-339.

21. Murray, H. A. Explorations in personality. New York: Oxford Univ. Press., 1938.

22. Postman, L. J. The experimental analysis of motivational factors in perception., in Current theory and research in motivation, a symposium. Univ. Nebr. Press, 1953, 59-108.

23. Postman, L. J. Toward a general theory of cognition., in Social psychology at the crossroads., ed. by Rohrer, J. H., and Sherif, M. New York: Harper and Bros., 1951.

24. Rosenzweig, S., and Flemming, E. Apperception norms for the Thematic Apperception Test. II. An empirical investigation. J. Pers., 1949, 17, 483-503.

25. Symonds, P. M. Criteria for the selection of pictures for the investigation of adolescent fantasies. J. abn. soc. Psychol., 1939, 34, 271-274.

26. Thompson, C. E. The Thompson modification of the Thematic Apperception Test. Rorschach Res. Exch. and J. proj. Tech., 1949, 13, 469-478.

27. Weisskopf, E. A. An experimental study of the effect of brightness and ambiguity on projection in the Thematic Apperception Test. J. Psychol., 1950, 29, 407-416.

28. Weisskopf, E. A., and Dunlevy, G. P. Bodily similarity between subject and central figure in the TAT as an influence on projection. J. abn. soc. Psychol., 1952, 47, 441-445.

29. Weisskopf, E. A., Joelson, E. and Money, L. Facial similarity between subject and central figure in the TAT as an influence on projection. J. abn. soc. Psychol., 1953, 48, 341-344.

30. Wekstein, L. A preliminary outline for a fantasy projection technique as a clinical instrument. J. Psychol., 1945, 19, 341-346.

31. Wendt, H. W. Personal communication, 1954.

32. Wittenborn, J. R., and Eron, L. D. An application of drive theory to TAT responses. J. consult. Psychol., 1951, 15, 45-50.