Clinical Psychology Review, Vol. 1, pp. 319–336, 1981
Printed in the USA.

0272–7358/81/030319–18$02.00/0

THE THEMATIC APPERCEPTION TEST: A REVIEW

Julia R. Vane

Hofstra University

ABSTRACT. *The development of the Thematic Apperception Test and subsequent variations in the method of administering, scoring and interpreting the test, as well as standardization attempts by Eron, Murstein, Dana, Arnold, and McClelland and Atkinson are covered. At present only about two-thirds of the doctoral programs in clinical, counseling, and school psychology are teaching the TAT, and the majority of these programs are teaching nonstandardized techniques of interpretation. If the TAT is to meet the criteria of adequate reliability and validity, a large-scale effort is needed to develop a standardized test that may be easily scored and interpreted. This would be worthwhile, because research suggests that the TAT method elicits material regarding attitudes and motivation that is not adequately elicited by other methods.*

The Thematic Apperception Test is a controversial instrument which has generated a great deal of research and many different systems of analysis and scoring. It has its strong critics and strong adherents, but both would agree that the test is a clinician's delight and a statistician's nightmare.

The Thematic Apperception Test was introduced originally in an article published in 1935 by Christiana Morgan and Henry Murray. Murray and colleagues at the Harvard Psychological Clinic published further material regarding the test in 1938 and 1943. The purpose of the test was to enable the examiner to analyze the unconscious fantasies of the test taker by means of evaluating the stories told to a set of pictures. It was thought that when individuals were confronted with ambiguous social situations, as shown in the pictures, and required to make up stories about them, the personalities of the story tellers would be revealed in the process. Murray believed that the value of the Thematic Apperception Test, or TAT as it came to be called, lay in its capacity to reveal things that the individual is unwilling to tell or unable to tell because the material is unconscious.

The TAT was endorsed enthusiastically by many psychologists almost as soon as it appeared. Articles by those interested in its potential as a clinical instrument and those who saw it as a possible research tool began to appear in 1941, and by

Requests for reprints may be addressed to Julia R. Vane, Psychology Dept., Hofstra University, Hempstead, N.Y. 11550.

1950 over 100 articles and several books had been published regarding it. These articles ranged from new methods of scoring and administration to reports of research using the TAT to study language variations, aspects of child development, social attitudes, delinquency, and personality abnormalities.

THE ORIGINAL TEST

Murray worked on several revisions of the test and a final set of 31 pictures was developed. Eleven were designed for all sexes and ages, seven for girls and women only, seven for boys and men only, one for women only, one for men only, two for boys only, and two for girls only. In addition, there was one blank card to be given to men and women. The cards were to be administered in a particular order in two sessions, one week apart, with 10 cards given at each session, for a total of 20 cards. Each card was to be given to the examinee individually, and the examinee was asked to make up a dramatic story for each picture which would include the thoughts and feelings of the characters, and to tell what led up to the story and how it ended. The stories usually are recorded by having the examiner write what the story teller says as the stories are told. Tape recordings and stenographic transcripts have been used as well.

Murray's method of evaluating the material obtained from the TAT was based on psychoanalytic theory, and he believed that one of the major uses of the test would be to hasten the process of therapy. He assumed that evaluation of the stories would reveal what would take many months to uncover in psychoanalysis. No overall quantitative or numerical scores were proposed. Seven categories were to be evaluated in each story: namely, the characteristics of the hero or main character; needs or drives evident in the stories; the press or force of the environment; the inner states and emotions of the hero; the interests of the hero; the outcome of the stories and a general overall evaluation of the theme of the story derived from a combination of the need, press, and outcome. The hero in the story was considered to be the individual with whom the storyteller would be most closely identified. This method of evaluation was developed using the protocols of 50 Harvard undergraduates, and these were the normative data that were originally supplied (Murray, 1938).

Those psychologists who began to use the test as a clinical tool to assess personality and predict behavior from this assessment found that Murray's method of giving and scoring the test was time consuming, and as result shortened versions began to appear. Within a decade, almost no one was adhering to the recommended practice of giving 20 cards in two sessions, but instead anywhere from five to twelve cards were given in only one session. Sometimes, parts of Murray's assessment system were grafted onto other systems and totally different systems developed. The scope of these variations is attested to in a book published in 1951 by Schneidman. The book was based on an attempt to establish the reliability of the TAT as well as to compare it to another test, the MAPS. Experts in the field were asked to interpret the same protocol based on responses to eleven TAT cards, give the results, and explain their method of analysis. The criterion for selecting the experts was that they had published a method or technique for interpreting or analyzing the TAT. A number of experts agreed to cooperate with the study and to analyze the TAT "blind", (i.e., knowing only the sex, age, and marital status of the subject). This group included many psychologists who were and still are well known in psychology. They were Magda Arnold, Betty Aron, Leopold Bellak, Leonard Eron,

Reuben Fine, Arthur Hartman, Robert Holt, Walter Joel, Shirley Jessor, Seymour Klebanoff, Sheldon Korchin, Jose Lasaga, Julian Rotter, Helen Sargent, David Shapiro, Percival Symonds, and Ralph White. There were other psychologists who had published material on the TAT but either they were not available or did not choose to involve themselves. These included John Atkinson, David McClelland, Morris Stein, and Sylvan Tompkins.

The analyses contributed to the book by the experts are impressive, not only for the diversity of thought, but also for the amount of work that had been invested in the development of a variety of systems. Several psychologists had quantitative scoring systems; others had complicated notations; still others had attempted to develop normative data.

VARIATIONS

Schneidman (1951) reviewed the evaluations made by these experts and reduced them to 802 separate items of personality description; each of these 802 items was subsumed under one of 38 categories. Despite this large number of items, the author stated that the amount of agreement of the TAT interpretations with one another and with the behavioral data regarding the subject "seems to be remarkable," (Schneidman, 1951, p.307). He admitted, however, that this was a qualitative impression and how the TAT worked and how the clinical evaluations were arrived at was to be answered in the future. Unfortunately, the reader is left with no clear idea of which system might be the most desirable in terms of reliability or validity.

Murray himself seemed to be ambivalent about the results, for in commenting on the interpretation of the experts he stated, "those who laughed at probabilities and stretched inference to the limit succeeded in seizing the fruits of truth.... The fact that those who gave their intuitions freest play came closer the the mark can hardly be hailed as a prophetic sigh. For some of the psychologists who in this experiment held most conscientiously to the roots of a particular conceptual scheme may be contributing most, now and later, to the development of an instrument which will someday be unanimously accepted as the one which yields the highest percentage of valid diagnoses. Thus, in all likelihood, the more methodical tortoises will, as usual, be the first to arrive at the scientific goal—a reliable system instead of unreliable flashes of the imagination," (Schneidman, 1951, p. ix)

GENERAL USE

Despite Murray's belief that the TAT would be developed into a reliable test, and despite the fact that by 1971 nearly 1800 articles had been published relating to it (Reynolds & Sundberg, 1976), it remains today an impressionistic instrument in the eyes of many psychologists, and not a test at all in terms of standardization data, reliability, or validity. Its popularity as a clinical instrument has continued to be high. This is attested to by a number of studies that indicate that it is among one of the most frequently used tests by clinicians in the United States and other countries of the world (Brown & McGuire, 1976; Gonzales, 1977; Lubin, Wallis & Paine, 1971; Lapointe, 1974; Sundberg, 1961; Wade, Baker, Morton, & Baker, 1978). Not only is the TAT used as a clinical instrument, but it has been used in research in almost every country in Europe, in Russia, India, China, South Africa, and many countries of South America and Asia.

It would appear that with such extensive research and use, there would be

adequate data available to provide a substantial foundation for the TAT as a standardized test. Unfortunately, because of the variations in type of administration, scoring, and selective use of only some of the cards, much of the research is not comparable and cannot be used as a basis for demonstrating validity and reliability. There are studies, however, which suggest that when used in certain ways, the TAT reveals important aspects of personality, and that within these limits the instrument may meet many of the criteria of a good test.

STANDARDIZATION ATTEMPTS

In attempting to evaluate different studies, one of the questions that arises is: Can results obtained using group administration be compared to those obtained using individual administration? A number of authors have studied this problem (Eron & Ritter, 1951; Eron, Sultan, & Auld, 1955; Lindzey & Heinemann, 1955; Lindzey & Silverman, 1959; Sarason & Sarason, 1958; Terry, 1952). The results indicate that longer stories are told when the test is individually administered, but that there is little difference in the content of the stories despite the difference in the number of words. Lindzey and Silverman (1959) found no differences for variables such as achievement, affiliation, transcendence, number of themes, dominance, involvement, sex, quality, number of figures, and compliance with instructions. There were a larger number of rejections with group administration and more alternative themes and comments elicited under individual administration. In view of the small differences obtained between these two types of administration, comparison between studies using either of these methods seems reasonable.

ERON'S RESEARCH

Eron (1950), who was one of the original contributors to Schneidman's book, worked to provide objective psychometric procedures to make the TAT a more efficient clinical instrument. To this end he evaluated 3,000 TAT stories taken from 150 male subjects, divided into groups of 25 of similar age and IQ. Two groups were normal college students. There was one group each of non-hospitalized neurotics, hospitalized neurotics, schizophrenics, and general neuropsychiatric patients. All were given the 20 male TAT cards in two sittings with a ten minute interval between sessions. Standard directions were given, and responses were recorded by a trained examiner. Using the same method, Eron (1953) also collected TAT stories from 60 normal women and later from a group of Navy recruits using a group administration of 20 cards (Eron & Auld, 1954).

Eron evaluated the stories in a number of different ways, making comparisons between the normal and disturbed groups in terms of emotional tone, outcome, themes, and perceptual distortions. His results showed that many perceptual distortions originally thought to be characteristic of abnormality were found to be common among normals. For example, although it had been thought that sexual misidentification of characters in the TAT cards was a sign of sexual confusion, some of the figures, such as the so-called male in Card III was identified as a girl or woman by 40 to 60% of each group of normals and disturbed. Other perceptual distortions, originally thought to be characteristic of schizophrenics, were found to be common among normals as well, such as seeing the violin in Card I as a gun or typewriter, and adding a happy ending to a sad story. Another belief, namely

that individuals who made up sad stories were depressed, was negated by the finding that the great majority of stories told by normals and hospital groups were sad. Eron concluded that this resulted from the stimulus characteristics of the cards themselves.

From his analysis of the stories, Eron (1950) delineated 6,210 themes. Themes of violence, hostility, death, restriction, guilt, and frustration were among the most frequent and appeared in all groups, both normal and disturbed. The theme "pressure from parents" was number one in both groups of normals, nonhospitalized neurotics, and hospitalized neurotics; second in general neuropsychiatric patients and third in hospitalized psychotics. Eron also found that normals gave more symbolic stories, more abstract stories, and more stories which continued from one card to the next, and more stories that were designated as unreal. Although there were some intergroup differences with respect to themes, "the outstanding finding of this study seems to be the similarity in fantasy content among our various groups of subjects" (Eron, 1950, p.26). As the result of this work of Eron, most users of the TAT concluded that it was not a good instrument for differentiation of normals from those with psychiatric disorders.

Eron's list of normative themes were based upon the manifest content of the stories, and symbolic or covert significance was disregarded. The themes were classified according to equilibrium and disequilibrium, based upon the tension or adjustment of the characters in the story. These two categoreis were divided into interpersonal, intrapersonal, and impersonal, and then further subdivided for direction of the relation, such as parent, partner, peer, or sibling. A list of 161 categories of themes was developed. Tables were published (Eron, 1953, Eron et al., 1955, Zubin, Eron, & Schumer, 1965) indicating themes that appear in 40% or more of the normal group, and which themes were characteristic of less than one percent, less than 10%, and less than 20% of the normative group. Also published was a table indicating unusual formal characteristics, such as alternative themes for the same picture, central character not in the picture. In addition there was a table indicating usual deviations for each card and a five point scale for rating emotional tone, ranging from very sad to very happy, and a seven point scale for rating outcome ranging from failure to great success. These scales were shown to have good interrater reliability. In studies of emotional tone and outcome done on populations of college men and women, Navy men, adolescent girls and psychiatric patients the correlations ranged from .79 to .90 (Eron & Auld, 1954; Eron et al., 1955; Garfield & Eron, 1948; Liccione, 1955; Sarason & Sarason, 1958).

According to Eron's system (1950), a story such as the following would be classified as: Disequilibrium, Interpersonal Parent Pressure.

> The boy doesn't want to play the violin, but his mother makes him do it and tells him he cannot go out to play unless he gets the piece perfect. The boy begins to cry but he practices as he is told.
>
> A story such as: The boy doesn't like to practice but he dreams of being a great violin player, so he practices hard would be classified as: Disequilibrium, Intrapersonal, Aspiration.
>
> A story such as: The boy is very happy because he is able to study the violin which he loves and which he enjoys practicing would be classified as: Equilibrium, Impersonal, Favorable Environment.

Eron suggested that through the use of these scales, checklists, and norms, it was possible to tabulate all the essential elements of a TAT protocol and to isolate

those elements which were either deviant or peculiarly characteristic of the subject. These items then could be used to furnish the basis for inferences to be made about the subject in personality interpretation. How the TAT interpreter used these elements to derive a meaningful personality description of the individual, Eron believed would depend on the theoretical orientation of the individual clinician or researcher.

MURSTEIN'S METHOD

Murstein, who has long been interested in making the TAT a useful clinical instrument, followed in Eron's path in trying to establish norms which could be used as a basis of comparison for the individual protocol. Murstein (1963, 1965) wrote and edited two books in which the TAT was given considerable attention. In 1972 he published a system of scoring based upon stories to the 20 TAT cards given by normal college men and women. There were 3,890 usable stories, which Murstein classified according to who was in the picture, what was going on, what was happening in the story, why it was happening, outcome of the story and affect shown by the characters. A notational system was developed. For example a story in which a man was expressing negative affect for a woman would be shown as: man$\xrightarrow{-}$woman. If the woman was unhappy as the result of the man's negative affect, the notation would appear as: man$\xrightarrow{-}$woman $-$. A story of a man and woman who were comforting one another because they had lost a son would be shown as: (son $-$ X$\longrightarrow -$ man $+\underset{+}{\rightleftarrows}+$ woman $-$.

In tabulating the material, Murstein, like Eron before him, found that most of the stories were negative in tone, but had positive endings. He also found, as had Eron, that many of the intuitive ideas held by some clinicians regarding the significance of certain responses to the cards were not supported. For example, although Murray had considered the figure in Card 3 BM to be a male, 90% of these college men and women saw the figure as a girl or woman.

Murstein found that of the possible 182 comparisons made between the stories told by the different sexes, 45 were significant. Among these differences, women were found to give more positive stories and more positive endings than men. In addition, Murstein found large differences among the stories told to the various cards. He concluded that in view of these differences between cards, knowledge of the stimulus seemed mandatory before one could determine the extent of projection in the TAT response. He felt his norms offered a framework against which to compare the TATs of similar college adults and that deviations in story construction from the normative responses would have important implications for personality assessment.

DANA'S APPROACH

While Murstein and Eron collected information based on 20 cards for each sex in order to develop normative data, a number of other authors used only a selected number of cards. Dana (1955, 1956a, 1956b, 1959) believed that the previously developed series were cumbersome and felt there was a need to devise a system that was simple and could be scored by a clerk. He conducted a series of studies using five cards. He gave cards 2, 3BM, 4, 6BM, and 7BM to 50 normal, 50 neurotic

and 50 psychotic men (1955). He then cross validated these results with a group of 30 normal, 30 neurotic and 30 psychotic men (1956a). He also collected data from 50 normal, 50 nonhospitalized neurotic, and 50 psychotic women, giving them cards 2, 3GF, 4, 6GF and 7GF (1956b).

Dana's scoring system involves three categories: Perceptual Organization (PO), Perceptual Range (PR), and Perceptual Personalization (PP) (1959). Perceptual Organization involves the ability to follow directions given when the test is administered. It is scored for presence or absence of seven categories: card description, present behavior, past events, future events, feelings, thoughts, and outcome. Perceptual Range is scored by giving a plus or minus for presence or absence of three stimulus properties listed for each card. These stimulus properties were chosen for each card from the responses given by 90% of the normal group and include, for example, the mention of the young girl, woman and man; the farm or field; and books or school for Card 2. Perceptual Personalization is scored a plus for expressions that have no obvious reference to the story told and include comments such as: "I can't figure this out"; questions such as "what's this supposed to be"' criticism such as "this picture is silly"; personal references such as "this reminds me of my mother" or any other remarks that have no connection with the story.

Although these three categories relate only to the formal aspects of the cards, Dana found that by using them he could predict 82% of the normal men and 90% of the normal women, 86% of the neurotic men, 65% of the neurotic women, and 100% of the psychotic men and 74% of the psychotic women. His scorer reliabilities were high: 76% for PP, 91% for PR, and 95% for PO. Dana provided t-scores for each category and noted that the scores could not only differentiate between normal and psychiatric groups, but also could represent a degree of psychiatric illness. He believed that use of such an easily scorable, objective system did not imply abandonment of content analysis, but was merely a formal aid in this process.

ORIENTATION OF McCLELLAND AND ATKINSON

Two psychologists who worked with the TAT almost from its inception are McClelland and Atkinson. As early as 1948 they had published material on the projective expression of needs, utilizing the TAT. Other articles followed and in 1953 they published a book "The Achievement Motive" in conjunction with Clark and Lowell. This book contained a summary of research on the achievement motive conducted over a six year period by a number of psychologists who worked with the authors: first at Wesleyan and then at Harvard.

In the book is outlined a theory of motivation and much research relating to the arousal and measurement of motivation and its relation to behavior. McClelland and Atkinson were not interested in the TAT as a clinical instrument, but used it as a means of measuring motivation. According to Atkinson, he had been involved in a "number of uniformly unsuccessful attempts to devise a new and better assessment device for this motive (achievement) than the TAT but none were successful" (Atkinson & Feather, 1966, p.299).

Because McClelland and Atkinson and their associates were not concerned originally with the TAT per se, but saw it as a means to measure fantasy which would give indications of achievement motivation, they used only three of the original TAT cards and supplemented them with five other cards. Some of the eight had high cue value for achievement; that is, they elicited more achievement motivation

than other cards. The cards were combined into two sets of four, each containing cards of high and low cue value. Stories from four cards are considered adequate to assess achievement motivation. McClelland and Atkinson recommend using a group procedure in which the pictures are projected on a screen and the subjects write their own stories with an imposed time limit. They developed a system for scoring that was derived from an analysis of the content of the stories told. According to McClelland, "We assume that the thought processes of an individual are in part determined by his present state of motivation and that in complying with a request to write imaginative stories he reveals the content of his dominant thoughts at the time of writing and hence, indirectly, his state of motivation" (1953, p. 194). McClelland and Atkinson originally had been interested in need affiliation and need succorance as well, but these areas have not been as fully developed as the need achievement area. One of their co-workers, Veroff (1958) developed a system of analyzing the need for power, both personal and social.

The McClelland and Atkinson scoring system (1953) for need for achievement (n Achievement) using the TAT cards is based upon determining whether any of the characters in the story has an achievement goal. This is indicated by evidence that the character wants to perform better or cares about performing better. If this is established, the story is scored for further evidence of achievement motivation. A total of 11 points may be obtained for one story, depending upon whether the story contains a variety of different statements relating to achievement. The two stories that follow give some idea of how the scoring system is used (McClelland et al. 1953, McClelland, Atkinson, Clark, & Lowell, pp.336, 342).

The young fellow is showing anxiety about an impending operation. Some malady has been diagnosed. The doctor has prescribed an operation to remove the source of illness and the boy wonders about the advisability of such a plan. The operation will be performed and successfully.

The boy has visions of becoming a surgeon and has projected his mind to the day when he will be performing operations. He has had a concern for people in the past and desires to be of service to mankind. The boy is thinking of the day when he will be a successful surgeon. He is looking forward to that day very seriously. It is likely that he will strive and reach his goal.

The first story is not scored for achievement and receives a zero. The second story is scored for achievement and receives a score of +4 because in addition to giving evidence of an achievement goal, there is evidence of a need for achievement and an anticipatory goal state and the plot is elaborated in such a manner that achievement is the central theme of the story.

The need for affiliation and power are scored in a similar fashion. To be scored for affiliation, one of the characters in the story must want to be with someone else or to enjoy mutual friendship. If this is established, the story is then scored for a total of seven points relating to aspects of affiliation. To be scored for need for power, one of the characters in the story must desire to have an impact or to make an impression on other individuals. If this is established the story is then scored for 11 aspects relating to the power motive.

The scoring system developed by McClelland and Atkinson takes considerable time and practice to master, and Atkinson characterizes a judge with limited scoring experience as one who scored only 400 stories. Two minutes per story is considered

sufficient time for scoring when the criteria have been learned. Reliability coefficients of .95 and .96 are reported for various judges scoring four story protocols of 30 subjects. Such high scorer reliability requires that the judges agree in advance to score only what the manual covers and not to score a particular category if in doubt. Students are reported to require scoring experience and rechecking of scoring with at least 120 protocols before scoring reliability reaches .90.

Other research (McClelland & Atkinson, 1953) indicated that motivation scores increased by preceding administration of cards with a stressful situation in which pressure for achievement was high, and that motivation scores decreased when administration was preceded by a relaxing situation in which there was no pressure for achievement. Some individuals score high on achievement motivation even under relaxed conditions with low cue pictures. These individuals are considered to be responding primarily to internal cues of unusual strength. The authors also found that although the high cue pictures elicited more achievement motivation than low cue pictures from both high and low achievers, high achievers tended consistently to have higher overall scores than low achievers.

Although Atkinson was not primarily interested in the TAT as a test in 1953 when "The Achievement Motive" was first published, in his preface to the 1976 edition he acknowledges the value of the TAT.

> I refer to the sensitivity of thematic apperception or fantasy to experimental arousal of motivation; to its ability to assess relatively stable and general individual differences in the strength of motives as expressed behaviorally in a variety of ways and to its ability to reflect individual differences in the antecedent formative life histories of individuals.... The implications of this are that content analysis of the stream of operant imaginative thought (thematic apperception) has a more solid theoretical foundation than ever before and must, I believe, be considered the most promising and virtually untapped resource for future study of personality. (1976, preface)

Atkinson expands on these sentiments in an article in the American Psychologist in 1981 in which he proposes using this method, along with others, to study a new conceptualization of personality.

To check for achievement motivation in the general population, Veroff (1959) gave six TAT cards to a representative sample of the adult population of the United States, consisting of 1,619 men and women. The stories were scored for need for achievement, affiliation, and power. The results showed that in the youngest age group, 21 to 34 years, there was a simple positive relationship between occupational status of the men and the achievement motivation scores. For the middle aged, 35 to 49 years, there was a slight diminution of achievement motivation for the men in the professional and managerial classes, but an increase for those in the lower occupational levels. In the oldest group there was a decline in general motivation except in the managerial group. Veroff explains these results as an interaction between occupational status, age, and need for achievement, and indicates that it is clear that there is not a direct relation between achievement motivation scores and occupational status.

Roe (1953a) who studied the TAT stories of eminent biologists, social scientists, and physicists found a similar result. She had expected that there would be indications of strong achievement motivation in these stories, but this did not occur. She asked McClelland and his associates to examine her data.

They agreed that the sort of need for achievement which my subjects must have had, judging by what they have actually done, is not reflected in their TAT protocols, with a few exceptions. It is very possible that the fact that they have achieved has reduced the intensity of the need to do so. (p. 193)

Another confounding factor that Veroff, Roe, and others have found is that achievement motivation scores based on the McClelland and Atkinson scoring tended to be related to the length of the story. Roe found, for instance, that because physicists tended to give short, unimaginative stories, they scored lower in achievement motivation compared to the social scientists, who were inclined to give elaborate and longer stories.

One of the more impressive aspects of the work done by Atkinson and McClelland has been the expansion of research, using this method, not only to educational and business institutions in the United States, but to other countries of the world. McClelland particularly extended the use of this method to the field of economics and promoted training workshops in many different countries. In his book, "The Achieving Society," (1961) he advances the belief that achievement orientation is an identifiable trait, measurable through the use of thematic apperception and other methods, and that achievement orientation has been associated with economic growth in all societies throughout the ages.

The number of studies carried out that attest to the validity of this concept are too numerous to review, but a recent study published in 1979 by Singh indicates the extent to which use of this method has spread through the world. In this study the author used six of the McClelland and Atkinson cards and obtained stories from 200 Punjabi farmers who who were middle class independent entrepreneurs. The achievement rating of the stories, using the McClelland and Atkinson scoring and the Sentence Completion Test scores (Rogers & Neil, 1966) for achievement, were compared with the growth rate of the farms within the previous five years. Only the TAT stories showed significant differences between high and low growth rate farmers. Indeed there was a correlation of +.68 with per acre increase in farm output. A second study, this time using owner-operators of small-scale Punjabi industrial organizations, showed a high correlation between TAT scores and production rates.

ARNOLD'S STORY SEQUENCE ANALYSIS

Another psychologist who developed a quantifiable method of scoring the TAT is Arnold. Arnold was one of those who took part in the study carried out by Schneidman in 1951 and has continued to show interest in the TAT through research and publication. Arnold's method and rationale are outlined in "Story Sequence Analysis" (1962). Her system is based upon extracting an import from the content of the story told to TAT cards. She rejects the notion put forth by others that the theme of the story should be analyzed. She states,

We never can tell whether themes revealed in the TAT accompany behavior or are an alternate to behavior; whether themes missing in the TAT indicate lack of corresponding need, its blocking by ego-defenses, or its being acted out in reality. It almost seems as if we had to agree with Lazarus, who insists that we will never be able to predict behavior from the TAT alone. (1962, p. 12)

This conclusion by Lazarus that Arnold mentions is based on the sometimes contradictory results of studies with the TAT using theme analysis. For example, TAT stories of hungry individuals showed more food themes than stories of sated individuals, but when hunger was prolonged, number of themes of hunger or eating dropped. In studies of aggression, it was found that although aggressive themes increased when storytellers were subjected to frustration before the test was given, aggressive themes increased principally in a group that scored low on a hostility scale, while high scorers tended to have fewer aggressive themes. In addition, stories by aggressive adolescents contained no more aggressive themes than did stories of well behaved adolescents. It was found, however, that aggressive boys told more stories in which aggression was not punished.

With these studies in mind, Arnold proposes that the outcome of the story and the way the story is told is what illustrates the storyteller's motivation. She suggests condensing each story into an import that leaves out incidental details but preserves the kernel. The story import is said to show how the storyteller thinks people usually act and how they should act, what actions are right or wrong, what actions will lead to success and to failure, what can be done when danger threatens, and what things to strive for. In trying to formulate the import, Arnold suggests all theoretical preconceptions be set aside, because one should not try to interpret the story, guess what particular childhood experience a picture might recall, or determine with whom the storyteller identifies. The import is written from the point of view of the main character without an assumption that the main character is the storyteller. Since it is the action, plot and outcome that matter, not the picture upon which the story is based, nor even the characters in it, Arnold eliminates the need for normative data for each picture. She recommends using 13 cards from the original TAT set namely, 1, 2, 3BM, 4MF, 6BM, 7BM, 8BM, 10, 11, 13MF, 14, 16 and 20. She states, however, that she has obtained as good imports from pictures found in magazines and from short descriptive sentences such as: "Tell me a story about a boy and a violin," as from the TAT cards.

Based upon this orientation, Arnold developed a five point scoring system for outcomes of stories ranging from +2 to −2. These scores are added to give a motivation index, which is used to measure high or low achievement and positive or negative attitudes toward right and wrong and human relations. According to Arnold, high achievers tell stories in which success and happiness follow actions for ethical, well-intentioned, rational, or religious motives, and indicate that omitting to act for such motives brings failure and unhappiness. Successful, happy people are optimistic and prefer altruistic, ethical values, take responsibility for their acitons, believe wrongdoing deserves punishment, and that amends should be made. They have a concern for injustice, but are not cynical. In terms of human relations, high achievers tell stories in which people cooperate with others and contribute to the success or happiness of others.

In contrast, the low achiever tells stories in which the characters have no goals or ones that are shifting, uncertain, or antisocial. Success, if it does occur, is a matter of chance or fate. Wrongdoing is shown to be caused by external factors, but may be amended by admitting fault and expressing sorrow, or may be covered up by those who love the wrongdoer. Wrongdoing often is shown as a matter of personal relations and social conscience. Stories of low achievers indicate that lack of love leads to unhappiness, failure, or difficulties, and that rejection has exaggerated effects. For these reasons, the characters conform to the opinions of others

and yield for the sake of peace or social relations. The characters act as though one is doomed to failure unless others help, even before much of an effort has been made. Love is seen as the solution to many problems and if there is no love, blind rebellion against authority is supported. The two stories that follow indicate Arnold's scoring.

The man is talking to his friend about robbing a bank. He decides to do it even though his friend refuses to take part. He robs the bank, is caught and put in prison.

Such a story would have an import "Crime does not pay" and would be scored positively as +2.

A story with a similar theme to the same card has a different import and scoring.

Two men plan to rob a bank. They are aware it is dangerous and wrong, but they want the money. They hold up the bank and get away with $100,000. They never get caught.

The import for this story is "Even though it is dangerous and wrong, success in crime is possible." This story is scored −2.

If the above stories were scored for themes, both would indicate a concern with crime. However, if these stories are scored as imports, a different approach is shown and the indications are that one storyteller is saying that crime is something to be avoided because it will be punished, whereas the other indicates an awareness that crime is wrong, but that it may succeed. From a thematic viewpoint both storytellers are oriented to success, but the first rejects crime as a means because it is not seen as leading to success, whereas the second indicates that crime might be something to consider because it might lead to success.

Arnold's method involves looking not only at imports and scoring each story, but requires evaluation of the sequence of stories, because the imports taken in sequence given a connected statement of the storyteller's principles of action and motivational pattern. To be successful with Arnold's method requires good clinical skills, for the motivational pattern derived from the sequence analysis is used to hypothesize how an individual would react to situations in life. The results cannot be used in a simplistic fashion to indicate that one individual will commit a crime and the other will not. If one considers the two stories, it is possible that neither storyteller would ever rob a bank, but if other stories in the sequence expressed attitudes similar to those in the two stories, one might be justified in hypothesizing that the teller of the first story would be unlikely to be involved in something dishonest, whereas the teller of the second story might become involved if it was felt it would be successful, if the need was strong enough and the possibility of getting away with it existed.

Arnold's system clearly has ethical and social normative aspects, which she acknowledges and defends, stating that too much of personality research has been drawn from the records of abnormal and disturbed individuals. It is apparent, also, that the high achiever as described by Arnold, has characteristics that are congruent with an individual with an internal locus of control (Rotter, 1960), whereas the low achiever has characteristics of an individual with an external locus of control. The fact that this aspect of personality may be derived from the TAT has been demonstrated by Dies (1968), who found that, by using a five point scoring system, nurses with internal and external locus of control, as measured by the I/E

scale, could be accurately differentiated. Eighty percent of the nurses were correctly classified on this dimension.

Early use of the Arnold system with a group of 300 teachers of grades 4 through 12 in 14 different schools (Burkhard, 1958) and a group of 60 enlisted Navy men (Petrauskas, 1958) showed that the scoring system significantly differentiated teachers who received high ratings from students from those who received low ratings. It also differentiated enlisted Navy men who were awaiting court martial proceedings from those who had no infractions of the military code. In a study of 91 college men and women (Garvin, 1960), the correlation between the Arnold scoring and grade point average was shown to be +.85 for men and +.83 for women. This was considerably higher than the correlation between grade point average and intelligence, which was +.63 for men and +.50 for women. In a study of 20 executives, Steggart (1961) found that 19 had positive motivation scores, which was not surprising since all were executives. But those who were classified as high achieving executives had scores ranging from 3.3 to 3.8, whereas low achieving executives obtained significantly lower scores ranging from 1.7 to 3.2. In another study (Honor & Vane, 1972), 50 high achieving and 50 low achieving high school students were classified with 95% accuracy using Arnold's system, whereas only 75% accuracy was achieved using a paper and pencil test of achievement motivation.

CURRENT STATUS

As may be noted from the foregoing, Arnold, Atkinson, and McClelland, and Dana, Eron, Murstein and Veroff all have invested much time and effort to develop systems to evaluate TAT stories so that the method might be considered not only useful, but also reliable and valid. The question to be answered therefore is, have these efforts answered some of the criticisms that have been leveled at the TAT with respect to reliability and validity.

In terms of reliability, almost all the foregoing authors report good interrater reliability, providing the scorers have adequate experience. In a review of a great number of interrater studies of reliability, using concepts such as narcissism, exhibition, sex conflict, Murstein (1972) reports reliability coefficients ranging from .37 to .90, most in the upper ranges.

Entwisle (1972), in examining fantasy based measures of achievement motivation (i.e., the McClelland and Atkinson system), spends most of the article documenting the low internal consistency estimates of reliability. On the basis of a study in which 665 ninth graders were given four especially developed pictures, there was 92% interscorer agreement, but intercorrelations on the individual picture scores with one another were low, ranging from −.01 to .19. When the tests were rescored using a 0 or 1, depending upon whether achievement imagery was present or absent and the remaining ten scoring categories were ignored, scores obtained with each method were similar. Correlations were .88 for all black students and .92 for all white students in the ninth grade sample. As might be expected, interitem correlations continued to be low, and Entwisle suggests that using four pictures is equivalent to having a test with four items, which would automatically result in low reliability. Entwisle gives other examples of previous research in which low reliability coefficients occurred in studies using the TAT and challenges the high results of some previous studies done by McClelland and Atkinson (1948, 1953). She argues that her studies show that the mean value of reliability defined as test

homogeneity tends to be .30 across many studies. Entwisle also reviews studies in which the results of motivation scores obtained using TAT like cards and other tests of achievement motivation have low correlations.

In order to obtain high internal consistency coefficients for the TAT, it would be necessary for a theme such as aggression, not only to be expressed in one story, but in every story. This assumes that the story teller sees every situation in the same light and is a single dimension individual or that the test is a single dimension test. Of course, this is not the case. It is probable that achievement and affiliation may be scored for the same story, since they frequently appear together. Usually, if the set of 11 or 12 TAT cards used by Arnold is given, a trend is shown in the stories. There may be a major theme of achievement in the initial stories, usually stimulated by the picture of the boy with the violin, followed by the picture of the girl on the farm. Both stories often include themes of parent pressure and themes involving determination to make one's own way in life, or to conform to the expectations of others, or conflict about the attempt to do both. These stories frequently are followed by stories of interpersonal relations, usually beginning with the picture of a man trying to break away from a woman, followed by a woman looking into an empty room, followed by a picture of an older man and younger woman, and then a picture of a younger man and older man. These cards have certain cue values for certain themes. From the point of view of Arnold's method the theme per se is not important, but how the problem is resolved is important in determining the achievement score.

Test-retest reliability is a difficult area for the TAT because there have been many studies to show that the content of the stories changes with time. This is not surprising because a retest is psychologically not the same as an initial test. Adults and children tend to change their responses, not only on the TAT stories, but on intelligence tests which are considered to have good reliability and validity. In a recent study (Vane & Motta, 1980) 90 to 100% of a group of kindergarten children changed answers on a vocabulary test when the retest was only one day later, and 40 to 80% of another group of kindergarten children were inconsistent in responding to Stanford Binet and Wechsler Scale items when the retest interval was only 15 minutes. It is interesting that the MMPI has a so-called "lie detector" scale built into it to determine when subjects give contradictory answers to similar items. Since this happens with considerable frequency it would appear that people tend to change their answers either purposefully or inadvertently while a similar item appears on the same or a repeated test. It is not surprising, therefore, that people tend to tell different stories to a repeat of the same or equivalent set of TAT cards.

Most users of the TAT acknowledge that the reliability will be low for the foregoing reasons. Atkinson (1981) speaks of this problem and claims that "the validity of thematic apperceptive measures of motives does not depend on reliability; and correlations among different behavioral expressions of the same trait are expected to be quite modest even prior to consideration of measurement error" (p. 117).

Criticisms of validity of the TAT are associated with the fact that presence of certain themes or scores for certain traits do not directly relate to behavior of individuals tested. In other words, an individual may tell stories in which much aggression is present, but does not show this in his actions. Many clinicians have no difficulty reconciling this discrepancy because the protocol would be interpreted to mean that the individual really is aggressive, but is afraid to be so. Because of

this fear, he has built a successful defense against aggression and thus appears meek and mild. Occasional outbursts of aggression would be explained as a result of breakdown of defenses. On the other hand, if an aggressive individual told stories with many aggressive themes, he would be considered an aggressive individual. In other words, when aggression appears in the stories, there is a tendency to consider that the storyteller is basically aggressive. And when anxiety appears in the story, the teller is considered to be anxious. Such interpretations have certain face validity, because almost all individuals are anxious or aggressive at times.

This belief that behavior described in the story is characteristic of that trait in the storyteller is a common one. It is tied to trait theory and the traditional approach to personality. It is based on the assumption that an individual's personality consists of a combination of basic traits that will be manifested independently of the situation and that these traits will appear consistently across time. Within such a theoretical approach, test signs of underlying traits are considered to be more valid indicators of the "real" person than the observable behavior, which may be inconsistent with the test signs. Such a theory is impossible to validate or to prove wrong and accounts for much research that indicates that personality tests such as the TAT are invalid and unreliable.

There are those who challenge trait theory, among them Atkinson (1981) and Thorne (1973a, 1973b, 1974), because their research suggests that manifestations of personality characteristics vary over time in the same person and under different circumstances. Atkinson notes that the early learning theorists (Hull, 1935; Tolman, 1932) made the distinction between what an animal does and what it has learned. More recently, Bandura (1969) has carried out a series of experiments that indicate children learn much behavior, but that they usually display it only when in the proper stimulus situation. Therefore, certain behaviors are situation specific. Atkinson believes not only that a new theory of motivation and personality is needed, but that his recent research indicated that classical test theory is not applicable to tests such as the TAT.

THE FUTURE

Whether Atkinson's views will prevail remains to be seen, but at present trait theory in combination with an analytic or dynamic orientation and reliance upon individual interpretation rather than the use of one of the specific quantitative systems for interpretation of the TAT is widespread. This is shown by the results of a survey inquiring into the extent to which the TAT is being taught in psychology programs throughout the country today. A questionnaire was sent to the directors of the 163 APA accredited Doctoral Programs in Clinical, Counseling, and School Psychology listed in the December 1980 issue of the American Psychologist. Returns were received from 91 directors or 56% of the programs. Of this number, 18 indicated they did not teach projectives at all and eleven said they taught projectives but not the TAT. Of the remaining 62 programs, 54 required a course in projectives and 8 recommended it as an elective. A representative comment was made by one professor who stated; "I cannot allow more than a very few hours to the TAT. If it were replaced with a thematic procedure of a more efficient format, it would deserve more thorough coverage" (Webb, Note 1).

The question, "In this course what method does the instructor teach to analyze the TAT?" gave the following choices: Murray, Eron, McClelland, and Atkinson,

Murstein, Arnold, Dynamic, Psychoanalytic and Other. Because many of the respondents checked several methods, the numbers exceed 100% of the sample. In reply, more than half, or 44 instructors, said they used a dynamic method of analysis. This was followed by 25 who used a psychoanalytic method. Murray's need-press system was used by 20. Of all the systematic scoring systems, none was used by more than ten. This was Murstein's method, followed by six each for Eron and McClelland and Atkinson and two for Arnold. The approach of Bellak (1954) and of Tompkins (1947) are being taught in three programs each.

The results suggest that the great majority of students, who will be the practicing psychologists of tomorrow, will be using the TAT intuitively or will not be using it at all. Despite the fact that the TAT is at present a popular clinical instrument, it is not being taught in one-third of the Doctoral programs in Clinical, Counseling, and School Psychology. Unless this trend is reversed, it is probable that the TAT will be relegated to a position of a research instrument, used only for special types of projects and that fewer and fewer clinicians will use it. The critics of the instrument have been effective in making clinicians feel uncertain about its value, and the complexity and time consuming aspects of the test have worked to make clinicians wonder whether it is worth the time and effort to teach and use a method so open to question. This combination of factors has operated to discredit the TAT in the eyes of almost all of those with a scientific orientation, with the exception of a few who believe it has the potential to become the useful, reliable, and valid clinical instrument that Murray envisioned when he developed it. Time will tell whether these believers will be able to rescue it from possible oblivion.

The likelihood of this occurring may rest upon an overall attempt to coordinate the findings of the many who have contributed studies in this area. Too many reviews that have been published tend to report negative and positive findings without critical evaluation. This is understandable, for critical evaluation is difficult because of the enormous number of studies. If such a task is to be undertaken, it would be important to consider only well designed studies that are similar enough in terms of method and population to provide solid evidence for or against valid interpretations of the TAT. If it were possible to obtain some of the original protocols from researchers such as Roe, Dana, and others, then evaluate them using the different systems previously discussed, it might be possible to assess the usefulness of some of the systems. Additional data need to be gathered in a systematic fashion with a view to establishing a test that will offer information to the clinician and researcher commensurate with the time and effort expended to obtain it. Unless this is done, it is doubtful that the TAT will survive as a popular clinical instrument.

REFERENCE NOTE

1. Webb, W. W. Personal communication, May, 1981.

REFERENCES

Arnold, M. B. *Story sequence analysis.* New York: Columbia University Press, 1962.

Atkinson, J. W. Studying personality in the context of an advanced motivational psychology. *American Psychologist,* 1981, **36**, 117–128.

Atkinson, J. W., & Feather, N. T. *A theory of achievement motivation.* New York, Wiley: 1966.

Bandura, A. *Principles of behavior modification.* New York: Holt, Rinehart, & Winston, 1969.

Bellak, L. *The Thematic Apperception Test and the Children's Apperception Test in clinical use.* New York: Grune & Stratton, 1954.

Brown, W. R., & McGuire, J. M. Current psychological assessment practices. *Professional Psychology*, 1976, **7**, 475–484.

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

Dana, R. H. Clinical diagnosis and objective TAT scoring. *Journal of Abnormal and Social Psychology*, 1955, **50**, 19–24.

Dana, R. H. Cross-validation of objective TAT scoring. *Journal of Consulting Psychology*, 1956, **20**, 33–36. (a)

Dana, R. H. An application of objective TAT scoring. *Journal of Projective Techniques*, 1956, **20**, 159–163. (b)

Dana, R. H. Proposal for objective scoring of the TAT. *Perceptual and Motor Skills*, 1959, **9**, 27–43.

Dies, R. R. Development of a projective measure of perceived locus of control. *Journal of Projective Techniques and Personality Assessment*, 1968, **32**, 487–490.

Entwisle, D. R. To dispel fantasies about fantasy-based measures of achievement motivation. *Psychological Bulletin*, 1972, **77**, 377–391.

Eron, L. D. A normative study of the Thematic Apperception Test. *Psychological Monographs*, 1950, **64**, No. 315.

Eron, L. D. Responses of women to the Thematic Apperception Test. *Journal of Consulting Psychology*, 1953, **17**, 269–282.

Eron, L. D., & Auld, F. A study of Thematic Apperception Test stories and sentence completion of subjects in Operation Hideout. *Bulletin of Medicine and Surgery*, Navy Department, MRL Report, No. 243, February 1954.

Eron, L. D., & Ritter, A. M. A comparison of two methods of administration of the Thematic Apperception Test. *Journal of Consulting Psychology*, 1951, **15**, 55–61.

Eron, L. D., Sultan, F., & Auld, F. The application of a psychometric scoring procedure to a group modification of the TAT (N-TAT). *Journal of Consulting Psychology*, 1955, **19**, 83–89.

Eron, L. D., Terry, D., & Callahan, R. The use of rating scales for emotional tone of TAT stories. *Journal of Consulting Psychology*, 1950, **14**, 473–478.

Garfield, S. L., & Eron, L. D. Interpreting mood and activity in Thematic Apperception Test stories. *Journal of Abnormal and Social Psychology*, 1948, **43**, 338–345.

Garvin, J. A. *A Thematic Apperception Test study of non-intellective factors related to academic success on the college level.* Doctoral dissertation, Loyola University, Chicago, 1960.

Gonzales, J. M. Psychological tests in Columbia. *Revista Latinoamericano de Psicologia*, 1977, **9**, 429–435.

Honor, S., & Vane, J. Comparison of TAT and questionnaire methods to obtain achievement attitudes of high school boys. *Journal of Clinical Psychology*, 1972, **28**, 81–83.

Hull, C. L. Special review: Thorndike's fundamentals of learning. *Psychological Bulletin*, 1935, **32**, 807–823.

LaPointe, R. E. The use of psychological tests by Ontario psychologists. *Ontario Psychologist*, 1974, **6**, 75–82.

Lazarus, R. S. A substitutive-defensive conception of apperceptive fantasy. In J. Kagen & G. S. Lesser (Eds.) *Contemporary issues in thematic apperceptive methods.* Springfield, Illinois: Charles C. Thomas, 1961.

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

Lubin, B., Wallis, R., & Paine, C. Patterns of psychological tests used in the United States: 1935–1969. *Professional Psychology*, 1971, **2**, 70–74.

Lindzey, G., & Heinemann, S. H. TAT: individual and group administration. *Journal of Personality*, 1955, **24**, 34–55.

Lindzey, G., & Silverman, M. Thematic Apperception Test: techniques of group administration, sex differences and the role of verbal productivity. *Journal of Personality*, 1959, **27**, 311–332.

McClelland, D. C. *The achieving society.* Princeton, N.J.: Van Nostrand, 1961.

McClelland, D. C., & Atkinson, J. W. The projective expression of needs. I. The effect of different intensities of the hunger drive on perception. *Journal of Psychology*, 1948, **25**, 205–232.

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

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L., *The achievement motive.* New York: Wiley, 1976.

Morgan, C. D., & Murray, H. A. A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry*, 1935, **34**, 289–306.

Murray, H. A. (Ed.) *Explorations in personality*. New York: Oxford University Press, 1938.

Murray, H. A. *Thematic Apperception Test manual*. Cambridge, Mass: Harvard University Press, 1943.

Murstein, B. L. *Theory and research in projective techniques*. New York: Wiley, 1963.

Murstein, B. L. *Handbook of projective techniques*. New York: Basic Books, 1965.

Murstein, B. L. Normative written TAT responses for a college sample. *Journal of Personality Assessment*, 1972, **36**, 104–147.

Petrauskas, F. B. *A TAT and Picture Frustration study of naval offenders and non-offenders*. Doctoral dissertation, Loyola University, Chicago, 1958.

Reynolds, W. M., & Sundberg, N. D. Recent research trends in testing. *Journal of Personality Assessment*, 1976, **40**, 228–231.

Roe, A. *The making of a scientist*. New York: Dodd, Mead, 1953. (a)

Roe, A. A psychological study of eminent psychologists and anthropologists and a comparison with biological and physical scientists. *Psychological Monographs*, 1953, **67**, #2. (b)

Rogers, E. M., & Neil, R. E. *Achievement motivation among Columbian peasants*. Unpublished paper, Department of Communications, Michigan State University, 1966.

Rotter, J. Generalized expectancies of internal control of reinforcement. *Psychological monographs*, 1960, **80**, No. 609.

Sarason, B. R., & Sarason, I. G. The effect of type of administration and sex of subject on emotional tone and outcome ratings of TAT stories. *Journal of Projective Techniques*, 1958, **22**, 333–337.

Schneidman, E. S. *Thematic test analysis*. New York: Grune & Stratton, 1951.

Singh, S. Relationships among projective and direct verbal measures of achievement motivation. *Journal of Personality Assessment*, 1979, **43**, 45–50.

Steggart, F. X. *An analysis of some personal and executive characteristics of participants in a university program for federal personnel*. Doctoral dissertation, Loyola University, Chicago, 1961.

Sundberg, N. D. The practice of psychological testing in clinical service in the Untied States. *American Psychologist*, 1977, **32**, 616–624.

Terry, D. The use of a rating scale of level of response in TAT stories. *Journal of Abnormal and Social Psychology*, 1952, **47**.

Thorne, F. C. The clinical priority of psychodiagnosis. *Journal of Clinical Psychology*, 1973, **29**, 382. (a)

Thorne, F. C. The existential study. *Journal of Clinical Psychology*, 1973, **29**, 387–392. (b)

Thorne, F. C. The new factorial ball game. *Journal of Clinical Psychology*, 1974, **30**, 104–105.

Tolman, E. C. *Purposive behavior in animals and men*. New York: Appleton-Century, 1932.

Tompkins, S. *The Thematic Apperception Test: The theory and technique of interpretation*. New York: Grune & Stratton, 1947.

Vane, J. R., & Motta, R. Test response inconsistency in young children. *Journal of School Psychology*, 1980, **18**, 25–33.

Veroff, J. A scoring manual for the power motive. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society*. Princeton, N.J.: Van Nostrand, 1958.

Wade, T. W., Baker, T. B., Morton, T. L., & Baker, L. J. Status of psychological testing in clinical psychology: Relationship between test use and professional activities and orientation. *Journal of Personality Assessment*, 1978, **42**, 3–10.

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