

Measuring Personality at a Distance: Development of an Integrated System for Scoring Motives in Running Text¹

David G. Winter
University of Michigan

Abstract

A new integrated method for scoring achievement, affiliation-intimacy, and power motive imagery in verbal running text (speeches, interviews, literary works, etc.) is introduced as an aid to personality research at a distance. Systematic studies demonstrate that the new scoring system has interscorer and temporal reliability, validity, and convergence with scores from the original TAT-based scoring systems. Group and individual studies of political leaders, as well as studies of conflict escalation, illustrate its range of application. Overall, the performance of the new integrated running text system, which is nonreactive and can be applied to any naturally-occurring verbal material, suggests an alternative to the testing procedures of traditional personality research and assessment.

Personality research is often limited by the way in which personality variables are measured. These limits are especially prominent in the study of individual lives. Often it is impossible to administer traditional psychological tests or carry out an interview. Examples include cases where testing is impossible, such as studies of political leaders, prominent persons, or historical figures who are no longer alive; and situations where testing is possible but undesirable because it is highly reactive and arouses test-taking or self-presentational sets, suspiciousness, defensiveness, or anxiety. Personality researchers, therefore, have begun to overcome this problem by developing indirect methods of measuring personality that can be applied “at a distance,” as a supplement to the more conventional methods of psychobiography (Hermann, 1968). In general, these methods are nonreactive, so that they can give repeated measurements over time without repeated testing; and they are flexible, so that they can measure personality characteristics in a variety of contexts.

Because of widespread interest in understanding the relationship between personality and politics, at-a-distance techniques are probably most widely used in political psychology (e.g., Hermann, 1977), but they have also been applied to literary figures (Porter & Suedfeld, 1981), notable persons, and in fact to college students and other traditional groups of personality research subjects, as an alternative to traditional testing. Prominent among the variables measured in this way are motives (Hermann, 1979, 1980a, 1980b), integrative

¹ A later version of this paper is published in A. J. Stewart, J. M. Healy, Jr., & D. J. Ozer (Eds.), *Perspectives in personality: Approaches to understanding lives* (pp. 59-89). London: Jessica Kingsley, 1991. I am grateful to Abigail Stewart for advice and discussion of several issues; to Joseph M. Healy, Jr. for advice in the final formulation of the scoring system and actual scoring; and to John W. Atkinson, Dan P. McAdams, David C. McClelland, and Joseph Veroff for generously providing data from their own previous research studies. This chapter was prepared with support from the John and Mary R. Markle Foundation, which support is gratefully acknowledged.

complexity (Suedfeld & Piedrahita, 1984; Suedfeld & Rank, 1976), and temperament (Historical Figures Assessment Collaborative, 1977; Simonton, 1986).

This paper reports the development of a single, simplified and integrated system for scoring the achievement, affiliation-intimacy, and power motives, all at once, in both the traditional Thematic Apperception Test (TAT) stories and also any other imaginative verbal material such as interviews, speeches, autobiographies, literary works, and mass media. Since all three motives are scored at once, and since the new scoring system can be learned in about the same amount of time required to learn one of the original systems, the time and expense of scorer training and scoring are considerably reduced.

Motivation involves the domain of personality concerned with goals and goal-directed actions, in contrast to other domains of personality such as cognitions or beliefs, temperament, or social learning history. In recent decades, personality psychologists have developed methods of measuring several important human motives through content analysis of fantasy productions or other imaginative verbal material (see Atkinson, 1958; 1982). The *achievement motive* involves a concern for excellence and unique accomplishment, and is associated with restless activity, moderate risk-taking, using feedback or knowledge of results, and entrepreneurial activity (see McClelland, 1961). The *affiliation-intimacy motives* involve a concern for close relations with others. Sometimes they predict interpersonal warmth and self-disclosure, but under conditions of threat or stress, they can produce a “prickly,” defensive orientation to others (see Boyatzis, 1973; McAdams, 1982). The *power motive*, a concern for impact and prestige, leads both to formal social power and also to profligate, impulsive actions such as aggression, drinking, and taking extreme risks (see Winter, 1973; Winter & Stewart, 1978). These three motives are drawn from Murray’s (1938) comprehensive taxonomy. They are not the only important human motives, but they do involve many of the most common and significant human goals and concerns.

Background

The achievement, affiliation, intimacy, and power motive scoring systems were originally developed to code the brief imaginative stories, each usually about 100-150 words long, that people tell or write in taking a TAT (see Atkinson, 1958). Scoring categories were defined by observing how experimental arousal of the motive in question affected the contents of TAT stories. (See Winter, 1973, chap. 3, for a detailed account of this process.) Each scoring system consists of a generic category of motive imagery, scored as either present or absent in each story, and from six to ten subcategories, each of which can be scored as present or absent only if motive imagery has already been scored for the story. The total score for a story is the number of imagery and subcategories scored as present in that story.

With considerable modification and adaptation, several researchers applied these traditional scoring systems to measure the motives of political leaders or historical figures “at a distance” (Donley & Winter, 1970; Hermann, 1980a, 1980b; Winter, 1976, 1980, 1982; Winter & Stewart, 1977a), to score people’s dream reports and essays (LeVine, 1966), and to

assess motives in a variety of historical and literary documents (Berlew, 1956; Bradburn & Berlew, 1961; Child, Storm, & Veroff, 1958; Cortes, 1960; deCharms and Moeller, 1962; Giliberto, 1972; Henley, 1967; McClelland, 1958, 1961, 1975; Straus & Houghton, 1960). Typical materials scored in these studies--speeches, press conference transcripts, and lengthy excerpts from literary works, all conveniently labeled as "running text"--are usually much longer than TAT stories and do not have any consistent natural divisions into shorter segments. Often these materials are less imaginative than TAT stories, with only a rudimentary plot line. Hence the scoring systems originally developed to code TAT stories had to be substantially modified in order to be applied to running text.

In making changes, most researchers followed the original practice of Berlew (1956, also cited in McClelland, 1958, p. 531), who dropped the subcategories and scored only motive imagery per 100 lines of text. Donley and Winter (1970) elaborated this correction for variability in lengths of text by expressing scores in terms of motive images per 1,000 words. Later experience in scoring political speeches led Winter and Stewart (1977a, pp. 48-49) to propose some extensions and simplifications of the scoring definitions for achievement, affiliation and power motive imagery.

While the results of most of these previous studies have been consistent with the research findings of TAT-based studies of individuals, the convergence and equivalence of the original motive scoring systems with the various modified, adapted versions has never been systematically demonstrated. This chapter first introduces a single codified and integrated version of the motive scoring systems for application to TAT's or any kind of running text. Then, data from several earlier studies are rescored and reanalyzed in order to explore the psychometric credentials of the new integrated running text scoring system and to establish its convergence with the three original scoring systems. Finally, results from several new studies are also presented or cited in order to demonstrate the wide range of usefulness of the new system.

The results presented here should make it possible to use this new scoring system with psychometric confidence, both to score TAT's with greater simplicity and less expense, and also to investigate many problems and topics--ranging from the intensive study of individuals to studies of personality processes in conflict escalation and other complex social processes--which up to now could not be easily studied with the original motive scoring systems. Examples will be given to illustrate this range of applications.

The New Integrated Scoring System

The integrated system for scoring motives in running text is based on a codification and simplification of the existing motive scoring systems. It can be applied to any verbal material that is at least in part imaginative or "aspirational," rather than being purely factual; or that contains statements about goals, actions, or wishes: TAT stories, interviews, and press conference transcripts, speeches, dialogue, and a variety of literary forms including novels, stories, plays, poems, songs, and mass media programs. Table 1 presents a brief

outline of the contents of this scoring system.² All three motives are scored at once. Only motive imagery is scored; subcategories are dropped. For each motive, the original definitions of motive imagery were simplified, clarified, and rephrased in language that was as broadly applicable (that is, not tied to particular picture cues) and as clear and economical as possible. Over the years, the original motive scoring systems had become “encrusted” with modifications, exceptions, and special points rarely used in TAT scoring and irrelevant to running text. These were either eliminated in the rephrased general definitions, or else omitted completely. These changes were based on the author’s experience of twenty-five years of scoring and teaching the original scoring systems and developing the revised power motive scoring system, as well as conferences with several expert motive scorers.

 Insert Table 1 about here

Definitions of Motive Imagery

Several examples for each motive will illustrate the process of redefinition, simplification, and codification. Readers interested in more detail can obtain the complete scoring system as noted in footnote 1. For example, the original definition of achievement motive imagery--“competition with a standard of excellence”--contained further glosses such as “competitive activity . . . with affective concern over goal attainment or certain types of instrumental activity” and “meeting self-imposed requirements of good performance,” and further elaborations in extensive footnotes (see Atkinson, 1958, pp. 181-183). This was simplified to “adjectives or phrases which positively evaluate performance, such as ‘good,’ ‘better,’ ‘best,’” which is the working definition that most expert scorers actually use. “Long-term [i.e., career] involvement” was a frequent achievement theme in the earliest TAT research, probably because of the particular pictures and subjects used (see McClelland, Atkinson, Clark, & Lowell, 1953, p. 113, 118-119). It was scored as achievement imagery without further explicit concern for excellence, but only by making certain assumptions, under certain circumstances, and with several hedges and cautions (Atkinson, 1958, p. 184). Because long-term-involvement by itself is so tortuously linked to achievement concerns, and because it is rather rare in most TAT stories, it was simply dropped from the new integrated running text system in the interests of simplicity and efficiency. At the same time, the achievement and power aspects of competition, which were somewhat muddled in the two original scoring systems, were more clearly distinguished: references to the *quality* of performance are scored for achievement, while references to the *prestige or impact* derived from the outcome are scored for power.

The original definition of affiliation imagery was concern about “establishing, maintaining, or restoring a positive affective relationship with another person...a relationship

² This brief outline is not adequate for scoring purposes. Copies of the full scoring system, together with practice materials and expert scoring, are available at cost from the author, Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109-1109.

most adequately described by the word friendship.” Further glosses referred to “warm, companionate quality” and “the desire to be liked or accepted or forgiven.” This was rephrased as “expression of positive, friendly feelings toward other persons,” a definition that also reflects much of the spirit of McAdams’s (1980) scoring definition for the intimacy motive (conceptualized as related to but distinct from affiliation). Because of this overlap of scoring content, as well as because of the results to be reported below, the revised scoring definition will be labeled “affiliation-intimacy” in the rest of this chapter. Finally, themes of “unity” and “good relations” among people often occur in political and cultural verbal material; since these usually involve affiliative concerns, specific language regarding them was added in elaborating the definition of affiliation-intimacy imagery.

In the case of power motivation, distinctions which had often confused scorers in the past were clarified and elaborated. For example, concerns for autonomy are not scored unless they are elaborated to include additional concern for power; and power as impact on others is scored, while power as inherent strength or capacity (as in “powers of reasoning”) is not scored.

After all modifications, extensions, and revisions were made, and after several different kinds of material involving different formats and topics had been scored, the entire new integrated running text scoring manual was once again rewritten, in order to organize everything in the most logical and coherent form, and to cover issues that arise in scoring many different kinds of material.

Scoring Conventions

Rules and conventions about units of scoring and units of analysis were based on those introduced by Donley (1968), Donley and Winter (1970), and Winter and Stewart (1977a). In written text, the sentence is the unit of scoring. A single sentence can be scored only once for the imagery of any particular motive, although it can be scored for two or more different motives. Also, if imagery for the same motive occurs in two consecutive sentences, it can be scored only once, unless the two occurrences are separated by imagery of another, different motive. If imagery for a motive occurs in three consecutive sentences, it could be scored in the first and third sentences, and so forth, for every other consecutive sentence in which it occurs.

These rules have two purposes. First, they simplify scoring decisions in cases where what seems to be a single motive is spread out over two sentences. Here are some examples involving achievement motive imagery from TAT stories:

The inventors are becoming very emotional about its outcome. Will it be a success? Yes, they will discover a new metal which will revolutionize industry. (The first two sentences seem to be part of the same image, involving anticipation of achievement. According to the rule, only one image can be scored. If that image is assigned to the first sentence, then the third

sentence, which introduces a new achievement theme of unique accomplishment, can also be scored.)

They probably need more advertising about sales of their products, and in the past they didn't succeed so well. They are probably thinking about ideas to promote more sales. (The first sentence is only scorable because of the second sentence; both are really part of the same image of counteraction following failure.)

Poor management has caused business to slump. They are trying to rebuild the business. (The second sentence is only scored for achievement because of the first sentence, which by defining the previous situation as failure makes it clear that rebuilding is a counteraction.)

These rules also distinguish repetition that merely clarifies the writer's or speaker's meaning (often because of the format used, as in the above examples) from repetition extending over several sentences that indicates special concern about a given motive theme. This latter kind of repetition is more common in rhetoric, such as political speeches, than in imaginative TAT stories.

Once the text has been scored, the resulting scores are expressed as motive images per 1,000 words. (Alternative ways of correcting for length, such as counting the proportion of text sentences containing motive imagery, do not yield correlations with criterion behaviors as high as those obtained using images per 1,000 words.)

Example of Running Text Scoring

The following passage from Abraham Lincoln's 1861 inaugural address illustrates how the integrated running text system is applied:

<u>Ach</u>	The course here indicated will be followed unless current events and experience shall show a modification or change to be proper, and in every case and exigency my <u>best</u> discretion will be exercised, according to circumstances actually existing and with a view and a hope of a peaceful solution of the national troubles and the
<u>Aff</u> <u>Pow</u>	<u>restoration of fraternal sympathies and affections</u> . That there are persons in one section or another <u>who seek to destroy the Union</u> at all events and are glad of any pretext to do it I will neither affirm nor deny; but if there be such, I need address no word to them. To
<u>Aff</u> <u>Pow</u>	those, however, <u>who really love the Union</u> may I not speak? Before entering upon so grave a matter as <u>the destruction of our national fabric</u> , with all its benefits, its memories, and its hopes, would it not be wise to ascertain precisely why we do it?

Psychometric Studies: Description of Data Sets,

Subjects, and Procedures

Data previously collected from a variety of other studies were rescored with the new system, and the results were analyzed in order to provide information about the psychometric characteristics and validity of the new system. All scoring was done by a trained scorer who had demonstrated mastery of the new system (Category Agreement above .85 on each motive with the scoring of an expert; see below in “Results” section for details). In all cases, the scorer was not aware of the original motive scores, scores on other variables, and the experimental condition or other classification of all subjects.

Practice sets of sample TAT stories used to train scorers. Stories in these practice sets probably have more variance in motive imagery than do most groups of TAT stories, so they are most appropriate for a first study of convergence between the two systems. Three sets of thirty stories each for achievement (Atkinson, 1958, pp. 701-719), affiliation (Atkinson, 1958, pp. 736-760), and power motivation (Winter, 1973, pp. 274-296) were rescored according to the new integrated running text system, in order to establish the relationships between the original systems and the integrated running text system at the level of the individual TAT story.

Study of students from “Ivy College.” TAT’s given to 234 male students entering college in 1960 (Winter, McClelland, & Stewart, 1981) were rescored to establish the relationships between the original systems and the integrated running text system at the level of the person.

Motive arousal experiments. TAT stories from several original motive arousal experiments, in which one group (“aroused”) was given experimental arousal of a motive and another (“neutral”) group was given no such arousal, were rescored: Achievement (“success-failure,” “failure,” and “relaxed” conditions from McClelland, Clark, Roby, & Atkinson, 1949); affiliation (Atkinson, Heyns, & Veroff, 1954); intimacy, a new system designed to reflect the positive desire for warm, close interpersonal states in contrast to the more instrumental and defensive concerns of the affiliation motive (McAdams, 1980, study 1); and power (Veroff, 1957; Winter, 1973, pp. 62-63). These data sets were used to determine whether scores from the integrated running text system could successfully differentiate the groups of “aroused” and “neutral” stories which were the foundation and first source of validity of the original systems.

Study of the test-retest reliability of the power motive. Winter and Stewart (1977b) collected data from 70 male and female college students to study the temporal stability of the power motive over a one-week interval, with different groups given varying retest instructions: to tell the same stories, to tell different stories, or to tell whatever stories they liked without regard to whether they were the same or different. This data set was used to determine the temporal stability of the integrated running text system under these three retest instructional sets.

Study of leadership among Naval officers. Winter (1978, 1979) studied the relationship between motivation and supervisors' ratings of leadership in a sample of 22 male United States Navy officers. TATs had been scored with the original scoring system. The new system was used to score critical incident interviews conducted with each officer. This analysis was designed to determine whether motive scores based on verbal material from a situation more like ordinary life and less like a psychological test (and scoreable only with the new system) show stronger relationships to behavior than would motive scores based on traditional TATs. The critical incident interviews were conducted by trained Navy interviewers, lasted about an hour and a half, and covered important success and failure experiences in each subject's present job (see Flanagan, 1954). Using a seven-point scale, each subject's "total competence as a leader and a manager, based on performance in current position" was rated by that subject's immediate superior officer.

Results of the Psychometric Studies

Learnability and Interscorer Agreement

Several persons with no previous motive scoring experience learned the new integrated running text scoring system. After about fifteen hours of practice, each achieved very high levels of category agreement with expert scoring. Category Agreement figures on the final three sets of practice materials, combined across all three motives, were above .85 (see Winter, 1973, pp. 248-249). These figures are quite comparable to those reported for the original scoring systems (see Smith & Feld, 1958), suggesting that the integrated running text scoring system is learnable, objective, and therefore usable for research purposes. With the new system, scorers can be trained to score all three motives in about the time required to learn the original scoring system for one motive, which also suggests a gain in the efficiency of scorer training.

 Insert Table 2 about here

Convergence with Original Scoring Systems

Table 2 presents the correlations between motive imagery scores from the integrated running text system and the original scoring systems for two data sets, including calculations by story and by person. All of these correlations between the two systems are highly significant, suggesting that there is considerable overlap or convergence, even though the running text system uses only imagery, permits multiple imagery scoring per TAT story, and dispenses with the subcategories. (The correlations by story, from the practice sets, are about ten points higher than those by subject, probably because the practice sets have a larger proportion of highly-saturated stories and a greater variance of scores.) Nevertheless, the two systems are by no means identical, and there is a good deal of room for divergence and perhaps differential prediction to action.

Insert Table 3 about here

Differentiation of Aroused and Neutral Stories

For each motive arousal experiment, stories from “aroused” and “neutral” conditions were mixed together and scored blind to condition. In the achievement motive arousal study, the “success-failure” and “failure” conditions, both of which originally aroused the achievement motive, were also combined and contrasted to the “relaxed” condition. Results are shown in Table 3. The integrated running text scoring system clearly differentiates aroused and neutral groups of stories from the original achievement motive arousal experiments. For data from the affiliation motive arousal experiment there were essentially no differences, but the integrated running text system did work well on stories from the intimacy motive arousal experiment.

Taken together with the convergence data reported in Table 1, these results suggest that although affiliation-intimacy scores from the new system may converge with those from the original system, the actual affiliation arousal experiment of Atkinson, Heyns, and Veroff (1954) may have generated complex and ambivalent affective states in addition to concerns with positive, warm interpersonal relationships. (Such a conclusion was also drawn by Boyatzis, 1973, and McAdams, 1980.) Among subjects in the control group of McAdams’ arousal experiment, affiliation-intimacy imagery as scored by the integrated running text system correlates as highly with original intimacy motive scores ($r = .41, p < .001$) as it does with original affiliation motive scores ($r = .40, p < .001$). We may conclude, therefore, that the affiliation-intimacy scoring of the integrated running text system is probably located somewhere between the often-defensive instrumentality of the original affiliation scoring system and the direct goal-state experiencing of the intimacy scoring system (see McAdams, 1980, pp. 430-431).

The integrated running text scoring system significantly differentiated aroused and neutral groups of stories in Winter’s (1973) power motive arousal experiment, and showed nonsignificant results in the predicted direction for stories from Veroff’s (1957) original experiment.

Overall, the integrated running text scoring system is valid in the sense of differentiating aroused and neutral stories from the original arousal experiments, although the differences are on the whole less than those obtained with the original scoring systems. In part this is probably due to the simplifications and omissions in the integrated running text scoring system. On the other hand, since the original systems were actually developed from these arousal experiment TAT stories, they may be excessively tailored to differences in those specific story sets, differences which are not of general validity or usefulness.

Reliability

In their experiments on power motive reliability, Winter and Stewart (1977b) and Lundy (1985) found that test-retest correlation coefficients were strongly affected by the instructions given to subjects on the retest. When told to tell the same stories, power motive reliability was quite high ($r = .61$), but subjects instructed to tell different stories showed reliability scarcely exceeding chance levels ($r = .27$). Because test-retest reliability was also quite high in a third group of subjects explicitly told not to worry about whether their stories were the same or different, Winter and Stewart concluded that the usual TAT retest situation implicitly invokes a self-instruction to change stories, thus producing spuriously low coefficients of reliability. (See also Atkinson, 1982, and McClelland, 1981, for a further discussion of the reliability of projective measures.)

 Insert Table 4 about here

When these same stories were scored for all three motives with the integrated running text scoring system, the results were similar to those of the Winter and Stewart (1977b) study, as shown in Table 4. The “same” condition produces high test-retest correlations across all three motives. The “no instruction” condition produces lower correlations, but ones that are usually higher than in the “different” condition. Pooled correlations for the “same” and “no instruction” conditions are significant and on the high side of those usually reported for the original scoring systems (see Winter & Stewart, 1977b), suggesting that the reliability of the integrated running text scoring system is at least as good as that of the original systems.

Other evidence suggests that its reliability is much higher than this, especially when applied to verbal material other than TAT stories. Scoring press conferences with the integrated system, Winter (1980) reported split-half reliabilities of .62, .77, and .71 for achievement, affiliation-intimacy, and power motives, respectively. Hermann (1979, 1980b) scored affiliation and power motive imagery with a system similar to the integrated running text scoring system, and reported odd-even correlations above .70 for both motives. Winter (1982) also reported a concordance coefficient (3-way rank order correlation) of .74 for the motive scores of presidential candidates on three different occasions. These latter findings support the argument of Winter and Stewart (1977b) that TAT test-retest correlations are spuriously low because of artifacts within the “test” and/or “retest” situation.

Validity: Predicting Leadership Ratings

The study of leadership among United States Navy officers (Winter, 1978, 1979) offers evidence that the integrated running text scoring system, when applied to “ordinary life” verbal material, has predictive validity superior to that of the original scoring system applied to TATs. On theoretical grounds, power motivation should predict leadership performance. Using TATs scored with the original scoring system, however, the correlation with superior officers’ ratings of “total competence as a leader and a manager” is a nonsignificant .05. When the critical incident interviews are scored with the running text system, however, the correlation is .46 ($N = 22$, $p < .05$). While the sample is very small, these results give

increased confidence both in the new scoring system and in its application to “natural” verbal material.

Studies of Political Leaders: Subjects and Procedures

In order to demonstrate the application of the integrated running text scoring system to the study of political and historical figures, who cannot be tested directly and therefore who have to be studied “at a distance,” results and reanalyses from several other studies will be presented.

Southern Africa leaders. Winter (1980) studied 22 political leaders of all types (whites and blacks, heads of state, ministers, and nationalist leaders) from Africa south and east of the Congo River basin. Motives were scored from verbatim press conference transcripts of the period 1975-1977. Eight experts from various disciplines, experienced and knowledgeable about southern Africa, were asked to rate 19 of these leaders on the three motives themselves and on the traits of “war disposition” (defined as “likely to support the initiation, continuation, or escalation of armed conflict”) and activity (“invests great energy in leadership and is more likely to act when challenged than to wait”).

American presidents 1789-1981. The previous studies of motive imagery among twentieth-century American presidents (Donley & Winter, 1970; Winter & Stewart, 1977a) were extended back to George Washington’s 1789 inaugural, yielding a sample of 34 presidents who delivered inaugural addresses at their first inauguration. (Tyler, Fillmore, Andrew Johnson, Arthur, and Ford are not included because they were never elected and inaugurated in their own right, although they all doubtless made remarks of some kind after taking the oath of office. Scores for Theodore Roosevelt, Coolidge, Truman, and Lyndon Johnson are based on the speeches after they were inaugurated in their own right. Because of their brief tenures in office, William Henry Harrison was omitted from all analyses, and Garfield from all analyses except, of course, assassination attempts.) All speeches were rescored according to the final version of the integrated running text system.

Dependent variables, mostly devised and discussed by Winter and Stewart (1977a), include the following: (1) a composite rating of presidential greatness (based on sums of standardized ratings of strength, prestige, activity, and accomplishments) and (2) a rating of idealism, both taken from a poll of over 500 scholars of American history (Maranell, 1970); (3) the number of “great decisions” (momentous decisions, not necessarily wise ones) made (Morris, 1967); (4) assassination attempts (Clarke, 1982); (5) involvement in a war (excluding the continuous and ill-demarcated Indian conflicts), taken from the compilation of Richardson (1960), with additions for 1945-1980 according to his definitions; (6) war avoidance, or peaceful handling of crises that easily could have escalated into war (taken from Small, 1980); (7) number of foreign military interventions not involving war (Emerson, 1973); (8) arms control, defined as the conclusion of an agreement with one or more major powers to limit or reduce actual or potential major weapon systems (starting with McKinley, who was president during the first major international disarmament conference at The Hague in 1899, and continuing through Carter; revised and extended from Winter & Stewart,

1977a); (9) scandal, as reflected both by resignation of a cabinet or major staff member for political misconduct, and also by the number of pages per year of office in Woodward's (1974) compilation of information about presidential scandals; and finally (10) categorization by Barber (1977) as "active-positive" or "active-negative."

Data for some dependent variables were available for all presidents from Washington through Reagan; other measures (such as Maranell's ratings, Barber's categorizations, and many of the historical measures) were only available on some of the presidents. The varying numbers of cases for each variable are noted in the presentation of results in Table 6.

Results of Studies of Political Leaders

Southern Africa Leaders

Among the 22 southern African political leaders, there was virtually no relationship between motives as scored from the verbatim press conference materials and motives as rated by the judges (r 's = -.14 for achievement, .32 for affiliation, and -.15 for power). Which scores predict behavior? Because these leaders are in such different (and changing) political and personal situations--ranging from head of state to exile--there are no objective measures of political behavior that are comparable across all leaders, as used in the study of American presidents. For example, "war entry" would mean quite different things for a head of state, a cabinet minister, or an exiled nationalist guerilla leader. When expert judges, experienced and knowledgeable about southern Africa, are asked to make ratings of "war disposition," or the propensity to resort to violence, however, they are probably able to take account of these differences in situation, opportunity, barriers, and external forces. (Though it seems reasonable that people rated as having a "war disposition" actually do engage in more warlike behavior when other things are equal, nevertheless this connection should be demonstrated by further studies of leaders in more homogeneous situations.)

 Insert Table 5 about here

Table 5 presents the main significant results, which have to do with power motivation. On the basis of both theory and previous research, both "war disposition" (propensity to support armed conflicts) and "activity" should be related to power motivation. As Table 5 shows, power motive imagery scores based on the press conference transcripts are significantly related to both dependent variables, while the judges' ratings of "power motive" are essentially uncorrelated with both dependent variables. Leaders who score high in power motivation are rated as more warlike--more likely to resort to armed force--and more active than leaders who score low in power motivation. For them more than for other leaders, then, war is (in Clausewitz's famous phrase) the "continuation of politics by other means."

These findings are especially important because the motive imagery scores based on press conferences do not correlate with judges' ratings of these "same" motives. On grounds

of shared method variance, or judges' implicit theories of personality, one might have expected the two sets of judges' ratings--of "power motive" and "war disposition"--to be highly correlated. That is, a judge who has already rated a leader as high in power motivation, and who believes that power concerns are one of the personal characteristics that make war more likely, would tend to rate that same leader as high in "war disposition" also (or vice versa). In such cases, significant relationships occur in part because they already exist in the minds of raters as well as (or rather than) in reality (see Mischel, 1968). In the present case, however, the results were obtained across methods instead of within methods.

American Presidents

The extension of previous studies of American presidents was designed to show how the integrated running text scoring system can be applied to score the motives of historical figures and other people who can only be studied at a distance. Table 6 presents the results for the presidents from George Washington to Ronald Reagan. They are consistent not only with prior research on twentieth century American presidents, using an early adaptation of the motive scoring systems (see Donley & Winter, 1970; Winter & Stewart, 1977a, Table 2.4), but also with the broad trend of previous laboratory research using the traditional scoring system applied to TATs.

 Insert Table 6 about here

For example, the power motive predicts perceived presidential impact both in terms of specific "great decisions" and also in terms of historians' overall ratings of greatness. Power motivation also predicts war entry and assassination attempts. (These results may be linked: historians may admire the "warrior hero" president, while the power-motivated "great" presidents, in turn, may arouse strong emotions and so provide an especially strong focus for the wide variety of disturbed impulses of potential assassins; see Clarke, 1982.) Interestingly enough, it is also related to avoiding war in a crisis situation. Thus power motivation overall may be associated with dramatic, crisis-oriented, perhaps confrontational foreign policy, which may end peacefully but which can end in war (see Hermann, 1980b). Consistent with this, power-motivated presidents seem to show the positive, energetic and productive approach to the job that Barber (1977) characterized as the "active-positive" character type.

Affiliation-intimacy-motivated presidents tend to be drawn into political scandals, as reflected by both actual resignations and also the extent of historians' commentary on scandal during their administration. Perhaps they are readily influenced and therefore vulnerable to self-serving "friends" (see Hardy, 1957; Walker & Heyns, 1962). On a more positive note, presidents high in affiliation-intimacy seem to be especially concerned with arms limitation and control. In the subsample of 14 presidents for which this variable was scored, affiliation-intimacy and power are highly correlated ($r = .59$). When the effects of power motive imagery are partialled out, the correlation of affiliation-intimacy with arms limitation

agreements becomes .57 ($p < .05$). In terms of peace and war, arguably the most important issues facing any president, the affiliation-intimacy and power motives thus appear to have opposite effects. (See McClelland, 1975, chap. 9 for other evidence on this point.)

Achievement motivation shows a curious pattern. While the achievement motive generally predicts success in entrepreneurial ventures (see McClelland & Winter, 1969, chapter 1), in presidential politics it does not. What is the explanation for the difference? Achievement-motivated presidents do seem to approach the office with visions and aspirations (historians' ratings of "idealism"). They themselves want to identify and carry out the one best solution to every problem and follow the one best path to every goal--as achievement-motivated Jimmy Carter titled his campaign autobiography, *Why Not the Best?* Where authority is hierarchically concentrated, as in situations of corporate and industrial discipline, this may work.

Politics, however, is more diffuse, decentralized, and conflicted; it requires bargaining, compromise, and aggregating the inchoate interests of diverse groups. As the political columnist Martin Nolan vividly put it, "CEOs accustomed to barking orders are seldom successful at schmoozing their way through the compromise carousel of Washington" (Boston Globe, July 28, 1986, p. 19). In such situations, achievement motivation may only lead to impatient, frustrated perseverance. The relationship between presidential achievement motivation and foreign (nonwar) military intervention suggests one possible outcome: a calculated (low- or moderate-risk) foreign intervention, which will rally the American people and so restore the president's freedom to act. In any case, these presidents show the self-defeating "rigidification" that Barber (1977) described in the "active-negative" character type.³

Case Studies of Individuals: Kennedy and Nixon

Research using data from Richard Nixon and John F. Kennedy illustrate how motive scores obtained at a distance can be used in studies of individuals. In the process of adapting the running text scoring system for direct scoring of television programs, films, and other media, Winter and Healy (1983) scored nineteen major speeches and the first thirteen press conferences of President John F. Kennedy. (With media materials of this type, scores are expressed as proportions of 15-second time units scoring for motive imagery.) The speech scores were analyzed by type of speech (speeches given as a response to domestic or international threat, or non-threat general speeches), and the press conference scores were related to the pattern of presidential activity and time-use on adjacent days.

Winter and Carlson (1988) validated the motive scores from Richard Nixon's 1969 inaugural through a systematic analysis of the course of Nixon's everyday behavior and

³ Barber (1977) argues the reverse: that active-negative presidents are power-driven, while active-positives want to achieve. Since his analysis refers to manifest actions and results, rather than to latent motives, there is no necessary conflict with the present results. In addition, Barber's use of "achievement" and "power" motives is probably different from the scoring definitions.

personality as reported in the memoirs of his aides. After using Nixon's motive profile to explain several paradoxes of Nixon's political career, they concluded that motive scores obtained in this way could be used in psychobiography, as the basis for constructing personality portraits.

John F. Kennedy: Motives, Threat and the Use of Presidential Time

The 19 Kennedy speeches fall naturally into two groups: "threat" speeches, composed and delivered under urgent time pressure as immediate responses to specific threats (racial incidents or crises in Berlin and Cuba), and "nonthreat" or general speeches, such as the State of the Union messages or speeches on specific occasions. On theoretical grounds, we would expect threat speeches to be higher in power motive imagery, since they are the president's attempt to arouse public response to the threat of others' power. As shown in Table 7, these four speeches are in fact significantly higher in power motivation, and significantly lower in achievement imagery.

 Insert Table 7 about here

Kennedy's spontaneous responses to questions at his first 13 domestic press conferences (January 25 through July 19, 1961) were also scored for motive imagery. (Questions and answers after Kennedy's June 2 speech in Paris were not scored, because of the very different setting. Opening statements at press conferences were not scored on the grounds that they were prepared and not spontaneous.) Do these scores reflect Kennedy's changing motive states during the early weeks of his presidency? Using information from the president's official appointment books at the Kennedy Presidential Library, Kennedy's schedules on the two business days before and two business days after each press conference day were then classified and aggregated to give an indication of how Kennedy spent his working time during the period surrounding each press conference. Table 8 presents the relationships between these categories of time use and Kennedy's motive imagery.

 Insert Table 8 about here

If power motivation is related to Barber's "active-positive" category, as shown in Table 5 above, then those times when Kennedy's power motivation is relatively high should be times when he enjoyed the job more and so worked harder and longer. As shown in Table 8, the higher Kennedy's power motive imagery, the earlier he arrived at the Oval Office, the longer was his total working day, and (perhaps as a result) the longer his mid-day breaks (usually for lunch, a nap, or swimming) in the living quarters of the White House. In terms of use of working time, power motivation is associated with time spent giving speeches and formally unscheduled time, both of which are consistent with earlier laboratory research findings about visibility and autonomy, respectively, in power-motivated men (see Winter,

1973, pp. 133, 182-184); and time spent in scheduled visits of friends. Perhaps this latter finding reflects the pleasure that power-motivated people take in cultivating a loyal nucleus of devoted, and nonthreatening, followers (Winter, 1973, p. 114).

Are these findings an artifact of external threats? During times of threat presidential power motivation does seem to be higher (see the analysis of speeches presented in Table 7 above), and it is reasonable to suppose that external threats also lead the president to work longer hours. Two of these first 13 press conferences were held during times of special threat (April 21, shortly after the Bay of Pigs fiasco, and June 28, during the Berlin crisis), and omitting these from the analysis leaves the results reported in Table 8 substantially unchanged. In fact, the correlation of total duration of working day with power imagery rises to .65 ($p < .05$), while that with achievement imagery becomes $-.48$. Thus the way in which Kennedy spent his day seems in part a reflection of his changing motive states rather than an artifact of external threats.

Achievement motivation imagery, on the other hand, is associated with later arrival at work and shorter noontime breaks, perhaps reflecting an efficiency orientation and even a slight distaste for the work of the presidency. (Recall from Table 7 above that achievement-motivated presidents tend to be classified “active-negative.”)

Explaining Richard Nixon

The motive scores for Nixon’s 1969 inaugural address, expressed in both raw and standardized form, are shown in Table 9. His profile is high achievement and affiliation-intimacy, and average power motivation. Although individual motive scores change over time, the basic shape of this profile is reasonably stable, as shown by the scores for the 1973 second inaugural at the bottom of the table.

 Insert Table 9 about here

Is this motive profile valid? Perhaps the first way to answer this question is to determine whether it enables us to make forecasts about Nixon’s presidency. Obviously we now know a lot about what Nixon did and did not do as president; so the real question is whether knowledge of Nixon’s motive profile, at the time of his first inauguration in 1969, would have enabled us to make accurate predictions about his presidency, based on the relationships between presidential motives and presidential behaviors observed in previous research. Did Nixon act like a president high in achievement and affiliation-intimacy motivation and average in power motivation?

Validation with presidential outcomes. In January 1969, we could have made four specific and objective predictions about Nixon’s presidency, based solely on the research relating presidential motive scores to presidential behaviors and outcomes (summarized in Table 6) and Nixon’s scores (as shown in Table 9). These predictions are exhaustive, in the

sense that they take account of all correlates of presidential motives that can be objectively measured. Each prediction, along with the relevant evidence from Nixon's presidency, is discussed below.

(1) *The United States would be unlikely to enter a war.* Among other presidents (and other political leaders), power motivation is associated with war entry and aggressiveness. Nixon's power motivation, however, is only average. This prediction is confirmed: although Nixon continued and even expanded the Vietnam War, which he "inherited," the United States did not enter a new war during his presidency.

(2) *The United States would be likely to conclude a major-power treaty for the limitation or abolition of a major weapons system.* Ever since the Hague Conferences of the Nineteenth Century, arms treaties have been concluded by presidents who score low in power and high in affiliation-intimacy. Such an outcome seems consistent with their desire for close, cooperative relations, in a climate of "safety," as well as their disinterest in power. Nixon's affiliation-intimacy score is very high, and his power motivation only average. The prediction is amply confirmed by the Seabed Treaty, the first SALT Treaty, and even by the ratification (almost fifty years late) of the 1925 Geneva Protocol against the use of poison gas.

(3) *There would be a scandal of sufficient magnitude to cause resignations of staff and/or cabinet members.* Perhaps the vulnerability to scandal of affiliation-motivated presidents arises from an excessive desire to go along with (or not to antagonize) self-seeking advisors, or perhaps such presidents are easily influenced. In any case, Nixon's affiliation-intimacy score is very high, and his power motivation (which seems to act as a slight "protection" against scandals, or at least the discovery of scandals) is only average. In terms of resignations, Watergate was of course the biggest political scandal in American history to date.

(4) *There would be no assassination attempt on the president.* Leaders high in power motivation seem to attract assassins, perhaps as a result of their very charisma or ability to arouse strong emotions in other people. Nixon's power motive score is only average, but the prediction is a little more difficult to evaluate. Given that there are usually some threats to the president, what constitutes a "serious" assassination attempt? According to John Ehrlichman, for example, both "Billy Graham and Jeane Dixon, among others, [warned] that his life was in danger" (1982, p. 363). Yet no one fired shots at Nixon, as they did at Kennedy, Ford, and Reagan. Still, Clarke (1982) recounts the curious story of Samuel Byck, who in February 1974 was shot and killed after trying to hijack an airliner at the Baltimore airport. According to a tape recording that Byck left behind, he planned to crash the plane into the White House (Clarke, 1982, p. 134). On a conservative analysis, we would have to say that this prediction was not confirmed.

(5) One final prediction, involving Barber's (1977) typology, is more subjective, although Barber spells out the criteria for classifying presidential character in considerable detail. Given the relationships reported in Table 6, we would have predicted that *as*

president, Nixon would be classified as “active-negative” and not “active-positive.” Power-motivated presidents are active and love their work. Achievement-motivated presidents, in contrast, seem to become frustrated and impatient. It almost seems as though they would be happier running a company than a country. Nixon’s achievement imagery score is very high, and his power imagery only average. These paired predictions are confirmed by Barber’s own classification: Nixon showed the pattern of rigidification in a losing cause that is characteristic of the active-negative type--someone who is energetic and active in the job, but who has a fundamental dislike for the game of politics (1977, pp. 345-442, 457-484).

Overall, then, four out of five reasonably objective predictions that could have been made in 1969, solely on the basis of Nixon’s motive profile, would have been confirmed during the years of his presidency, and the fifth has a somewhat unclear status. (Discriminant function analysis, based on all presidents except Nixon gives the same results as the more informal presentation presented above.) In terms of its use in forecasting the Nixon presidency, then, the motive profile shown in Table 9 has predictive validity.

Validation in everyday life. Winter and Carlson (1988) validated Nixon’s 1969 motive profile in the texture and style of Nixon’s everyday behavior, gathering evidence for motive-related behavior from the published record of Nixon’s own autobiography, other major biographies, and the memoirs of his principal associates who worked most closely with him and presumably knew him best. First, they compiled a systematic inventory of all reasonably well-established behavior correlates of each motive, drawn from textbooks, handbook chapters, review articles, and the *Psychological Abstracts*. Then, they selected books by and about Richard Nixon, including his own two autobiographies (Nixon, 1962, 1978/1979) and the memoirs of six former close aides who worked directly with Nixon during his presidency and who continued to have positive or at least neutral feelings about him.

Each memoir was carefully searched for facts or comments about *enduring* behavior patterns of Richard Nixon. The resulting lists of Nixon behaviors were then matched with the list of motive correlates. Overall, 65 specific behaviors or background characteristics associated with the three motives were studied. Table 10 brings together the results of this analysis, summarizing the record of confirmations, disconfirmations, and inconclusive evidence for these 65 behavior correlates. Eliminating cases where the evidence was inconclusive, Nixon shows almost all of the correlates of the achievement and affiliation-intimacy motives, as would be expected by his very high scores on those two motives. He shows only some correlates (slightly fewer than half) of the power motive, as would be expected by his average score.

 Insert Table 10 about here

To express this impressionistic summary in more quantitative and testable terms, it is first necessary to frame the appropriate hypotheses. By chance alone, what proportion of

motive-related behavior correlates should apply to any given person? The right-hand columns of Table 10 show two approaches. First, we could assume, as a kind of “null hypothesis,” that about half or 50% of the behavior correlates of any given motive should apply to any given person. Using the Chi-square test as a measure of goodness of fit between the proportions expected and obtained (as shown in the “50:50 confirmation” column), there are two significant departures from 50 percent: that is, Nixon showed significantly more than 50 percent of both the achievement and the affiliation-intimacy correlates, but not more than 50 percent of the power correlates. In other words, Nixon displays more achievement and affiliation-intimacy behaviors (but not power behaviors) than average, just as he scored higher than average in achievement and affiliation-intimacy motive imagery (but not power imagery).

An alternative hypothesis is that the percent of motive-related behavior correlates displayed by any given person will be roughly equal to that person’s motive score expressed in percentile terms. Using the Chi-square test again as a measure of goodness of fit between expected and obtained proportions, we find in this case no significant departures from expectation, as shown in the right-hand column of Table 10 (“percentile confirmation”). Because of the different kinds of “expectations” used in each analysis, one approach shows departures from expectation and the other shows no departures from expectations; but in fact, both analyses are consistent with each other. In other words, for each motive Nixon’s behavior (percent of motive correlates shown) fits what would be expected on the basis of his percentile score for that motive.

Using either null hypothesis, then, the results support the validity of the Nixon motive profile in the sense of predicting Nixon’s patterns of behavior. More broadly, the Kennedy and Nixon studies suggest that motive scores obtained at a distance with the integrated running text scoring system, can contribute to the intensive study of the individual case, as well as to more traditional personality research with groups.

Using Motive Imagery Scores to Study Social Processes

All of the findings presented so far involve the study of personality, either of groups or individuals. Three recent studies by Winter (1987a, 1987b, 1987c) illustrate further applications of the integrated running text scoring system to the study of social psychological processes. By comparing the motive profiles of American presidents (scored from their first inaugural addresses) to those of American society at the time of their election (scored from cultural documents; see McClelland, 1975), Winter (1987b) showed that the greater the congruence or fit between the two, the greater the president’s electoral success (percentage of the vote, margin of victory, and being re-elected). Winter interpreted these results as supporting Erikson’s theory that great leaders, with their own identities, conflicts, and needs, are “found and chosen by contemporaries possessed of analogous conflicts and corresponding needs” (1964, p. 204). Leader appeal, in other words, is a function of leader-follower personality match.

Leader *performance*, on the other hand, seems to be a function only of a president's motives themselves, rather than president-society congruence. As shown in Table 6 above, power-motivated presidents are rated more highly by historians, make more "great" decisions, enter more wars, and avoid more wars. None of these outcomes is associated with the degree of president-society motive profile congruence.

Motive imagery scores have also been useful in identifying a dynamic of conflict escalation. Winter (1987a) examined two different conflict situations: presidential campaign debates in 1960 and 1980, and the secession crisis and outbreak of the American Civil War in 1861. Key speeches and statements by leaders on each side, as reported by favorable and opposed media sources, were scored for motive imagery. In both cases, there was a significant tendency for media reports to accentuate the power motivation of the opposed side and to diminish the power motivation of the favored side. In a subsequent study of the British-German relations during July-August, 1914, at the outbreak of World War I, Winter (1987c) found that as the crisis grew more severe, the same tendency developed both in British and German newspaper reports and in diplomatic dispatches on both sides. Among political leaders and the general public, this reciprocal distortion of power motive imagery probably increases the sense of external threat, thereby making it easier to rationalize aggression as a reasonable response. Thus motive distortion may be one factor that can escalate conflicts to war.

Summary

This paper has presented a variety of psychometric and validity data intended to establish the usefulness of a new integrated way of measuring achievement, affiliation-intimacy, and power motive imagery in verbal text, including both traditional TAT stories and also speeches, interviews, literary works, and so forth. The new system is objective and can be learned by previously inexperienced scorers. It is easier and cheaper to use in research. It converges substantially (though not completely) with the original motive scoring systems. It has considerable predictive validity, especially when applied to "ordinary life" verbal materials, as in at-a-distance studies of historical figures and political leaders.

The fact that the new system can be applied to any kind of verbal material should enhance its research utility in several ways. First, it makes scoring of traditional TAT stories cheaper and easier. In addition, it expands the range of people and topics that can be studied with measures of human motivation.

Finally, the reliability and validity of the integrated running text scoring system when applied to naturally-occurring verbal material suggests a broader conclusion about assessment methodology. Even for people who could be tested in traditional ways, assessments of motives and other personality variables might be improved by scoring naturally-occurring verbal material instead of administering TATs or other "instruments" in contrived and constrained "testing" or "assessment" situations. Given the problems of traditional testing, many researchers are turning to sampling of behavior in natural settings as an alternative. The research reported in this chapter suggests that scoring verbal material for motives can be

thought of as a technique of “thought sampling” that has the same advantages but is even easier and cheaper to carry out. Perhaps it is an irony of research that techniques developed to study people who could not be tested in traditional ways may turn out to have psychometric and validity advantages for studying people who could be directly tested.

In any case, the new integrated running text scoring system is a highly generalized content analysis procedure that can be applied in a variety of research contexts to study human motivation, with confidence about the validity and reliability of what is being measured.

References

- Atkinson, J. W. (Ed.). (1958). *Motives in fantasy, action and society*. Princeton, NJ: Van Nostrand.
- Atkinson, J. W. (1982). Motivational determinants of thematic apperception. In A. J. Stewart (Ed.), *Motivation and society* (pp. 3-40). San Francisco: Jossey-Bass.
- Atkinson, J. W., Heyns, R. W., & Veroff, J. (1954). The effect of experimental arousal of the affiliation motive on thematic apperception. *Journal of Abnormal and Social Psychology, 49*, 405-410.
- Barber, J. D. (1977). *The presidential character: Predicting performance in the White House* (rev. ed.). Englewood Cliffs, NJ: Prentice Hall.
- Berlew, D. E. (1956). *The achievement motive and the growth of Greek civilization*. Unpublished honors thesis, Wesleyan University.
- Bradburn, N. M., & Berlew, D. E. (1961). Need for achievement and English economic growth. *Economic Development and Cultural Change, 10*, 8-20.
- Boyatzis, R. E. (1973). Affiliation motivation. In D. C. McClelland & R. S. Steele (Eds.), *Human motivation: A book of readings* (pp. 252-276). Morristown, NJ: General Learning Press.
- Child, I. L., Storm, T., & Veroff, J. (1958). Achievement themes in folk tales related to socialization practice. In J. W. Atkinson (Ed), *Motives in fantasy, action and society* (pp. 479-492). Princeton, NJ: Van Nostrand.
- Clarke, J. W. (1982). *American assassins: The darker side of politics*. Princeton, NJ: Princeton University Press.
- Cortes, J. B. (1960). The achievement motive in the Spanish economy between the 13th and 18th centuries. *Economic Development and Cultural Change, 9*, 144-163.
- deCharms, R., & Moeller, G. H. (1962). Values expressed in American children's readers: 1800-1950. *Journal of Abnormal and Social Psychology, 64*, 136-142.
- Donley, R. E. (1968). *Psychological motives and the American presidency*. Unpublished honors thesis, Wesleyan University.
- Donley, R. E., & Winter, D. G. (1970). Measuring the motives of public officials at a distance: An exploratory study of American presidents. *Behavioral Science, 15*, 227-236.
- Ehrlichman, J. (1982). *Witness to power: The Nixon years*. New York: Simon & Schuster.
- Emerson, J. T. (1973, April 11 and 12). Chronological list of 199 U. S. military hostilities abroad without a declaration of war, 1789-1972. Exhibit II in J. T. Emerson, War powers: An invasion of presidential prerogative. *Hearings before the Committee on Foreign Relations, United States Senate, Ninety-Third Congress*, pp. 120-151. Washington: Government Printing Office.
- Erikson, E. H. (1964). *Insight and responsibility*. New York: Norton.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327-358.
- Giliberto, S. M. (1972). *Motivation and the Methodist revival*. Unpublished honors thesis, Harvard University.
- Hardy, K. (1957). Determinants of conformity and attitude change. *Journal of Abnormal and Social Psychology, 54*, 289-294.

Henley, N. M. (1967). Achievement and affiliation imagery in American fiction, 1901-1961. *Journal of Personality and Social Psychology*, 7, 208-210.

Hermann, M. G. (1968). *Indirect methods of assessing personality*. Research Bulletin report. Princeton, N. J.: Educational Testing Service

Hermann, M. G. (Ed.). (1977). *A psychological examination of political leaders*. New York: Free Press.

Hermann, M. G. (1979). Who becomes a political leader? Some societal and regime influences on selection of a head of state. In L. S. Falkowski (Ed.), *Psychological models in international politics* (pp. 15-48). Boulder, CO: Westview Press.

Hermann, M. G. (1980a). Assessing the personalities of Soviet Politburo members. *Personality and Social Psychology Bulletin*, 6, 332-352.

Hermann, M. G. (1980b). Explaining foreign policy behavior using personal characteristics of political leaders. *International Studies Quarterly*, 24, 7-46.

Historical Figures Assessment Collaborative. (1977). Assessing historical figures: The use of observer-based personality descriptions. *Historical Methods Newsletter*, 10(2), 66-76.

LeVine, R. (1966). *Dreams and deeds: Achievement motivation in Nigeria*. Chicago: University of Chicago Press.

Lundy, A. (1985). The reliability of the Thematic Apperception Test. *Journal of Research in Personality*, 49, 141-145.

McAdams, D. P. (1980). A thematic coding system for the intimacy motive. *Journal of Research in Personality*, 14, 413-432.

McAdams, D. P. (1982). Intimacy motivation. In A. J. Stewart (Ed.), *Motivation and society* (pp. 133-171). San Francisco: Jossey-Bass.

McClelland, D. C. (1958). The use of measures of human motivation in the study of society. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society* (pp. 518-552). Princeton, NJ: Van Nostrand.

McClelland, D. C. (1961). *The achieving society*. Princeton, NJ: Van Nostrand.

McClelland, D. C. (1975). *Power: The inner experience*. New York: Irvington.

McClelland, D. C. (1981). Is personality consistent? In A. I. Rubin, J. Aronoff, A. M. Barclay, & R. A. Zucker (Eds.), *Further explorations in personality*. New York: Wiley.

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

McClelland, D. C., Clark, R. A., Roby, T. B., & Atkinson, J. W. (1949). The effect of the need for achievement on thematic apperception. *Journal of Experimental Psychology*, 37, 242-255.

McClelland, D. C., & Winter, D. G. (1969). *Motivating economic achievement*. New York: Free Press.

Maranell, G. M. (1970). The evaluation of presidents: An extension of the Schlesinger poll. *Journal of American History*, 57, 104-113.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Morris, R. B. (1967). *Great presidential decisions: State papers that changed the course of history* (rev. ed.). Philadelphia: Lippincott.

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

- Nixon, R. M. (1962). *Six Crises*. Garden City, NY: Doubleday.
- Nixon, R. M. (1979). *RN: The memoirs of Richard Nixon*. New York: WarnerBooks. (Original work published 1978)
- Nolan, M. (1986, July 28). Column. *Boston Globe*, p. 19.
- Porter, C. A., & Suedfeld, P. (1981). Integrative complexity in the correspondence of literary figures: Effects of personal and societal stress. *Journal of Personality and Social Psychology*, *40*, 321-330.
- Richardson, L. F. (1960). *Statistics of deadly quarrels*. Pittsburgh and Chicago: Quadrangle Books.
- Simonton, D. K. (1986). Presidential personality: Biographical use of the Gough Adjective Check List. *Journal of Personality and Social Psychology*, *51*, 149-160.
- Small, M. (1980). *Is war inevitable*. Beverly Hills, CA: Sage.
- Smith, C., & Feld, S. (1958). How to learn the method of content analysis for n achievement, n affiliation, and n power. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society* (pp. 685-818). Princeton: Van Nostrand.
- Straus, M. A., & Houghton, L. J. (1960). Achievement, affiliation, and cooperation values as clues to trends in American rural society, 1924-1958. *Rural Sociology*, *25*, 394-403.
- Suedfeld, P., & Piedrahita (1984). Intimations of mortality: Integrative simplification as a precursor of death. *Journal of Personality and Social Psychology*, *47*, 848-852.
- Suedfeld, P., & Rank, A. D. (1976). Revolutionary leaders: Long-term success as a function of changes in conceptual complexity. *Journal of Personality and Social Psychology*, *34*, 169-178.
- Veroff, J. (1957). Development and validation of a projective measure of power motivation. *Journal of Abnormal and Social Psychology*, *54*, 1-8.
- Walker, E. L., & Heyns, R. W. (1962). *An anatomy for conformity*. Englewood Cliffs, NJ: Prentice-Hall.
- Winter, D. G. (1973). *The power motive*. New York: Free Press.
- Winter, D. G. (1976, July). Why the candidates run. *Psychology Today*, pp.45-49, 92.
- Winter, D. G. (1978). *Navy leadership and management competencies: Convergence among interviews, test scores, and supervisors' ratings*. Unpublished paper, Wesleyan University and McBer & Company (Boston).
- Winter, D. G. (1979, September). *An introduction to Leadership and Management Effectiveness Training (LMET) theory and practice*. Paper presented at the American Psychological Association convention, New York.
- Winter, D. G. (1980). Measuring the motives of southern Africa political leaders at a distance. *Political Psychology*, *2*(2), 75-85.
- Winter, D. G. (1982). Motives and behavior in the 1976 presidential candidates. In A. J. Stewart (Ed.), *Motivation and society* (pp. 244-273). San Francisco: Jossey-Bass.
- Winter, D. G. (1987a). Enhancement of an enemy's power motivation as a dynamic of conflict escalation. *Journal of Personality and Social Psychology*, *52*, 41-46.
- Winter, D. G. (1987b). Leader appeal, leader performance, and the motive profiles of leaders and followers: A study of American presidents and elections. *Journal of Personality and Social Psychology*, *52*, 196-202.

Winter, D. G. (1987c, April). *Power motive distortion in British and German newspapers and diplomatic dispatches at the outbreak of World War I*. Paper presented at the meeting of the Eastern Psychological Association, Arlington, Virginia.

Winter, D., G., & Carlson, L. (1988). Using motive scores in the psychobiographical study of an individual: The case of Richard Nixon. *Journal of Personality*, 56, 75-103.

Winter, D. G., & Healy, J. M., Jr. (1983). *Coding the motive imagery content of television programs*. Unpublished paper, Wesleyan University.

Winter, D. G., McClelland, D. C., & Stewart, A. J. (1981). *A new case for the liberal arts: Assessing institutional goals and student development*. San Francisco: Jossey-Bass.

Winter, D. G., & Stewart, A. J. (1977a). Content analysis as a technique for assessing political leaders. In M. G. Hermann (Ed.), *A psychological examination of political leaders* (pp. 27-61). New York: Free Press.

Winter, D. G., & Stewart, A. J. (1977b). Power motive reliability as a function of retest instructions. *Journal of Consulting and Clinical Psychology*, 45, 436-440.

Winter, D. G., & Stewart, A. J. (1978). The power motive. In H. London & J. Exner (Ed.), *Dimensions of personality* (pp. 391-447). New York: Wiley.

Woodward, C. V. (1974). *Responses of the president to charges of misconduct*. New York: Dell.

Table 1

Brief Outline of the Integrated Running Text System for Scoring Motive Imagery

Achievement Imagery

- Reference to a standard of excellence, either directly by adjectives that evaluate performance or quality, or indirectly by actions which clearly suggest a concern with excellence.
- Negative affect or counteraction in reaction to failure or a lack of excellence.
- Success in competition with others.
- Unique accomplishment.

Affiliation-Intimacy Imagery

- Expression of warm, positive, friendly feelings toward others.
- Negative affect about separation or disruption of a friendly relationship.
- Affiliative, companionate activities.
- Friendly nurturant acts.

Power Imagery

- Strong vigorous actions which have impact on others: force, attempts to convince or persuade, unsolicited help, attempts to monitor or control.
 - Actions which directly arouse a strong positive or negative emotional state in others.
 - Concern for reputation or prestige.
-

Table 2

Convergence between Original and Integrated Running Text Systems

Data set	Correlations of motive imagery scores from the two systems for:		
	Achievement	Affiliation	Power
By story: Practice sets (90 stories for each motive)	.54 ⁴ ***	.72***	.64***
By person: "Ivy College" students (Winter, McClelland, & Stewart, 1981; N=234, five stories each)	.48***	.45***	.55***

*** $p < .001$

⁴ $r = .66$ if ten stories with Long-Term Involvement, which is scored in the original system but not in the integrated running text system, are omitted.

Table 3

Performance of the Integrated Running Text Scoring System in Differentiating Aroused and Neutral Stories from Motive Arousal Experiments

Arousal experiment	Running text motive imagery						A-N diff.	Sig.
	Aroused group			Neutral group				
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>		
<i>Achievement motive</i> ⁵								
Failure vs. relaxed	39	8.80	6.08				4.08	<.01
Success-failure vs. relaxed	39	9.47	5.78	39	4.72	4.56	4.75	<.001
Failure and success-failure vs. relaxed	78	9.14	5.93				4.42	<.001
<i>Affiliation motive</i> ⁶								
Sociometric rating vs. classroom	29	11.32	5.51	36	11.42	5.54	-.10	
<i>Intimacy motive</i> ⁷								
Fraternity/sorority initiation vs. classroom	41	16.49	6.30	42	13.00	6.44	3.49	<.02
<i>Power motive</i> ⁸								
Candidates awaiting election results vs. classroom	34	8.65	4.84	34	7.67	4.76	.98	
<i>Power motive</i> ⁹								
Viewers of Kennedy film vs. neutral film	47	11.14	5.22	44	8.37	4.59	2.77	<.005

Note. All scores are expressed in terms of images per 1,000 words.

⁵ from McClelland et al. (1949)

⁶ from Atkinson et al. (1954)

⁷ from McAdams (1980), study 1

⁸ from Veroff (1957)

⁹ from Winter (1973)

Table 4

Test-retest Reliability of Integrated Running Text Scores

Retest instruction condition ¹⁰	Test-retest correlations for:		
	Achievement	Affiliation- Intimacy	Power
Different ($N = 24$)	-.18	.35	.11
No instructions ($N=23$)	.01	.32	.39
Same ($N=23$)	.71***	.62**	.63**
Pooled “same” and “no instructions” ($N=46$)	.46**	.41**	.60**

** $p < .01$ *** $p < .001$

¹⁰ See Winter and Stewart (1977b) for exact definitions and instructions used in each condition.

Table 5

Relation of Motive Imagery Scores and Judges' Ratings to Theoretically Relevant Behavioral Dimensions among Southern Africa Leaders

Judges' ratings of:	Correlation with:	
	Power motive imagery scored from news conferences	Power motive as rated by judges
War disposition	.71***	.14
Activity	.51*	.10

* $p < .05$ *** $p < .001$

Table 6

Predicting Presidential Behavior from Motive Imagery in Inaugural Addresses

Presidential behavior/outcome variable	Correlation with:		
	Ach	Aff-Int	Pow
Composite rating of "greatness" (Maranell, 1970; $N=29$)	.09	.09	.40*
Rating of "idealism" (Maranell, 1970; $N=29$)	.51**	.19	.19
Great decision(s) cited? (Morris, 1967; $N=29$)	.09	.30	.51**
Assassination attempt? (Clarke, 1982; plus Reagan; $N=33$)	.09	.17	.40*
War entry ($N=31$)	-.03	.16	.52**
War avoidance (Small, 1980; $N=29$)	-.11	.22	.34
Number of nonwar military interventions (Emerson, 1973; $N=28$)	.44*	-.18	.08
Arms limitation treaty ($N=14$)	.13	.43	-.05
Scandal: Resignation of cabinet or high staff (Woodward, 1974, plus Nixon; $N=30$)	.15	.40*	.01
Sandal: Number of pages per year in office (Woodward, 1974; $N=29$)	.15	.40*	.01
Active-positive type (Barber, 1977; $N=11$)	-.07	.37	.86***
Active-negative type (Barber, 1977; $N=11$)	.84***	.03	-.32

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 7

Motive Imagery Differences between “Threat” and “Nonthreat” Presidential Speeches by John F. Kennedy.

Type of speech	Proportion of time-units scored for:					
	Achievement		Affiliation-intimacy		Power	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Threat speeches ¹¹ (<i>N</i> =4)	10.00	4.06	5.25	3.56	48.25	12.83
Nonthreat speeches ¹² (<i>N</i> =15)	23.40	9.41	10.33	7.01	30.67	8.96
Difference	-13.40		-5.08		17.58	
	<i>t</i> =2.99		<i>t</i> =1.33		<i>t</i> =2.99	
	<i>p</i> >.02		<i>p</i> =ns		<i>p</i> <.01	

¹¹ Report on the Berlin crisis (6/25/61), Cuban Missile speech (10/22/62), report on situation at the University of Mississippi (9/30/62), report on civil rights and situation at the University of Alabama (6/11/63)

¹² Inaugural address (1/20/61), State of the Union messages (1/30/61, 1/11/62, 1/14/63), address to Latin American diplomatic corps (3/13/61), message on urgent national needs (5/25/61), report on European trip (6/6/61), United Nations addresses (9/25/61, 9/20/63), address on nuclear testing and disarmament (3/2/62), commencement address at Yale University (6/11/62), report on the national economy (8/13/62), commencement address at American University (6/10/63), address on the test ban treaty (7/26/63), and address on the test ban treaty and tax bill (9/18/63)

Table 8

Relationships between Presidential Motive Imagery in Press Conferences and Use of Presidential Time on Adjacent Days

Presidential time use	Correlation with proportion of time-units in news conference scored for:		
	Ach	Aff-Int	Pow
Early arrival in oval office	-.48+	.27	.70**
Time during mid-day breaks spent in White House living quarters	-.63*	-.05	.48+
Total duration of working day	-.09	.05	.47+
Use of time during working day:			
Scheduled working appointments	-.39	.09	.03
Ceremonial events	.29	-.21	.19
Delivering speeches or "remarks"	.24	.05	.50+
Scheduled appointments with friends	-.30	-.55+	.72**
Unscheduled time in office	-.37	.07	.62*

+ $p < .10$ * $p < .05$ ** $p < .01$

Table 9

Motive Profile of Richard Nixon based on Inaugural Addresses.

Year	Raw scores: Images/1000 words			Standardized scores ¹³		
	Ach	Aff-Int	Pow	Ach	Aff-Int	Pow
1969	8.94	8.00	7.06	66	76	53
1973	7.75	4.43	6.64	60	57	51

¹³ Standardized on sample of first inaugural addresses of 34 presidents (Washington through Reagan); overall $M=50$ and $SD=10$.

Table 10

Actual versus Hypothesized Motive-Related Behaviors of Richard Nixon.

Motive	Behavior correlates:				Hypothesized % of correlates confirmed and goodness of fit, ¹⁴ under assumption of:	
	Con- firmed	Not con- firmed	Incon- clusive	% con- firmed	50:50 confirmation ¹⁵	Percentile confirmation ¹⁶
Achievement	23	3	2	88	50 ($p < .01$)	95 ($p = ns$)
Affiliation- Intimacy	13	0	1	100	50 ($p < .01$)	99 ($p = ns$)
Power	8	10	3	44	50 ($p = ns$)	62 ($p = ns$)

Note. Table from Winter and Carlson (1988).

¹⁴ Probability of difference between proportions of behavior correlates actually confirmed and proportions expected to be confirmed, using the Chi-square test.

¹⁵ Assuming 50% confirmed and 50% not confirmed.

¹⁶ Assuming that percent of behavior correlates confirmed is equal to the person's percentile score on the relevant motive. (As shown in the next-to-right-hand column, Nixon's percentile scores calculated from the standardized scores of Table 9 are 95 for achievement, 99 for affiliation-intimacy, and 62 for power.)