A META-ANALYSIS OF FIEDLER'S CONTINGENCY MODEL OF LEADERSHIP EFFECTIVENESS

by

ELLEN PATRICIA CREHAN

B.P.E., The University of British Columbia, 1954
M.ED., The University of British Columbia, 1979

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF EDUCATION

in

THE FACULTY OF GRADUATE STUDIES

Department of Administrative, Adult and Higher Education
(Educational Administration)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

March 1984

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Administrative, Adult, and Higher Education (Educational Administration)
The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date March 26, 1984

ABSTRACT

Fiedler's Contingency Model of Leadership Effectiveness is widely cited, yet highly controversial. The present study subjected Fiedler-based studies to a meta-analysis to determine whether a body of consistent findings would emerge. If such findings did emerge, the aim would be to develop a theoretical framework which would adequately explain them. If they did not, the study would try to explain their absence. The original pool of 402 documents contained 112 primary studies. Of these studies, 38 met the criteria of quality, comparability, and extent of adherence to Fiedler's Model needed for retention in the meta-analysis. These 38 studies contained 249 LPC-performance correlations. These correlations were first analyzed for frequency of support for Fiedler's Model using two indicators: (1) the primary study authors' stated conclusions and (2) directional congruency with Fiedler's predictions. Full or partial support was concluded by 15% and 63%, respectively, of the authors. Directional congruency yielded 54% support and 46% non-support. The second phase of the analysis, in which the within-octant correlations were partitioned by eight different variables, examined both the magnitude and direction using median correlations as the common metric. The extensive disparity

between the obtained medians and those predicted by Fiedler led to a third examination using Exploratory Data Analysis techniques. This analysis showed that, regardless of the variable used, there were within-octant differences which prevented not only combining the partitioned data sets but also continuing the meta-analysis. An analysis of the methodological variability in the studies yielded findings important to the continued use of Fiedler's Model. These findings led to several recommendations intended to standardize testing procedures. These recommendations included suggestions regarding greater standardization of score division methods for LPC and GAS, assessment of the three situational variables, and criteria by which to assess support for the Model. Until such standardization is achieved, the validity of Fiedler's Model can be neither confirmed nor disconfirmed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The completion of this dissertation would have been impossible without the advice, support, and assistance of many people. The financial support provided by the Social Sciences and Humanities Research Council of Canada and The University of British Columbia Graduate Fellowships is gratefully acknowledged.

I extend my deepest thanks and gratitude to the members of my Supervisory Committee: Dr. Jean Hills, whose thought-provoking comments gave rise to this study; Dr. Graham Kelsey, whose analytical and editorial skills shaped the final draft of the thesis; Dr. Todd Rogers, whose expertise in statistics and methodology enriched my knowledge; and Dr. Ian Housego, whose insightful observations broadened my thinking beyond the confines of the study itself. In addition to the Supervisory Committee, I also thank the members of the Examining Committee: Dr. Peter Frost and Dr. Bill Borgen, the University Examiners; Dr. Cecil Miskel, the External Examiner who participated in the final defense; and Dr. Doug McKie, who served as Chairman for the defense.

Apart from the immeasurable contribution of my Committee, there are six people who, in a variety of ways, helped me through my years as a graduate student. I express

appreciation to Dr. Harold Ratzlaff, who made me believe I could handle statistics; to Bruce McGillivray, who helped carry out the computer analysis and word processing; to my fellow-student, Carol Gibson, who so generously gave of her time to listen and to react to my ideas; to my cousin, Gertrude Langridge, who provided both moral and material support throughout my graduate career; to my friend Daphne Clarke, who expressed faith in my ability to succeed; and to my friend and critic Fran Takemoto, who shared with me the joys and frustrations of graduate study. So great has been the support and encouragement of both the Committee members and these special individuals, that whatever merit this study may have belongs as much to them as it does to me. To all of them I owe a debt of gratitude. Thankyou.

Chapter 1

BACKGROUND, PURPOSE, AND DERIVATION OF THE STUDY

The phenomenon of leadership has long captured the imagination and provoked the curiosity of both theorists and researchers. Theorists have used a variety of approaches in their attempts to describe what leadership involves and, to some extent, how it comes to be. Using these theoretical frameworks, researchers have tried to provide empirical evidence to support or refute the particular theory under investigation. As the research moved from "Great Man" and "Trait" approaches to "Situational" and "Behavioural" theories, it became increasingly apparent that leadership is a product of at least three components: the leader, the led, and the situation in which they all function. The complexity of that interaction is addressed by the contingency theories of leadership.

Hendrix (1976) has identified eight such models, each of which states that leadership effectiveness is dependent upon a different variable. For example, Hersey and Blanchard (1969) claim that the maturity level of the followers is the variable on which leadership effectiveness is contingent while House (1971) contends that it is worker motivation. Focussing on the leader rather than subordinates, Katz and Kahn (1966) identify hierarchical

level as the independent variable while Stogdill (1959, 1971) states it to be the role and role set of the leader. Dubin (1965) and Woodward (1965) take yet a different point of view, claiming that leadership depends upon the level of production technology. Each of these contingency theories suggests that leadership effectiveness depends upon a single variable. In this respect, the theory put forward by Fiedler (1967) is more complex in that it is the relationship between two variables which is contingent upon a third variable.

Following its first appearance in the leadership literature in 1964, Fiedler's Contingency Model of Leadership Effectiveness initiated a research tradition which has prevailed for two decades. Many of the findings yielded by this research are not only inconsistent across studies but are also contradictory to the predictions of the Model itself. These contradictions are central to the ongoing debate which surrounds Fiedler's Model. It is not the intent of this thesis to lend support to one side or the other of the controversy. Rather this thesis reports an attempt to examine systematically the findings which have prompted the debate in order to determine which of them are consistent, valid, and reliable across studies.

THE PURPOSE OF THE STUDY AND AN OVERVIEW OF THE THESIS

The purpose of this study was three-fold:

1. To perform a meta-analysis of the findings in Fiedler-based studies in order to try to establish which, if any, findings are consistent, reliable, and valid;

2. If consistent, reliable, and valid findings are discovered, then to try to formulate a theoretical framework which can include them and from which hypotheses may be generated for the testing of the theoretical framework;

3. If consistent, reliable, and valid findings do not emerge, then to try to account for their absence in order that further work with Fiedler's Model may avoid existing but perhaps unrecognized pitfalls.

In order to provide the necessary background, the balance of this chapter will present an explanation of Fiedler's Contingency Model. Chapter 2 describes, in some detail, the ways in which the Model is controversial. Chapter 3 explains meta-analysis as a method of inquiry and reports the outcomes of three previous integrative studies. Chapter 4 describes the procedures used in the collection and coding of studies and in the reduction of the data base. Chapter 5 discusses the comparability of the studies in the data base, explains the reasons for further data reduction, and reports on the reliability of study coding. Chapter 6 reports the results of a descriptive analysis which used two different criteria to assess the frequency of support for

Fiedler's Model. Chapter 7 reports the outcomes of an analysis which used median correlations to ascertain the extent of support for the Contingency Model. Chapter 8 reports the results of a further comparability analysis made necessary by the findings reported in the two previous chapters. Chapter 9 reports the results of an analysis of the methodological variability found to exist across the studies included in the data base. Chapter 10 summarizes the findings of the three major analyses (Chapters 7, 8, and 9), presents conclusions and offers some suggestions for the further use of Fiedler's Contingency Model.

FIEDLER'S CONTINGENCY MODEL

Fiedler's Contingency Model of Leadership Effectiveness postulates that effective leadership is contingent upon the matching of leadership style with situational favourableness[1]. "Leadership style", regarded by Fiedler as the key variable and main predictor, is measured by an instrument called the Least Preferred Coworker (LPC) Scale. "Situational favourableness" is

---

[1] Fiedler has recently changed the phrase "situational favourableness" to "situational control" apparently because the former description has been confused with task difficulty (1978:62). However, the original terminology dominates most of the relevant research and will therefore be used throughout this thesis.

conceptualized as the "degree to which the situation provides the leader with control and influence" (Fiedler, 1978:62). It is a composite of three disproportionately weighted variables: leader-member relations, task structure, and position power. To the extent that each is scored high or low on the appropriate instrument, the situation is seen as favourable or unfavourable for the leader. Based on the post hoc analysis of correlational studies, Fiedler dichotomized the three situational variables, used the combinations of values on those variables to construct a single, eight value index of situational favourableness, and matched those values with the correlations between LPC and performance.

Historical Background of LPC

The LPC scale evolved from Fiedler's original research interest "in the operational measurement of interpersonal relations" (1967:37). It should be noted Fiedler's early work was not in the area of leadership but rather in the therapeutic relationship between clinician and patient. One of the measures used to operationalize that relationship "happened to be a method which required the therapist to predict the self-concept of the patient in psycho-therapy" (1967:37). Although this measure showed the therapists' predictions to be both inaccurate and unreliable, it did reveal a positive correlation between the

competency of the therapist and the perceived similarity between therapist and the patient (Fiedler, 1967:38). This finding, together with the assumption "that the way in which one person perceives another will affect his relations with him" suggested that there may be a link between the way leaders perceive their co-workers in a task group and the effectiveness of that group's performance (Fiedler, 1967:38). It was this connection which ultimately steered Fiedler's research into the leadership arena.

Fiedler originally measured the leader's perceptions by means of two instruments: the most preferred and least preferred co-workers scales (MPC and LPC, respectively). The difference between the scores on these two scales yielded a measure called the Assumed Similarity of Opposites (ASo). Consistently high correlations between ASo and LPC scores led to using only the latter scale. The instrument was further refined by changing from the original Q sort technique to a Likert-type scale, and then to the present semantic differential format (Fiedler, 1967:37-44).

Measurement of LPC

According to Fiedler (1964, 1967), the LPC is a measure of "leadership style" (i.e., of personality, needs-disposition) or, more recently, of "motivational structure" (1978:60). The instrument is completed by asking the subject to think of all the people with whom he or she

has ever worked and identify the one person with whom it was most difficult to work. The subject then rates this most difficult (least preferred) co-worker on a series of bi-polar adjective pairs (e.g., friendly.....unfriendly, cold.....warm, boring.....interesting). Each adjective of a pair represents one extreme of an eight point range. Respondents indicate where on this range of eight points they would place their target person. The "positive" adjectives (e.g., friendly, warm) score "8"; the "negative" ones (e.g., cold, boring) score "1". Although the earlier LPC instrument contained 17 pairs of adjectives, more recent versions include either 16 items (Fiedler, 1967:41, 269; Fiedler and Chemers, 1974:75) or 18 items (Fiedler et al., 1976:7; Fiedler, 1978:61). Subjects are designated as "high LPC" or "low LPC" in accordance with one of two procedures.

The first method is based on mean item scores. Based on a sample of 320 subjects, Fiedler indicates that high LPC mean item scores range from 4.1 to 5.9 ($\bar{x}$=4.9, SD=0.82) while, for low LPC persons, the scores range from 1.2 to 2.2 ($\bar{x}$=1.8, SD=0.43); the mean and standard deviation for the entire sample were 3.32 and 1.39, respectively (1967:43-44). Normative data, based on 2,014 subjects, indicate a mean and standard deviation of 3.71 and 1.05, respectively (Posthuma, 1970:11).

The second method of determining "high" and "low" LPC is based on the total score of the items. On an 18 item

instrument, a score below 58 indicates a low LPC individual while a score above 63 denotes a high LPC person. Subjects whose scores fall between 58 and 63 are referred to as middle LPC (Fiedler et al., 1976:8; Fiedler, 1978:61). According to Fiedler, LPC is normally distributed in the population with a mean of 60 (1978:91).

For individual subjects, the recommended procedures for designating respondents as high or low LPC are quite clear. What is much less clear is the method of dividing LPC scores for a group of subjects in a research investigation. There are at least three methods either suggested or endorsed by Fiedler. First, the subjects could be labelled high and low LPC on the basis of a mean split using Fiedler's (1967) value of 3.32 or Posthuma's (1970) norm of 3.71. However, in several studies (e.g., Bons and Fiedler, 1976) the split has been made using the sample mean rather than the normative mean. Second, Fiedler (1967) has indicated that only the upper and lower thirds of the sample distribution of scores should be used. On this basis, a group of subjects would, for example, be high and low LPC only if their scores fell above 4.1 or below 2.2, respectively.

The third method of score division is similar to the use of distribution extremes. More specifically, Fiedler (1971d) endorses the procedure used in the Chemers and Skrzypek (1969) West Point study in which cadets "whose LPC

scores fell either one standard deviation above or below the mean" were chosen as leaders (Fiedler, 1971d:136). The reported mean and standard deviation on a 20 item instrument were 70.0 and 21.1, respectively (Skrzypek, 1969:7)[2]. In order to render this sample mean comparable with any of the three discussed previously, it has been transformed into an item mean score of 3.50 (70.0/20) with a standard deviation of 1.05 (21.1/20). Thus, high LPC leaders in this study had scores of approximately 4.5 or greater while low LPC leaders had scores of approximately 2.5 or less. These score designations are very similar to the cut-off values of 4.1 and 2.2 reported by Fiedler (1967) for the two extremes of the sample distribution. The cut-off points used in all three methods of score division are summarized in Table 1.1. The table also includes the score values suggested in Fiedler (1971d).

Meaning of LPC

It was noted earlier (p. 5) that the emergence of the LPC instrument as a measure of leadership style was more a function of happenstance than of design. What began as the

---

[2] It should be noted that the West Point study was conducted by Skrzypek (1969) and originally reported as a technical report. The study was published in 1972 under the joint authorship of Chemers and Skrzypek. Both citations are included in the list of references.

TABLE 1.1: HIGH AND LOW LPC SCORE DESIGNATION FOR DIFFERENT SCORE DIVISION METHODS[1]

Score Division Method (Group of Subjects)	High LPC (>) item $\bar{x}$	total	Low LPC (<) item $\bar{x}$	total
upper/lower thirds[2]	4.10	66	2.20	35
above/below 1 SD[3]	4.50	72	2.50	42
mean value (n=320)	3.32	53	3.32	53
normative mean (n=2,014)	3.71	59	3.71	59
mean value (Fiedler, 1971d)	5.00	80	2.00	34
sample mean (Bons & Fiedler, 1976)	3.88	62	3.88	62
normative score (Fiedler et al., 1976)	4.00	64	3.63	58

1. All item mean scores are comparable; total scores above broken line are comparable (16 item instrument); total scores below broken line are not comparable with those above (18 item instrument).

2. The mean values for high and low LPC within each extreme were 4.9 and 1.8, respectively (Fiedler, 1967:43).

3. The scores are the mean values for each group of LPC scores.

operationalization of interpersonal perception between therapist and patient ultimately became the relationship between a leader's perceptions of his or her least preferred co-worker and the performance of that leader's group on its primary task. Although Fiedler's interpretation of LPC has undergone some modification during the course of its evolution, the notion that it reflects an individual's motivational structure or needs-disposition has remained a constant theme since 1964. While other interpretations exist (e.g., cognitive complexity: Mitchell, 1970;

value-attitude: Rice, 1975), their consideration is well beyond the scope and purpose of this thesis. Therefore, what follows is a brief description of the meaning and interpretation of LPC by Fiedler himself.

Traditional interpretation. At its most basic level, the LPC score is thought to indicate that some respondents tend to separate the person with whom a task must be performed from the task itself while others either do not make that distinction or do so to a lesser degree (Fiedler, 1964:155; 1967:44). Given this "implicit personality theory", respondents who have high LPC scores perceive the person with whom it was least preferable to work as "reasonably nice, intelligent, [and] competent" (Fiedler, 1967:44). Respondents with low scores on the LPC instrument perceive the person with whom it was most difficult to accomplish the task as "uncooperative, unintelligent, [and] incompetent" (Fiedler, 1967:44). These differential responses are interpreted by Fiedler to mean that:

> The high-LPC individual... derives his major satisfaction from successful interpersonal relationships while the low-LPC person ...derives his major satisfaction from task performance (1967:45).

This interpretation of high and low LPC scores led, in turn, to the conceptualization of leadership style as "relationship-motivated" and "task-motivated", respectively. However, the incorporation of the word "major" in the

above-quoted passage suggests that the two styles do not represent the complete conceptual dichotomy indicated by the terms themselves. Fiedler notes that when the situation is threatening to the satisfaction of their needs, each type of leader will exhibit behaviours characteristic of the other, but for different reasons. That is to say,

> The high-LPC leader will concern himself with the task in order to have successful interpersonal relations, while the low-LPC leader will concern himself with the interpersonal relations in order to achieve task success (Fiedler, 1967:46).

While this conceptualization of LPC in terms of a leader's needs-disposition has a certain intuitive appeal, it has not been totally supported by the empirical evidence. Several studies, summarized by Rice (1975), indicated that high LPC leaders were sometimes perceived by others to behave in a task-oriented manner while low LPC leaders were seen to display relationship-oriented behaviour (Rice, 1975:100, 250). In an attempt to account for these interaction effects, Fiedler (1970b)[3] proposed the "motivational hierarchy" interpretation of LPC.

---

[3] There are two different years cited in the literature. The re-interpretation was first formulated and written in a technical report dated September, 1970. The published version appeared in 1972. Both citations are included in the list of references.

Motivational hierarchy interpretation. This interpretation rests on a series of assumptions explicated by Fiedler (1970b:4). Collectively, these assumptions suggest not only that people have different goals but also that their motivation to achieve them corresponds to the importance they attach to each goal. Moreover, Fiedler notes that when individuals function in a threatening or stressful situation, they tend to concentrate on those goals which are of primary concern. When the environment is less stress-inducing, attention can be directed toward satisfaction of secondary goals.

When LPC is interpreted within the context of these assumptions, Fiedler suggests that high LPC persons are motivated primarily to achieve successful interpersonal relations. Their secondary goals include "self-enhancement, prominence, and esteem from others" (1970b:5). Central to the primary goal of low LPC persons is "the feeling of accomplishment derived from the knowledge that the job was well done" (1970b:5). Their secondary motivation focusses on interpersonal relations particularly when such relations will contribute to task achievement. Stated in general terms, the relationship between LPC and leader behaviour is contingent upon the nature of the work situation. According to Fiedler (1970b, 1978), it is this contingency which

accounts for the incongruence between the LPC score and observed behaviour found in some studies. The motivational hierarchy interpretation would thus suggest that when a low LPC leader is reasonably assured of task accomplishment, attention can be turned toward the interpersonal dimension. Under these circumstances, the low LPC leader's behaviour may be perceived by others as relationship-oriented. Likewise, the high LPC leader who feels that the group is functioning well may become concerned with task-related matters in an attempt to seek prominence. Such leader behaviour will likely be seen by others as task-oriented. The motivational hierarchy interpretation has been tested and, like the traditional interpretation, found to be lacking in empirical support (Rice, 1975).

Historical Background of Situational Favourableness

The situational favourableness dimension of the Contingency Model was the final outcome of Fiedler's (1967) effort to develop a taxonomy of groups. The development involved a series of three steps. The first step differentiated between social and task groups, only the latter being the concern of the Model. The second step categorized task groups as counteracting, coacting, and interacting. Each type of group is said to differ both in the dynamics required of its members to accomplish a particular task and in the role played by the leader. The

leader of a counteracting group, which is "typically engaged in negotiation and bargaining processes", functions as a "moderator or negotiator" to facilitate conflict resolution (Fiedler, 1967:20-21). In coacting groups, the members work relatively independently of each other to achieve the common task. The leader functions to motivate each individual member to perform to his or her maximum ability and, at the same time, prevent "destructive rivalries and competition" (Fiedler, 1967:19-20). By contrast, the members of an interacting group must work in close cooperation with each other. The accomplishment of their task depends upon "close coordination [and] interdependence of group members" (Fiedler, 1967:18-19). In this type of task group, the leader's role is one of "directing, channelling, guiding,...and coordinating the group members' work" (Fiedler, 1967:19).

As a result of extensive research with interacting groups, Fiedler developed a classification scheme in terms of the situational factors which would be "most likely to affect the degree of influence which the leader will enjoy over group behavior" (1967:22). The use of "degree of influence" as the basis for classifying situations is consistent with Fiedler's definition of leadership as being:

> an interpersonal relation in which power and influence are unevenly distributed so that one person is able to direct and control the actions and behaviors of others to a greater extent than they direct and control his (1967:11).

The Situational Variables

Fiedler selected three situational factors (variables) which he believed to be "of major importance" in the classification of interacting groups (1967:22). They are the interpersonal relationship between leader and members, the structure of the task, and the leader's position power. A brief description of each of these variables follows.

Leader-member relations. This variable, regarded by Fiedler (1967) as the most important of the three situational factors, refers to the leader's "affective relations with group members", to the "acceptance which [the leader] can obtain", and to the "loyalty which [the leader] can engender" (1967:29). The nature of this interpersonal relationship "is at least in part dependent upon the leader's personality" (1967:29). The primacy of this particular situational variable seems to have been based in part on common-sense observation, in part on the outcome of an unpublished study (Fishbein, 1966 cited in Fiedler, 1967:29), and in part on Hemphill's (1961) analysis of leader behaviour in terms of attempted, successful, and effective leadership acts (cited in Fiedler, 1967:31). Fiedler postulates that the quality of the leader's

interpersonal relationships with group members will affect the amount of control and influence the leader can exercise.

The leader-member relations variable may, according to Fiedler (1967, 1978), be operationalized in one of two ways. The sociometric preference method relies on the perceptions of group members. They are asked to respond to a series of questions regarding, for example, their choice of leader or coworker for a similar task in the future[4]. Fiedler does not indicate how the responses should be interpreted or scored in order to designate leader-member relations as good or poor. The alternate, and much simpler, method is the leader-completed Group Atmosphere Scale (GAS) — a semantic differential instrument containing ten pairs of bipolar adjectives each on an eight point scale. The leader is asked to describe the atmosphere of his or her group in terms of each of these adjective pairs. The instrument is very similar in item content to the LPC. Procedurally, the GAS score may be expressed either as a total or as an item mean score (Fiedler and Chemers,

---

[4] Fiedler gives three examples of the types of questions used in the sociometric determination of leader-member relations. The questions suggested are: (1) "If your group were asked to perform a similar task again, who in your group (or in your organization) would you want to have as your supervisor or leader?"; (2) "Who in your organization would you most like to work with?"; (3) "Who in your group had the most influence on the outcome of your deliberations?" (1967:31).

1974:65). No specific recommendation is given regarding the method of dividing the GAS score to reflect the level of leader-member relations. However, since a median split — or trichotomizing the sample distribution when there is a sufficient number of groups — is suggested for the sociometric preference rating, it would seem reasonable to follow the same procedures for the GAS scores.

There is one final observation to be made about the measurement of leader-member relations. Fiedler states that the sociometric method is the more appropriate in real-life organizations while the GAS is more suitable for use with short-lived, ad hoc groups assembled for laboratory-based research (Fiedler, 1967; Fiedler and Chemers, 1974). Fiedler bases this choice on the assumption that people who work together over time, as is the case in ongoing organizations, know each other well enough to provide a valid estimate of the leader-member relations dimension.

Task structure. This variable is considered by Fiedler (1967) to be the second most important determinant of the leader's control and influence with respect to group members. To the extent that the task is structured or unstructured, the leader's control is increased or decreased. For example, the leader of a group whose task is structured — that is, routine in nature and characterized by "standard operating procedures" — is in a more

favourable position to exercise control over the subordinates (Fiedler, 1967:26; Fiedler and Chemers, 1974:66). By contrast, the leader whose group is engaged in an unstructured task — one characterized by vague procedures and multiple approaches — "usually has no more knowledge than his members, and ... therefore no advantage over them" (Fiedler and Chemers, 1974:66).

The operationalization of task structure is based on Shaw's (1963) research on the classification of tasks (Fiedler, 1967:28). Of the ten dimensions identified by Shaw, Fiedler has selected the four

> which indicate the extent to which the leader is able to control and supervise his group members by virtue of the fact that the task is structured or capable of being programmed (1967:28).

To the extent that a given task is characterized by decision verifiability, goal clarity, procedural simplicity, and solution singularity, it is deemed to be structured (Fiedler, 1967:28; Fiedler and Chemers, 1974:67). Fiedler's procedure for assessing task structure asks raters or judges (not the leader) to evaluate each dimension on an eight point scale in which lower and higher scores indicate less and more structured tasks, respectively. A cut-off value of 5.0 is recommended to distinguish structured from unstructured tasks (Fiedler, 1967:149; Fiedler and Chemers, 1974:68).

One further point should be made concerning the way in which Fiedler views the task structure variable. It is

important to remember that this component of situational favourableness is conceptualized in terms of the task which the group is required to perform. It is the leader's function to ensure task achievement (Fiedler, 1967:26). This formulation seems to indicate that the leader's task is both different and separate from that of the group. To the extent that the group successfully accomplishes its task, the leader is said to be effective.

Position power. This variable is regarded by Fiedler as the least important component of situational favourableness (1967:144). The variable refers to the degree to which the position itself permits the leader to gain compliance from group members. According to Fiedler, position power is "highly related to French and Raven's (1956) concepts of legitimate power and reward-and-punishment power" (1967:22). In other words, the organization vests the position — as distinct from the incumbent — with, for example, the right to hire and fire, to promote and demote, and to direct and evaluate (Fiedler, 1967:23; Fiedler and Chemers, 1974:68). Fiedler contends that the more position power the leader has, the greater will be that leader's control and influence over group members (1967:25). It should be noted that Fiedler regards both position power and task structure as organizationally determined attributes of the group. Leader-member

relations, on the other hand, are established personally by the leader.

To operationalize the variable, Fiedler recommends use of a checklist containing "various indices of position power" (1967:23). Although, prior to 1978, Fiedler provided no specific instructions regarding the assessment of position power, it was implicit both in his discussion and in the checklist questions that it should be evaluated by persons other than leader (1967:23). Fiedler (1978) specifies that, because "self-assessments are subject to distortion", both position power and task structure "should be assessed by the leader's superiors" (1978:65). However, in neither case does Fiedler provide a cut-off score by which to determine strong and weak position power. While the incomplete procedural information for this variable may well result in considerable deviation in its measurement, Fiedler indicates that

> it is in fact only rarely necessary to rate leadership positions in work contexts. Practically all managers, supervisors, foreman, and superintendents in business and industry have high position power. Practically all committee chairmen and leaders of groups of colleagues tend to have low position power (Fiedler and Chemers, 1974:69).

The Contingency Model

Having determined that the Model would address task rather than social groups and having categorized the task groups into three types (counteracting, coacting, and

interacting), Fiedler then identified three variables (leader-member relations, task structure, position power) by which the interacting groups could themselves be classified. This second-level classification was based on the "assumption ... that different types of groups require different types of leadership" (Fiedler, 1967:32). Two points of clarification are needed in order to understand this statement. First, "different types of leadership" refers to leadership style as measured by the LPC instrument. Second, "different types of groups" refers to the differentiation of interacting task groups in terms of the three situational variables.

To create this classification scheme, Fiedler (1964, 1967) dichotomized the situational variables into good and poor leader-member relations, structured and unstructured tasks and strong and weak position power. These dichotomized values were then combined to yield eight different configurations called "octants". The eight octants were then arranged in an order which, for Fiedler, represents the amount of control and influence provided for the leader by each configuration. Thus, Octant I — which combines good leader-member relations with a structured task and strong position power — provides the most favourable situation for the leader. The least favourable situation is represented by Octant VIII in which leader-member relations are poor, the task unstructured and the position power weak.

The configuration for each of the eight octants is shown in Figure 1.1. Any given interacting group may be classified

Situational Variable	Octant I	II	III	IV	V	VI	VII	VIII
LMR	+	+	+	+	-	-	-	-
TS	+	+	-	-	+	+	-	-
PP	+	-	+	-	+	-	+	-

FIGURE 1.1: CLASSIFICATION OF GROUP TASK SITUATIONS

LMR = leader-member relations
TS = task structure
PP = position power
\+ = good, structured, strong
\- = poor, unstructured, weak

into one of these octants "by first ordering it on leader-member relations, then on task structure, and finally on position power" (Fiedler and Chemers, 1974:69). Once the groups were assigned to the appropriate octants, Fiedler correlated the group performance measures with the leaders' LPC scores to determine the relationship between the two variables.

The Contingency Model itself was induced from the diverse research findings relating group performance to leader LPC score. Correlations linking these two variables emerged from some 18 studies conducted by Fiedler and his

associates between 1951 and 1964. To order these data, each LPC-performance correlation coefficient (Spearman rho) was plotted according to the octant in which the group had been classified. Median correlations, shown in Table 1.2, were then calculated for each octant to generate the curve shown in Figure 1.2. These medians then became Fiedler's

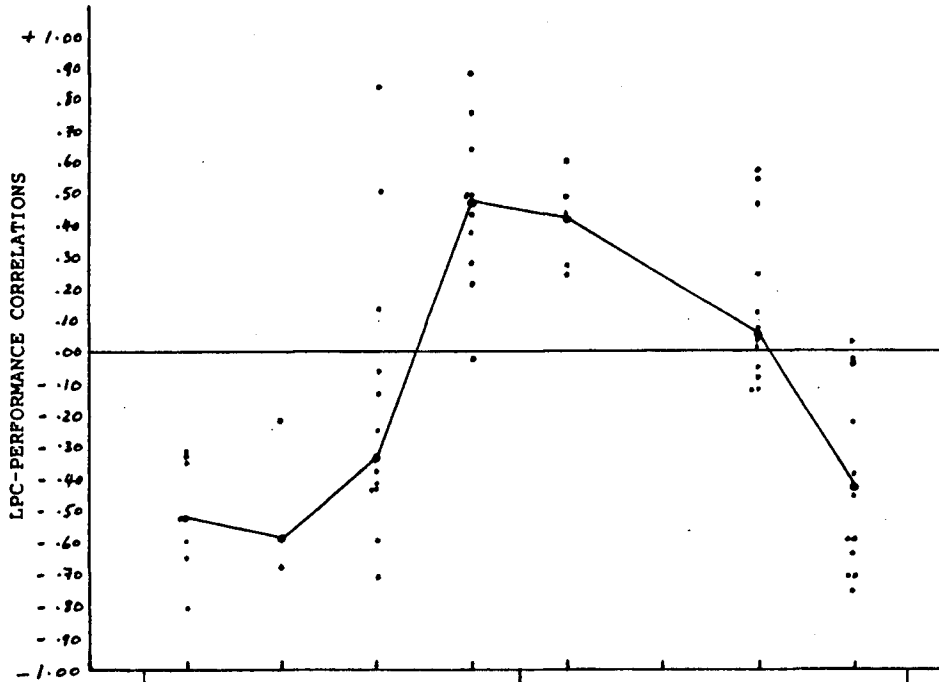
TABLE 1.2: MEDIAN CORRELATIONS BETWEEN LEADER LPC AND GROUP PERFORMANCE

Octant	LMR	TS	PP	Median rho	n
I	+	+	+	-0.52	8
II	+	+	-	-0.58	3
III	+	-	+	-0.33	12
IV	+	-	-	0.47	10
V	-	+	+	0.42	6
VI	-	+	-	(+)*	0
VII	-	-	+	0.05	12
VIII	-	-	-	-0.43	12

(adapted from Fiedler, 1967:142)

LMR = leader-member relations
TS = task structure
PP = position power
\+ = good, structured, strong
\- = poor, unstructured, weak
n = number of correlations included in median
* Fiedler predicted only direction for Octant VI

predicted value for each octant in subsequent studies. The positive correlations (above the solid line in Figure 1.2) between leader LPC and group performance indicate "that groups under relationship-oriented leaders (high LPC) tended to perform better than groups under task-oriented leaders



	I	II	III	IV	V	VI	VII	VIII
LMR	Good				Poor			
TS	Structured		Unstructured		Structured		Unstructured	
PP	Strong	Weak	Strong	Weak	Strong	Weak	Strong	Weak
LPC	Low	Low	Low	High	High	High	High	Low

FIGURE 1.2: LEADER LPC-GROUP PERFORMANCE CORRELATIONS BY OCTANT (adapted from Fiedler, 1967:146; Fiedler and Chemers, 1974:80)

Note: the lines join the median correlations which constitute the prediction for each octant; the small dots indicate the correlations from which each median was calculated.

(low LPC)" (Fiedler, 1967:145). The negative correlations (below the solid line) indicate "that groups under task-oriented (low LPC) leaders performed better than did groups under ... relationship-oriented leaders [high LPC]" (Fiedler, 1967:145). More specifically, the curve indicates that groups with low LPC leaders tend to perform more effectively when the situation is either favourable (Octants I, II, III) or unfavourable (Octant VIII). Groups with high LPC leaders tend to perform more effectively in situations of moderate favourableness [i.e., Octants IV, V, VI, VII]

(Fiedler, 1967:169). Thus, according to Fiedler, the Contingency Model is

> a theory ... which states that the group's performance will be contingent upon the appropriate matching of leadership style and the degree of favorableness of the group situation for the leader, that is, the degree to which the situation provides the leader with influence over his group members. The model suggests that group performance can ... be improved either by modifying the leader's style or by modifying the group-task situation (1967:151).

Although the majority of Fiedler's research has been based on interacting groups, there is some evidence to suggest that the Model may also be used to predict the conditions under which leadership may be effective in coacting groups. Because members of coacting groups tend to perform their tasks independently, Fiedler (1967) suggests that the leadership functions required for effective group performance may be somewhat different — in emphasis, if not in kind — from those required for effective interacting groups. These functions — identified as motivating, training, and quasi-therapeutic — further suggested that coacting groups should be subdivided into task and training/teaching groups. According to Fiedler, coacting task groups serve to increase "the ability or competence of the individual trainee" (Fiedler and Chemers, 1974:87). Fiedler reports the results of several studies which indicate that the Contingency Model generalizes to coacting task groups but not to training groups (Fiedler and Chemers, 1974:88). However, that the Model may not generalize to

coacting training groups is one of the less important problems associated with Fiedler's formulation. The ways in which the Contingency Model has raised problems are examined in the next chapter.

Chapter 2

THE PROBLEMATIC NATURE OF FIEDLER'S MODEL

During the past thirty years some 350 studies have been designed to derive, validate, and extend the Contingency Model. Collectively, they have yielded a vast array of empirical findings. That the findings and the conclusions based on them should be inconsistent, and sometimes contradictory, is not surprising when it is realized that the Model is an atheoretical formulation. Because the data used to derive the Model have also been used to support it, the Model gives the impression of having a theoretical soundness which it, in fact, lacks. At least in part, the absence of underlying theory to guide the research has resulted in disparity not only among the empirical findings themselves but also between the findings and Fiedler's formulation.

A Perspective on Theory

That Fiedler's Model was inductively derived from post hoc analyses of correlational data does not mean, in and of itself, that the model is atheoretical. There are those (e.g., Margeneau, 1950) who would suggest the correlational formulations are theoretical. Given the general agreement in the literature (e.g., Rudner, 1966;

Brodbeck, 1968; Monahan, 1975) that a theory is, at the very least, a set of interrelated terms or concepts, a correlational theory becomes theoretical by definition. Further, there are those (e.g., Suppe, 1979) who would suggest that, to qualify as theory a formulation must be able to predict certain outcomes which fall within its parameters. But, as Gale points out, prediction, together with control, "represents the lowest state of scientific knowledge" and does not contribute to an understanding of "the underlying explanation of why it works" (1979:63-64). As Hills observes, while "what correlates with what questions" are useful, it is also essential from the point of view of the advancement of knowledge to ask the why questions (1975:141). Moreover,

> It is only when empirical observations become logical consequences of a deductive system that they constitute a fully developed science, or an organized body of information... One can construct such a thing as a correlational science, but it is far less economical than a deductive science, and it suggests no hypotheses (Hills, 1975:119).

Viewed from this perspective, Fiedler's Model may be described as atheoretical. While the question of theoretical adequacy has been addressed by a number of authors (e.g., McMahon, 1972; Ashour, 1973a, 1973b; Kerr and Harlan, 1973; Korman, 1973; Schriesheim and Kerr, 1977), the lack of meaningful explanations for the relationships and predictions proposed by the Contingency Model is only one of of three problematic issues surrounding Fiedler's

formulation. The remaining issues can be grouped under two headings: methodological and empirical. The latter focus on the predictive validity and generalizability of the Model; the former on testing criteria and measurement of the variables comprising the Model.

Methodological Issues

Most of the controversy surrounding Fiedler's Model has focussed on the validity (construct, content, concurrent, and predictive) and reliability (internal, test/re-test, and response properties) of the LPC as a measure of leadership style (e.g., McMahon, 1972; Ashour, 1973a; Shiflett, 1973; Schriesheim and Kerr, 1977; Schriesheim, Bannister, and Money, 1979; Rice and Seaman, 1981). In view of its central position in Fiedler's formulation, it is not surprising that the LPC has been the subject of considerable research. However, this concentration on the "key variable" has perhaps tended to overshadow consideration of the three situational variables and of their relationship to both LPC and group performance (Fiedler and Chemers, 1974:73). Rather than focussing on the psychometric properties of the LPC itself, this discussion will emphasize its relationship to the other variables of Fiedler's Model.

Measurement issues involving LPC. It is suggested by

Shiflett (1973) that LPC is a better predictor of group performance when leader-member relations are good than when they are poor. For this suggestion to be accurate, the magnitude of the LPC-performance correlations in Octants I through IV should be greater than that in Octants V through VIII (see Table 1.2). This is not the case. Octants V and VIII both show stronger correlations than does Octant III (0.42, -0.43, and -0.33, respectively). Moreover, it contradicts the Model to isolate one situational variable from the configuration of three represented by each octant.

A second observation made by Shiflett (1973) concerns the method of dividing LPC scores. It will be recalled from the discussion in Chapter 1 that the division of LPC for a group of research subjects is not specified in precise terms. Shiflett points out that the use of extreme LPC scores prevents the normal distribution of "sampled leader LPC scores" (1973:430). Shiflett further notes that, while the effect of this procedure will be reflected when the Pearson r is used to correlate LPC and group performance,

> it will be obscured when using [Spearman] rho since the large gap between high and low LPC scores will be mathematically treated the same as the relatively small intervals between the scores at either extreme (1973:430).

This statistical observation raises some doubt regarding Fiedler's recommendation, discussed earlier, that extreme scores should be used to designate leaders as high and low LPC. The matter is even more problematic, given that the

Spearman rho is the recommended summary statistic for reporting the LPC-performance correlation (Fiedler, 1967).

There is another issue to be noted with respect to the measurement of LPC. In a laboratory study conducted by Graen et al. (1971a) leaders were not only randomly assigned to groups, their LPC scores were determined after the groups had been assembled. There are two important points to be made concerning this procedure. First, according to Fiedler, random assignment of individuals as group leaders will likely result in the uneven distribution of high and low LPC scores across the octants being tested (1971b:204; 1971d:139). An unequal distribution of LPC scores

> is not likely to give results which are predicted by the model since the model assumes roughly equal distribution of means and variances of LPC scores over the octants (Fiedler, 1971b:204).

Moreover, at least by implication, Fiedler is suggesting that non-division of LPC scores constitutes an unacceptable testing procedure.

The second point concerns the time at which leader LPC scores are measured. There is some evidence (e.g., Rice, 1975; Hosking, 1978) to indicate that the score is influenced not only by the leader's knowledge of the group's performance but also by his or her perception of leader-member relations when the Group Atmosphere Scale (GAS) is administered post-task. The possibility that the LPC score may be contaminated by either of these factors suggests that it should be determined before the group task

is begun. In this connection, it should be noted that despite the criticism by Ashour (1973a) and McMahon (1972) that LPC may not be independent of leader-member relations, Fiedler contends that "correlations between leader LPC and GA scores have been consistently around zero" (1973a:363).

Measurement issues involving situational variables. In addition to the possibility that leader perception of both group performance and leader-member relations may influence LPC scores, it has been suggested by several authors (e.g., Graen et al., 1970; McMahon, 1972; Ashour, 1973a) that the GAS score itself may also be contaminated by leader knowledge of group performance. This criticism raises an important question concerning the testability of the Model. It is clear that LPC can and should be measured at some time prior to the manipulation of the situational favourableness variables to avoid contamination by leader perception of either leader-member relations or group performance. However, if GAS scores are also to be measured pre-task in order to reduce the influence of group performance, it becomes increasingly difficult to introduce a temporal separation into the administration of the LPC and GAS instruments and still obtain meaningful scores for either variable.

The earlier discussion of leader-member relations indicated that Fiedler does not provide specific guidelines

by which to divide GAS scores for a group of research subjects. In this connection, Ashour (1973b) notes that because cut-off scores for this instrument tend to be sample-specific, the results across studies will be disparate. Yet the study by Graen et al. (1971a) showed that trichotomizing the GAS scores and partialling out the middle third yielded results essentially the same as those generated when a median split was used.

While it is recognized that an investigator's decision regarding when to measure LPC and leader-member relations and how to divide LPC and GAS scores is governed by both the purpose and design of the research, differences in ways of approaching these measurement questions may be responsible, at least in part, for the lack of congruent findings across studies.

Leader-member relations, however, is not the only situational variable which poses methodological problems. Both task structure and position power have been criticized on methodological grounds. While part of that criticism arises from the lack of a specified cut-off score by which to dichotomize position power (e.g., Mitchell et al., 1970; Ashour, 1973b), some of it stems from the relationship between task structure and position power. Fiedler (1967:153) reports that the two variables are significantly correlated ($r=0.75$, $p<0.01$). McMahon (1972) suggests that the position power variable be eliminated but that its

consequences be incorporated "in more comprehensive treatments of the group atmosphere and task structure variables " (1972:708). McMahon also suggests, based on the evidence from 12 non-Fiedlerian studies dealing with tasks and technology, that task structure may be a more important determinant of a leader's control and influence than is leader-member relations.

In defending the primacy of leader-member relations, Fiedler (1978) points to the results of two studies. The first, conducted by Fishbein et al., using analysis of variance procedures, concluded that

> the leader-member affective relation was, as would be expected, the most important single determinant of expectations about the most effective leader's behavior (1969a:467).

The second study (Nebeker, 1975) compared a multiple regression solution, in which the predictors were the situational variables, "with the zero-order coefficients obtained by the theoretical combination" suggested by the Model (1975:292). Nebeker concluded that

> Fiedler's theoretical combination of the component variables of situational favorability is about as close to optimal ... as could possibly be expected (1975:292).

Although both these studies lend credence to Fiedler's disproportionate weighting of the situational variables, neither addresses the problematic issue of the interdependence between task structure and position power. The nub of the criticism is that to the extent that the two

variables are not independent, the measurement is contaminated.

Measurement issues involving situational favourableness. In addition to Barrow's (1977) observation that the use of non-Fiedlerian instruments to measure favourableness yields incongruent findings, there are two important criticisms regarding this dimension of the Contingency Model. The first focusses on the relationship between favourableness and influence; the second, on the variables which are excluded from the situational dimension.

In their critical evaluation of the Model, Mitchell et al. (1970) have noted that there is no evidence indicating that the leader's "actual influence" varies as a function of the extent to which the situation is favourable for the leader (1970:14). A similar observation by Ashour (1973a) points to absence of evidence indicating that a given amount of situational favourableness is equal to the same amount of control and influence. Both these statements seem to indicate that situational favourableness lacks construct validity. There appears to be a need to validate, by some other measure, the belief that situational favourableness is a continuum along which control and influence decrease from Octant I to Octant VIII.

The second criticism of situational favourableness concerns not the three variables which comprise the

dimension but rather the variables which are excluded from it. While this criticism may be due in part to the absence of any theoretical rationale for the present variables, it may also be the result of Fiedler's oft-repeated statement that leader-member relations, task structure, and position power are not the only factors which determine the amount of control and influence for the leader (Fiedler, 1964:183; 1967:151; 1978:66; Fiedler and Chemers, 1974:70). However, while it may be desirable to incorporate new variables such as stress or group member LPC (Mitchell et al., 1970) or member-member relations (McMahon, 1972) or leader experience, training, and intelligence (Fiedler, 1978) or job satisfaction, cohesion, and morale (Hosking and Schriesheim, 1978), the Model — in its present formulation — is not sufficiently flexible to accommodate any of them (Barrow, 1977). Nor would it be scientifically sound simply to delete one variable merely to add another in a random, ad hoc manner. The criticism is more fruitfully interpreted as pointing to the need for research designed to refine the measurement of situational favourableness by comparing the relative strengths of the existing variables with those referred to by Mitchell et al. as the "neglected variables" (1970:16).

Testing criteria. The final methodological issue to be discussed in this review deals with the criterion or

criteria by which a study is deemed to provide supportive evidence for the Contingency Model. The acceptability of any set of findings seems to rely on only two factors: the direction and magnitude of the correlations reported by a researcher. Conspicuously missing is any reference to a generally acceptable level of statistical significance (e.g., $p<0.05$). The absence of that standard has been discussed in most of the critical reviews of the Model (e.g., Graen et al., 1970, 1971b; Mitchell et al., 1970; McMahon, 1972; Ashour, 1973a, 1973b; Hendrix, 1976; Schriesheim and Kerr, 1977). These authors all refer, in various terms, to the acceptance of correlations in the hypothesized direction as an alternative to statistical significance. Fiedler has defended this practice by stating that nonsignificant results can be meaningfully interpreted when a "sufficient number of correlations from comparable studies are high enough and in the same direction" (1973a:357). While the approach may yield significant results across several studies, it sheds no light on either the acceptability of findings from individual studies or on what constitutes sufficient magnitude.

It seems to be clear that, in practice, directional congruency is the only required standard for successful replication of one or more octants of the Model. The extent to which the magnitude of obtained correlations approximates Fiedler's predicted values appears to have been much less

important as a criterion for the acceptability of a set of findings. If it is conceded that the problem of small sample sizes frequently associated with the use of groups as the unit of analysis militates against statistically significant results, then the congruency of both magnitude and direction should be applied as testing criteria. Clearly the magnitude of a directionally incongruent correlation is irrelevant in terms of the Model. However, when the direction of the obtained value corresponds with that of Fiedler's predicted value, the magnitude criterion should be taken into account when drawing conclusions regarding support for the octant(s) under consideration. In addition to this array of methodological criticisms are those which, for purposes of this review, have been classified as empirical issues.

Empirical Issues

In light of the divergence of opinion which characterizes the ongoing debates between Fiedler and his critics regarding the methodological issues, it is not surprising that both the predictive validity and generalizability of the Model have also been challenged. Because many of the methodological issues tend to overlap with the empirical ones, only those criticisms which have not been included previously are reported in this section of the chapter.

Generalizability of the Model. It is generally accepted that one of the tests of the usefulness of a theory is its generalizability — its capacity to account for relationships beyond the specific occurrences for which it was originally developed (e.g., Kaplan, 1964; Harré, 1972; Kerlinger, 1973). Thus, to be useful, the Contingency Model should be applicable to task groups in a variety of organizational settings. Moreover, the results obtained in such tests should correspond with the Model's basic prediction that the effectiveness of high and low LPC leaders is contingent upon the situation in which their groups perform.

Apart from the assertion that the Model has been widely tested in a number of different organizational settings, there are remarkably few specific references by the critics to its generalizability. Ashour (1973a:348) notes that although there was a "broad sampling of settings and populations" used in the developmental studies, these samples varied from octant to octant. This diversity raised the possibility that the LPC-performance correlations could be "a function of the population sampled rather than [a reflection] of true differences between octants" (1973a:348). The extent to which the Model may or may not generalize to coacting task groups is not clear. Fiedler

reports that these "groups follow the same rules as interacting task groups" (Fiedler and Chemers, 1974:87).

One reason for there being little criticism with respect to the generalizability of the Model may be that it is difficult to separate generalizability from predictive validity. For example, if the results generated from testing one or more octants were deemed to be congruent with the predicted correlations, it would be reasonable to conclude that the Model does have predictive validity. Moreover, had those same results been obtained with a sample or setting or task group not often used, it would also be reasonable to conclude that the Model is generalizable. Given the possibility that these two tests of the efficacy of the Model have not been treated separately, the remaining discussion will focus on predictive validity.

Predictive validity. The critics seem generally to agree that the empirical evidence does not support the Contingency Model. It must be recognized, however, that none of the authors based his evaluation on all the studies which were available at the time of writing[1]. Moreover, in many instances, successive critics repeated the same

---

[1] One investigation did examine the findings from all studies. The results of that research by Graen et al., 1970 are reported in Chapter 3.

comments using the same observational bases.

Collectively, the reasons given for the lack of predictive validity are many and varied. For example, there are those who argue that supportive evidence is obtained only when there is rigorous adherence to Fiedler's prescribed methods (e.g., Ashour, 1973a; Barrow, 1977). In light of the earlier discussion regarding the difficulties arising from incomplete guidelines for measuring the component variables, it might be more accurate to say that the inconclusive evidence is attributable, not to methodological departures per se, but to the lack of more precise testing specifications.

A rather different, but potentially important, point arises from the exchange between Fiedler and Ashour. The pertinent argument focusses on whether the basis for determining predictive validity should be the results for individual octants or for all octants taken together. Fiedler (1973a) indicates that the decision should be based on the findings for all octants across studies by means of Stouffer's method of combining probabilities[2]. Fiedler argues that this procedure will overcome the problem of small sample sizes and will take into account results which are in the hypothesized direction (1973:358). By contrast,

---

[2] A recent study by Strube and Garcia (1981) was based on this procedure. The results are reported in Chapter 3.

Ashour contends that combining results from different octants "tends to obscure differences in the predictive power of the model among the eight octants" (1973b:370). Since each octant represents a different configuration of values on the situational variables and since there is a single predicted value for each octant except one (in Octant VI, only direction is predicted) it would seem to be more consistent with the formulation to combine the obtained correlations within rather than across octants. Remaining unresolved, however, is the question of how many octants must be tested simultaneously in order to declare the whole Model valid.

Conclusions

Given the present problematic nature of the Model, at least three conclusions about its continued usage seem possible. First, in keeping with the position exemplified by Ashour (1973a) and Schriesheim and Kerr (1977), the Model could be retired from the leadership theory arena. Second, in keeping with Rice's (1975) assumption that the Model does have "some degree of evidential validity", its continued usage in leadership research could be defended (1975:5). Neither of these two conclusions seems particularly useful. Instead of generating more empirical findings which cannot always be accounted for by the present formulation, it would seem more fruitful to examine the existing findings in order

to attempt to develop a theoretical framework which will account for them.

To this end, a meta-analysis of the Contingency Model research conducted between 1966 and 1981 was undertaken. The next chapter presents a discussion of meta-analysis as a method of inquiry. It also reviews three studies each of which has used a different approach by which to synthesize the results from research based on Fiedler's Model.

Chapter 3

THE METHOD OF INQUIRY: META-ANALYSIS AND INTEGRATIVE STUDIES

The research literature in almost every area of the social sciences is large and rapidly expanding. The size and complexity of any one part of that literature has created considerable difficulty not only in its review but also in the attempt to make sense out of frequently conflicting findings. In response to the need for a method whereby the research in any given area could be integrated into a whole more meaningful than the parts which comprise it, Glass (1976, 1978) proposed "meta-analysis". A discussion of that method of inquiry constitutes the substance of this chapter. In addition to a description and evaluation of meta-analysis, the chapter also reviews three studies each of which used a different approach by which to integrate the results from research based on Fiedler's Contingency Model. The chapter concludes with a statement regarding the use of meta-analysis in this study.

META-ANALYSIS

Meta-analysis is a method which permits the integration of quantitatively expressed findings from a number of original studies (primary analyses) all of which

have addressed essentially the same research problem. As an approach to the synthesis of empirical results across studies Glass states that meta-analysis

> connotes a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding literature (1976:2).

There are two dimensions which set this approach apart from the typical discursive type of review. First, meta-analysis is a study in its own right. Unlike a typical literature review, it is the whole study rather than just part of the study. Second, it applies statistical techniques to organize and extract "information from large masses of data that are nearly incomprehensible by other means" (Glass et al., 1981:22). In other words, meta-analysis is more than just a literary exposition which attempts not only to synthesize what is known about a given topic but also to identify the gaps in that knowledge. Meta-analysis is, at one and the same time, both the study and the method. To qualify as a meta-analysis in Glass' (1976, 1978) terms, the "study of studies" must fulfill three criteria: comprehensive coverage of the relevant literature, the derivation of a metric common to all studies and the consideration of covariance — the extent to which outcomes may or may not be affected by study characteristics (Glass et al., 1981).

Steps in Meta-Analysis

The conduct of a meta-analytical inquiry is a relatively straightforward undertaking. It is composed of four major and sequential steps: locating a sample of studies relevant to the topic of concern, recording the characteristics of each study in the sample, expressing the quantitative findings from each study in a metric common to all studies, and applying appropriate statistical techniques to determine the association of characteristics across all studies (Glass, 1976, 1978; Glass and Smith, 1979; Glass et al., 1981). An explanation of each of the four steps follows.

The location of studies. This phase of the research requires the investigator to identify as many studies as possible. Several authors (e.g., Glass, 1976, 1978; Rosenthal and Rubin, 1978) emphasize the importance of searching the dissertation abstracts, a source frequently overlooked or under-used by reviewers. Other sources helpful in locating pertinent studies include the ERIC data-base, the journal indices, the Social Sciences Citation Index, and what Glass calls "branching bibliographies" (1976:12). No matter how many studies the researcher is ultimately able to locate and access, they are unlikely ever

to include the total number of studies on the topic. For this reason, the collection of studies should be regarded as sample rather than the population (Glass, 1976, 1978; Glass and Smith, 1979).

Recording study characteristics[1]. The process of recording the study characteristics takes place in two phases. The first phase is a descriptive one and involves classifying the relevant variables onto a pre-prepared coding instrument. The coding instrument itself is set up to record information about the study (e.g., year, author, source), the variables of interest (e.g., sample size, instrumentation, type of subjects), and the major findings reported by the authors. Once the descriptive coding of all the primary analyses in the sample is complete, the second phase — "quantification of the findings" — may begin (Glass et al., 1981:93). This phase involves the transformation of the reported results into a common metric (Glass and Smith, 1979).

A common metric[1]. This phase of the meta-analysis

---

[1] Unless otherwise cited, this description of the coding process and the ensuing description regarding the common metric are based on Glass (1976, 1978) and White (1979).

involves the conversion of reported findings from primary studies which have used different designs (e.g., experimental, correlational) and summary statistics (e.g., Pearson r, F ratio, t value) to a common metric. Such a metric is essential in order to make comparable the findings across studies. For example, findings from experimental studies (treatment and control groups) "are probably best expressed as standardized mean differences between pairs of treatment conditions" (Glass et al., 1981:102). In this instance, all findings are transformed in a common metric referred to as an "effect size" which is defined as the mean difference between the treatment group and the control group divided by the standard deviation of the control group. Once computed, the effect sizes are then used to compare outcomes when the studies are grouped in terms of various characteristics.

However, when the primary investigations are correlational in design, the "effect size" may not be the most appropriate common metric. For such correlational studies, Glass indicates that the Pearson r should be used to express the final outcomes of the meta-analysis (Glass et al., 1981:147). This summary statistic is not always reported in the primary studies. The procedures for converting various summary statistics into the Pearson r are provided by White (1979:8-9), Glass et al., (1981:149-150), and McGaw and White (1981).

Statistical analysis. The descriptive and quantitative coding of the primary studies ultimately leads to the statistical analysis of the "relationships among the findings and characteristics of the studies" (Glass, 1978:362). This examination of covariance — of the extent to which study outcomes vary with selected study characteristics — constitutes what Glass (1981) calls "the essential character of meta-analysis" (Glass et al., 1981:21). The statistical analysis of the aggregated findings is analagous to the data analysis found in any primary study. The findings from the primary studies, once expressed in terms of a common metric, become the "raw" data for the application of the same methods which may be used to analyze the data collected for any primary investigation. Thus, according to its originator, "meta-analysis"

> is nothing more than the attitude of data analysis applied to quantitative summaries of individual experiments ... it is not a technique; rather it is a perspective that uses many techniques of measurement and statistical analysis (Glass et al., 1981:21).

A method which represents a new or different way of acquiring knowledge tends to draw a group of admirers who quickly become adherents and advocates. Interest in the novel approach tends to move rapidly from a simple curiosity to a fascination bordering on the apostolic. This phenomenon of immediate acceptance seems, all too often, to be accompanied by a lack of critical evaluation. As noted by Hedges and Olkin,

> the number of meta-analyses is growing exponentially, and is facing a danger of indiscriminate overpopularity (1982:160).

For this reason, it is important to consider those criticisms which do exist.

Criticisms of Meta-Analysis

According to Hedges and Stock (1983), there are two lines of criticism directed toward meta-analysis. The line which "stems from the lack of systematic statistical theory and formal statistical models" poses problems for the attention of statisticians and methodologists (1983:64). Since the refinement of meta-analysis is not the focus of the present study, it is sufficient to note that researchers such as Cooper (1982), Hedges and Olkin (1982), and Stock et al. (1982) are working toward resolution of these difficulties. Of greater importance for this particular inquiry is the second line of criticism which focusses on issues such as study comparability and quality, selection bias, and non-independent data (Glass et al., 1981). What follows is a brief description of each of these four criticisms and three observations regarding the limitations of meta-analysis as a method of inquiry.

Study comparability. Meta-analysis has been criticized on the grounds that it compares studies which

have used, for example, different methods of measuring similar variables or different types of subjects. Glass has responded to this criticism by noting that there is no need to compare studies which are the same because they would have yielded essentially the same results. Therefore, "the only studies which need to be compared or integrated are different studies" (Glass et al., 1981:218). However, Glass does not provide any guidelines to indicate how much difference between or among studies may be tolerable without loss of comparability.

Study quality. This criticism focusses on "the use of data from 'poor' studies" (Glass et al., 1981:220). It is noted by Hedges and Olkin (1982) that, while well-controlled and poorly-controlled studies yield differential outcomes when considered separately, there is little difference when they are analyzed together. Implicit in this observation is the decision to retain lower quality studies. In other words, the quality of any given study is not judged prior to its inclusion in the meta-analysis data base but rather is used as a covariate in the statistical analysis of the aggregated findings. As Glass states:

> the influence of "study quality" on findings has been regarded consistently as an empirical a posteriori question, not as a priori matter of opinion or judgment used in excluding large numbers of studies from consideration (Glass et al., 1981:222).

Those who would argue that a chain is as strong as its

weakest link would doubtless dismiss studies deemed to be inadequate. However, to eliminate such investigations is to assume not only that acceptable evidence can be yielded solely by unflawed studies but also that inferior studies can yield only less-than-useful results. As Glass notes, "many weak studies can add up to a strong conclusion" (Glass et al., 1981:220).

Selection bias. The main thrust of this criticism is directed primarily toward the inclusion of both published journal reports and unpublished dissertations and theses. There is empirical evidence to indicate that, regardless of study design, results reported in journals are stronger than those reported in dissertations. For instance, a summary of three meta-analyses by Kulik and Kulik (1982) revealed that the "average journal effect was 0.16 standard deviation units higher than the typical effect reported in dissertations" (1982:426). An identical effect was found by Smith (1980) in a summary of 12 meta-analyses (cited by Kulik and Kulik, 1982:426). According to Glass et al. (1981), published results are, "on the average, one-third standard deviation more disposed toward the favored hypotheses of the investigators" than are unpublished results (1981:66). Although these trends were based on the outcomes of experimental studies using effect size as the common metric, a similar pattern was found in

correlational analysis. For example, White (1979) showed that the magnitude of the mean correlation between SES and pupil achievement was lower in dissertations than in journals (0.20 and 0.25, respectively) (1979:17).

A second aspect of selection bias concerns the "chronological trends in research findings" (Glass et al., 1981:227). Based on the correlations between effect size and year of publication in six meta-analyses, Glass et al. (1981) show not only that more recent studies generate stronger results than do earlier ones but also that the difference is roughly equivalent to the bias between published and unpublished reports (1981:229). In light of the evidence indicating the extent of publication bias, Glass et al. (1981) stated that:

> it is appropriate to conclude that failing to represent unpublished studies in a meta-analysis may produce misleading generalizations. To omit dissertations because of their assumed lack of rigor is also unwarranted (1981:67).

To this conclusion, might well be added the recommendation that all meta-analyses should report the relationship between the aggregated findings and the year of publication for both journal reports and dissertations.

Non-independent data. The last of the criticisms of meta-analysis, as summarized by Glass et al. (1981), concerns the independence of multiple findings gleaned from a single study. This criticism, unlike the previous three,

concerns the statistical properties of the data which threaten the validity of the use of inferential statistics. Glass et al. (1981) suggest, as one answer, that each finding be regarded as an independent entity. They indicate further that the assumption of independence will err on the side of conservatism because:

> the effect of the dependence is almost surely to increase standard errors of estimates above what they would be if the same number of data points were independent (Glass et al., 1981:200).

Moreover, it is noteworthy that the majority of the meta-analyses reviewed for this chapter have apparently accepted the "untrue, but practical" assumption of independence (Glass et al., 1981:200).

Limitations of Meta-Analysis

It should be recognized, however, that meta-analysis is not without limitations. As an approach to the synthesis of research, meta-analysis cannot incorporate all kinds of data nor all the information needed for a complete assessment of the knowledge in any given subject area. The three limitations which follow are intended not as criticisms per se but rather as reflections on what meta-analysis cannot accomplish.

Direct and indirect evidence. The first limitation of meta-analysis concerns the nature of the data which can

be integrated. Jackson (1980) has noted that meta-analysis "can assess only relatively direct evidence on a given topic" (1980:452). That is to say, the findings to be aggregated are gleaned only from studies which have sought specifically to investigate the same area of concern. While the sharp framing of a meta-analysis is needed to delimit the scope of the inquiry, it is restrictive in the sense that it ignores any indirect evidence from primary studies on related topics. For example, an examination of only those studies using Fiedler's variables ignores whatever indirect but relevant evidence there may be about any one of these same variables in studies spawned by other models of leadership. The inability of meta-analysis to interweave both direct and indirect evidence is therefore seen as the first limitation of the method.

Masked evidence. The second limitation also arises from the sharp focus of meta-analytical studies. It concerns what Simpson (1980) and others see as the tendency to obscure small but potentially fruitful pieces of information which may in fact be present in the studies included in the meta-analysis. Two examples illustrate the point. First, the meta-analysis on class size and achievement (Glass and Smith, 1979) overlooks any other factors which may have had a moderating effect on the relationship between the two variables of interest. Second,

the meta-analysis linking class size with attitudes (Smith and Glass, 1980) focussed exclusively on the affective domain, omitting entirely the possibility of an interactive effect between class size, attitudes, and pupil achievement. The point to be made here is that if moderating and/or interactive effects do exist between or among variables, the information is masked by concentration on main effects.

Theoretical evidence. A third limitation arises from the "undeniably quantitative" nature of meta-analysis itself (Glass et al., 1981:22). It is an approach which creates numerical order out of empirical chaos. The limitation lies not with the capacity of meta-analysis to deal with large amounts of data but rather with its lack of capacity to deal with the theoretical dimensions underlying the data. It has been observed by Adair (1978) that quantitative summarizing overemphasizes the existence of effects at the expense of their meaning and significance. He states that:

> there seems to be a decided conceptual under-emphasis — i.e., an absence of theoretical statements tying together research in an attempt to explain the phenomenon (Adair, 1978:387).

If it is accepted, however, that the establishment of stable patterns of relationships among variables represents a first step in the process of theory development, meta-analysis may become a great contributor to the enterprise. If statistical integration can be used in a non-sporadic way to

create empirical order out of contradictory and inconsistent findings, then perhaps theoretical explanations could be developed to account for the aggregated results. To do so would be to strengthen the symbiotic relationship between theory and data. More will be said about this possibility following a discussion of the ways in which meta-analysis is presently being used.

Current Uses of Meta-Analysis

For the most part, meta-analysis has been applied to three different research designs (experimental, quasi-experimental, and correlational) to address three main kinds of questions. Meta-analyses of experimental or quasi-experimental primary studies have typically been designed to determine either the influence of one variable upon another or the effectiveness of one "treatment" relative to other "treatments". Studies which have sought to address questions concerning the influence of one variable on another include: Glass and Smith (1979): class size and pupil achievement, Luiten et al. (1980): advance organizers on learning and retention, Redfield and Rousseau (1981): teacher questioning behaviour on student achievement, Cohen et al. (1982): tutoring on educational outcomes, and Kulik et al. (1983): computer-based instruction in secondary schools. The effectiveness of different "treatments" has been examined in studies such as:

Hartley (1979): methods of individualized instruction in mathematics, Kulik et al. (1980): computer-based teaching versus conventional teaching, and Pflaum et al. (1980): different reading instruction methods.

Findings from correlational studies have been aggregated to ascertain whether an association between two variables of interest exists across studies and, if so, with what magnitude and in which direction. This type of meta-analysis is exemplified by: Hearold (1979): television and social behaviour, White (1979): SES and achievement, and Cohen (1981): college student ratings of instruction and achievement. A summary of these three current uses, classified in terms of purposes and questions, is shown in Table 3.1. More recently, studies focussing on the technical aspects of meta-analysis have begun to appear in the literature. These were discussed earlier in connection with the critical evaluation of meta-analysis.

A Different Use of Meta-Analysis

The studies cited above make clear that meta-analyses have generally had utilitarian and instrumental purposes. Investigators have typically been interested in substantiating the superiority of one method over another or in establishing a statistically significant empirical basis for a commonly held belief, such as class size affects learning. Very recently, a slightly different type of study

TABLE 3.1: CURRENT USES OF META-ANALYSIS[1]

Purpose	Questions	Examples
determine influence	much? little? none?	class size and achievement
compare treatments	better? worse? same?	different teaching methods
ascertain association	yes? no? maybe?	television and social behaviour

1. For an excellent comparison of meta-analysis with traditional integrative methods, see Cook and Leviton (1980).

has appeared in the meta-analysis literature (Johnson et al., 1983). In keeping with the present pattern, its functional purpose was to compare four different conditions of cooperation and competition. However, the authors' a priori development of a conceptual framework to be tested by integrating the pertinent findings represents a slight, but distinct, shift toward the inclusion of a theoretical dimension in meta-analysis. What has yet to be done is to use the aggregated data derived from a meta-analytic inquiry specifically for the purpose of theory development. Although there are three previous studies which have undertaken to synthesize research findings based on Fiedler's Contingency Model, none has attempted to use the

integrated results for this purpose. These investigations form the substance of the second part of this chapter.

A REVIEW OF PREVIOUS INTEGRATIVE STUDIES OF FIEDLER-BASED RESEARCH

Although none of the three prior studies used meta-analysis as defined by Glass (1976, 1978), all of them did attempt to synthesize the outcomes of research based on Fiedler's Contingency Model. The purpose of this review is three-fold: to indicate the approach used in each study, to report the outcomes, and to evaluate each investigation. The studies, classified in terms of approach, are presented in chronological order.

Pooling Correlations: Graen et al. (1970)

The majority of this journal report is devoted to an exposition and critique of Fiedler's Model. The authors conclude their evaluation with the suggestion that the Contingency Model could be tested by comparing the results from the "antecedent studies" (those used to derive the Model) with those from the "evidential studies" (those used to validate the Model) (1970:290). The 12 antecedent studies yielded a total of 63 LPC-group performance correlations; 83 correlations were found in the evidential studies.

The approach. These two sets of values were used in a 2 X 2 X 2 factorial analysis of variance in which the dichotomized independent variables were leader-member relations, task structure, and position power. The dependent variable was the covariance of LPC and group performance, the correlations having been "normalized using the Fisher Z transformation" (1970:292). Two ANOVA's were run, the first based on the antecedent and evidential correlations combined (n=146); the second, on the evidential correlations only (n=83). In addition to these two across-octant analyses, Graen and his associates used the t statistic to compare the antecedent and evidential correlations within each octant (except Octant VI for which, as was shown in Table 1.2, Fiedler predicted only direction).

The outcomes. The across-octant analysis of variance based on the combined sets of correlations revealed no significant main effects. There was a significant first-order interaction for leader-member relations with each of task and position power ($F=4.49$, $p<0.05$ and $F=22.00$, $p<0.01$, respectively). The authors conclude

> that something systematic is operating on the composite of antecedent and evidential results (Graen et al., 1970:292).

The same analysis on the evidential correlations yielded nonsignificant main effects and interactions leading Graen

et al. to conclude the absence of a systematic relationship between situational favourableness and the LPC-performance covariance.

The within-octant t tests showed that the antecedent and evidential mean correlations were significantly different one from the other in Octants I, II, IV, and V ($p<0.05$) and in Octant VIII ($p<0.01$). In addition, the pooled evidential correlations were directionally incongruent with Fiedler's predictions for Octants II, VI, and VIII (1970:292-293). As a result of both the across-octant and within-octant analyses, Graen et al. suggest that:

> the contingency model of leadership effectiveness clearly has lost the capability of directing meaningful research (1970:295).

While such a conclusion appears tenable on the surface, a more detailed examination of the outcomes may show that a more moderate viewpoint is warranted.

The evaluation. An evaluation of this study should reflect the time (1970) at which it was conducted. Obviously, Graen and his colleagues did not have the advantage of sophisticated methods such as meta-analysis (Glass, 1976, 1978). If this approach had been available to the authors, their investigation procedure, and subsequent findings, might have been quite different. However, viewed from the perspective of 1983, the Graen et al. study

contains several methodological weaknesses. For instance, the comparability of the studies in the evidential group was not clearly established (e.g., a study by Mitchell (1969) was not a validation test); the influence of study characteristics on the pooled correlations was not considered (e.g., variation in organizational settings); laboratory and real-life studies were not treated separately (Fiedler, 1967). Collectively, these weaknesses cast considerable doubt on the authors' conclusion that further use of the Model as a research tool be discontinued.

Formal Classification: Rice (1975)

Rice's research, which focussed on the interpretation of the LPC measure, exemplifies a different approach to the synthesis of data. The study is "meta-analytical" in the sense that it systematically classifies the findings gleaned from a random sample of 66 primary analyses in an attempt to ascertain what is already known about LPC in relation to other variables.

The approach. In order to classify formally variables which were empirically related to LPC, Rice (1975) modified the McGrath and Altman (1966) scheme developed to classify data from small group research. The classification system contained one dimension to "deal with [the] operational parameters of the data" and one to "represent

[the] substantive considerations" (1975:37). Each of these dimensions was subdivided into several categories to permit more precise grouping of the data. The unit of analysis was any reported statistical relationship between LPC and another variable. Through the "principle of operational concordance", it is possible to give an indication of the proportion of relationships which reach statistical significance.

The outcomes. Of the 1445 statistical relationships between LPC and another variable reported in the 66 primary analyses, 23% (334) were found to be significant (Rice, 1975:51). Clearly, the magnitude of Rice's study makes it impossible to summarize all the findings in concise form. There were, however, two outcomes which are particularly important because of their implications for the Contingency Model.

First, Rice found a consistently stronger relationship between LPC and leader performance than between LPC and group performance (1975:107, 130-131, 220). Since it is the latter relationship which is postulated by the Model, the finding that LPC was more strongly related to leader performance raises some doubt regarding the usefulness of LPC as the main predictor in Fiedler's formulation. With respect to this contradictory result, Rice concluded that:

> further research into the relationship between LPC and individual performance (and the relationship between individual and group performance) can potentially increase our understanding of the group performance effects central to the Contingency Model (1975:220).

However, it is also possible that in its present format, the Contingency Model may not be the most appropriate tool for generating an understanding of the complex leader-group interactions which account for variations in performance effectiveness.

Second, neither the traditional needs-disposition nor the motivational hierarchy interpretation of LPC was strongly substantiated by Rice's survey. Rather, the summarized data seemed to support Fiedler's earliest (1958) description of LPC as a measure of interpersonal attitudes (Rice, 1975:232). Therefore, Rice concluded that:

> The attitudes measured by the LPC scale apparently reflect a basic difference in the values of persons who score high and low on the scale. That is, high- and low-LPC persons appear to evaluate themselves and others in terms of different personal values (1975:232).

According to Rice, this "value-attitude interpretation of LPC" is compatible with Fiedler's (1964) position that the measure reflects a motivation-based "task versus interpersonal orientation" (1975:249).

The evaluation. By contrast with the parsimony of meta-analysis, Rice's (1975) classification system is

complex. That complexity is exacerbated by the unfamiliar terminology used to describe relationships (e.g., "non-additive lexicographic concordance" or "mode nonconcordance" or "batting averages"). The scheme does not permit the expression of the aggregated findings in terms of a single, easily understood value which indicates the strength of association between LPC and one variable relative to some other variable. A re-examination of Rice's data using meta-analysis may well be a fruitful undertaking.

Combining Probabilities: Strube and Garcia (1981)

Of the three studies which have attempted to synthesize the research findings generated from Fiedler's Model, the Strube and Garcia (1981) investigation is the only one which can be considered as a meta-analytic inquiry. However, the approach used by the authors was not based on the integrative methods suggested by Glass (1976, 1978).

The approach. To test the validity of the Model, Strube and Garcia (1981) conducted a meta-analytic inquiry, both within- and across-octants, using Rosenthal's (1978, 1979) methods for combining probabilities. The study, based on a sample which included both antecedent and validation LPC-performance correlations (n=33 and n=145, respectively), tested an hypothesis which reflected Fiedler's predicted direction for that relationship. More specifically, the

authors used Stouffer's formula for combining Z scores to generate the probability that the aggregated results occurred by chance alone (1981:310). The procedure was then used to analyze several different combinations of the data.

The outcomes. The results generated by Strube and Garcia (1981) were generally very supportive of the Model as a whole. There were, however, several instances in which the results within specific octants were not significant (i.e., $p>0.05$). For example, Octants III and VII were not supported when the antecedent data were analyzed (1981:311). When the validation evidence for interacting task groups was partitioned in terms of field and laboratory settings, only Octants I and IV in the field group and Octant IV in the laboratory group were statistically significant ($p<4.83 \times 10^{-4}$, 0.00164, and 0.00126, respectively) (Strube and Garcia, 1981:314). When these two groups were combined, the analysis revealed significance only in Octants I and IV ($p<3.02 \times 10^{-4}$ and 1.39×10^{-5}, respectively). However, when analyzed across all octants, the probability values for each setting and for the settings combined yielded highly significant results (1981:314). The analyses for coacting and training groups showed strong support for Octants V through VIII and for all octants combined (1981:317). When all the validation data were analyzed together (e.g., field and laboratory; interacting and coacting) the combined

probability was highly supportive ($p<2.99 \times 10^{-28}$, n=145) of the Model (1981:315). As a result of these findings, Strube and Garcia conclude that:

> The model as a whole was overwhelmingly supported .... Given that it took 13 years [1968-1981] to generate the present validation evidence ... it is unlikely that disconfirmation of the model is foreseeable in the near future (1981:316).

This conclusion seems to suggest that confirmation of the Model should be based on the outcomes for all octants taken together rather than on the confirmation of each individual octant within the Model. Thus, if confirmation is defined in terms of the whole, irrespective of the outcomes for the component parts, it becomes impossible to disconfirm the Model. Clearly, if every octant were supported by the aggregated findings, the only tenable conclusion would be one of confirmation. The point to be made is, that at least in conceptual terms, the Model can be confirmed in one of two contradictory ways: either some octants are supported by the evidence or all octants are confirmed by the evidence. The existence of a dual definition makes it impossible to draw any conclusion other than a supportive one. Given the evidence presented by Strube and Garcia (1981), the only defensible conclusion seems to be one of partial validation for the Contingency Model.

The evaluation. Apart from the comments concerning the conclusion drawn by Strube and Garcia, there are four

observations to be made about their study. It will be recalled from the description of meta-analysis in the first part of this chapter that journal articles reported stronger findings than those found in dissertations. Glass states that:

> If one integrates only "published" (meaning journal published) studies, the impression of support for the favored hypothesis will be artificially enhanced over what would be seen if the entire literature were integrated (Glass et al., 1981:66).

The Strube and Garcia research was based, almost exclusively, on published research. A review of the references used by the authors indicates that only two studies were unpublished — one, a bachelor's thesis (Bates, 1967); the other, a master's thesis (McNamara, 1967). No dissertations were included.

The second observation concerns the matter of study comparability. It was pointed out in the opening chapter that there was considerable variability in some of the procedures used in Contingency Model studies. Moreover, it was noted that Fiedler not only distinguishes between interacting and coacting task groups but also suggests that field and laboratory studies be treated separately. It would seem that none of these factors has been adequately taken into account in the Strube and Garcia study. Although, for example, the data for each task group and for each study setting were analyzed separately before being combined, the partitioned data sets were not tested for statistically significant differences. Rather, it was

assumed — presumably on the basis of observed similarity between the obtained results — that collapsing was a legitimate procedure. To the extent that the partitioned data sets may have been different, the final outcomes can only be regarded as dubious.

The third observation concerns the quality of the studies used in the data base. Again referring to the earlier discussion of meta-analysis, it will be remembered that Glass regards the covariance between study quality and findings to be "an important part of every meta-analysis" (Glass et al., 1981:222). There is no mention of either study quality or its influence on outcomes in the Strube and Garcia research. The systematic examination of covariance has been entirely omitted from the inquiry.

The fourth and final observation concerns the reliability of study coding. In connection with coding consistency Glass states that:

> The principal source of measurement unreliability in meta-analysis ... arises from different readers (coders) not seeing or judging characteristics of a study in the same way. Judge consistency or rater agreement is the most important consideration ... (Glass et al., 1981:75).

A careful examination of the Strube and Garcia study revealed no information of any kind as to either the coding procedure used or the consistency with which it was carried out. The absence of any measure of inter-rater reliability casts doubt on the credibility of the results reported by Strube and Garcia.

CONCLUSION

The stated purpose of this chapter was two-fold: (1) to explain meta-analysis as a method of inquiry and (2) to discuss three previous integrative reviews of Contingency Model research. The unstated purpose was to indicate that, although there have been prior attempts to synthesize the findings from studies based on Fiedler's formulation, none has used the more defensible perspective proposed by Glass. In addition, the research to date has lacked a thorough and comprehensive coverage of the literature as evidenced by the non-use of dissertations in both the Rice (1975) and Strube and Garcia (1981) studies. Further, the earlier inquiries have assumed study comparability, presumably because all the findings were generated by using only Fiedler's Model. Moreover, the thrust of the previous investigations has been directed, for the most part, toward establishing the predictive validity of the Contingency Model. The present study is directed rather to examining the extent to which the predictive validity of the Model exists. In so doing, an attempt was made to overcome the two problems noted above: to have a thorough, comprehensive coverage of the literature and to test rather than assume the comparability of studies to be included in the meta-analysis. The procedures used in the conduct of the present study are reported in the next chapter.

Chapter 4

PREPARATION AND PROCEDURES: DATA COLLECTION, CODING, AND REDUCTION

The conduct of a meta-analytic inquiry, like any other research study, includes the selection of a sample and the collection of data relevant to the problem under investigation. Since one of the purposes of this study was to determine which of the findings generated by Contingency Model research were consistent, reliable, and valid, it was necessary to identify as many Fiedlerian studies as possible. This chapter reports the procedures which were used in the preparatory stages of the meta-analysis. Although this step-by-step description of the preparation process conveys the impression of linearity, the stages were frequently non-sequential and overlapping. The chapter is divided into three main sections, the first two concerning data collection; the third, data reduction.

DATA COLLECTION: PHASE I

Data collection in meta-analysis takes place in two phases. The first phase is the selection of the sample studies; the second is the extraction of data from those studies. This section reports the procedures by which relevant studies were identified, collected, and organized.

Identification of Studies

At this early stage of the meta-analysis, no firm decision had been made regarding either the number of years or the specific years to be included in the data base. The problematic nature of the Model did however, suggest the possibility that Fiedler's developmental studies, conducted prior to 1965, might be useful in attempting to account for the empirical inconsistency reported in the literature. It was, therefore, deemed important to identify these studies as well as the later ones. For this reason, the process of study identification began with the references cited in Fiedler (1967:292-303).

Branching bibliographies. The reference information for each of the studies (n=55) cited in Fiedler (1967) was recorded on a separate filing card. For ease of cross-checking this initial group of studies with those gleaned from other sources, an alphabetical list of authors, source, and year was simultaneously maintained. This same process was applied to the bibliographies from the following nine sources: Fiedler (1964, 1971a, 1977, 1978); Fiedler and Chemers (1974); Fiedler et al. (1976); Rice (1975, 1978) and Schriesheim and Kerr (1977). Each of the references from these nine sources was cross-checked against the

alphabetical listing in order to avoid duplication. Together, the ten searches identified a total of 327 potentially useful documents. The pool included 181 journal articles (from 48 different journals), 38 book chapters (in 33 different books), 21 doctoral dissertations, 11 masters' theses, 19 papers (10 unpublished, 9 convention), and 57 technical reports.

Psychological Abstracts. In addition to the use of branching bibliographies, a delimited computer search of the Psychological Abstracts was conducted. There were two reasons for the delimitation. First, the initial method of study identification revealed that the inclusion of articles from journals published up to 1978 had been fairly comprehensive (as indicated by the frequency of citation across the sample of 48 journals) and, therefore, it was decided to search only the four years 1978 through 1981. Second, more than 65% (119/181) of the articles identified from the branching bibliographies were gleaned from only 25% of the journals (12/48) and, therefore, the decision was made to limit the search just to those twelve titles. In this connection, it should be noted that one of the 12 journals, namely the Educational Administration Quarterly (EAQ), is not included in the data base for the Psychological Abstracts. This particular journal was therefore searched manually in the EAQ indices for the same

four year period. Thus 11 journals were computer-searched using the following key words: Fiedler, contingency model, contingency theory, leadership effectiveness, least preferred coworker (co-worker), LPC, group performance, situational favorableness (favourableness), and situational control. The search revealed 23 documents (20 journal articles and three dissertations) not previously identified, thus raising the subtotal for the sample to 350.

Dissertation Abstracts International. This data base was the third source used in the identification of studies. A computer search spanning the years 1955 through 1981 was conducted. The four years prior to 1955 were not included because neither of the first two methods of study identification revealed any doctoral dissertations completed during that early interval. By using the same descriptors as those in the previous search, 36 dissertations were identified. Following the computer search, 16 additional dissertation titles were obtained — four from Fiedler (personal communication, 1981) and 12 from the reference lists of previously identified dissertations — raising the total to 76 dissertations. To facilitate the collection of the documents, a reference card was completed for each title identified. The ways in which the reference cards were used is explained following the section on study collection.

Collection of Studies

These combined searches yielded a total of 402 potentially useful documents. As each journal article or book chapter was located, it was checked for its relevance to the meta-analysis. Those studies which did not use at least part of Fiedler's Model were deleted from the pool of documents. Photocopies were made of all pertinent journal articles and book chapters.

The dissertations and theses were somewhat more difficult to collect. With the help of Interlibrary Loan, most of the master's theses and some of the dissertations were borrowed from the degree-granting institutions. In some cases, personal contact had to be made with the authors all of whom were kind enough to agree to a direct loan and return of their studies. The majority (n=62/76) of the dissertations, however, were purchased from University Microfilms International either on film (n=58) or in paper-back (n=4). All the technical reports and some master's theses were obtained directly from Dr. Fiedler at the University of Washington.

Organization of the Studies

Because of the large number of documents which formed the potential data base for the meta-analysis, it was necessary to develop a system which would not only classify

the materials in a logical way but also facilitate their retrieval. To this end, a system was designed, prior to data collection, using three guidelines, namely: simplicity, quick access, and accuracy checks. These criteria correspond to the three purposes underlying the organizational system.

The first purpose was to facilitate quick and easy retrieval of any document from the file. To this end, all photocopied materials were colour coded by both source and type. The second purpose was to prepare the documents for eventual computer analysis by assigning each of them a numerical code which corresponded to its particular colour code. By coding the colour and number simultaneously, the time required for document classification was considerably reduced since each item was handled only once. The final purpose was to ensure that all documents identified for potential use were, in fact, actually collected. Each of the purposes was served by both the Document and Reference Card Files.

Document file. This file contained all photocopied articles, dissertation single pages, and microfiche envelopes which had been identified and collected. The file was composed of thirty-one separate units. Each unit comprised one white divider indicating the year (1951 through 1981) and five coloured folders which denoted the

source from which the a particular item had been derived. The source and colours were: journals (blue), books (green), dissertations/theses (yellow), papers (red), and technical reports (pink). All items, prior to placement in the appropriate folder (i.e., year and source), were assigned two codes, one numerical and one colour, the latter using two self-stick labels. The first label, whose colour was the same as that of the folder, denoted the source. The second label, affixed to the right of the first, indicated the type of document (e.g., primary study, commentary/critique, unpublished paper). The numerical code, which translated the colour coding into computer-readable format, was written in beside the two self-stick labels. The source and numerical codes were: books (1), journals (2), technical reports (3 or 4 or 5 or 6), dissertations/theses (7), convention papers (8), and unpublished papers (9). The numerical codes for each type of document were: primary study (11), review of research (12), critique (13), exposition (14), commentary (15), response by Fiedler to a critique (16), doctoral studies (17), and master's theses (18). For example, a document labelled "blue-orange-211" denoted that it was found in a journal (blue=2) and was a primary study (orange=11). The complete coding system by which documents were classified is shown in Tables 4.1 and 4.2.

TABLE 4.1: DOCUMENT FILE SYSTEM BY SOURCE
(colour and numerical coding)

General Source	Colour Code	Specific Source	Colour Code	Number Code
Book	Green	—	—	1
Journal	Blue	—	—	2
	Pink	Univ. of Illinois	Pink	3
Technical	Pink	Univ. of Wash.	Purple	4
Reports	Pink	Univ. of Florida	Lemon	5
	Pink	Other Institutions	Black	6
Dissertation	Yellow	Doctoral	Purple	7
Thesis	Yellow	Master's	Yellow	7
Papers	Red	Convention	Flame	8
	Red	Unpublished	Red	9

TABLE 4.2: DOCUMENT FILE SYSTEM BY TYPE
(colour and numerical coding)

Type	Colour Code	Label Initials[1]	Number Code
Primary Study	Orange	PS	11
Review of Research	Orange	RR	12
Critique of Model	Lime	Cr	13
Exposition of Model	Lime	Ex	14
Commentary on Model	Lime	Co	15
Response by Fiedler to Critique	Lime	RF/Cr	16
Doctoral	Purple	—	17
Master's	Yellow	—	18

[1] the appropriate initials were written on the self-stick label itself; the use of more coloured labels would have reduced the simplicity of the system.

Reference Card File: Stage I (Identification). This file contained one card for each of the 402 documents which had been identified from all the searches. During this stage, the cards were filed by specific source (e.g., specific journal titles, library books, personal books, University of Washington technical reports). This procedure not only made the location of materials in the University Library more efficient, it also provided a way of ensuring that a particular document had been collected. Following the colour and number coding of a given document and its placement in the appropriate source folder, its reference card was similarly coded and immediately transferred to the Stage II file box. This gradual removal of cards from the Stage I file was designed to indicate that the document had indeed been located and photocopied.

Reference Card File: Stage II (Collection). The Stage II file was set up using a format identical in principle, and similar in structure, to that of the Document File (i.e., by year and source). The only structural difference involved the use of coloured labels affixed to the tabs of the index cards rather than coloured index cards. Each reference card, already coded by colour and number during the first stage, was then filed by source within the appropriate yearly unit. The cards remained in this Stage II file until the document itself had been read

and its variables coded.

Reference Card File: Stage III (Alphabetical).

Although the alphabetical filing of the cards was actually part of the second phase of data collection, it is reported here to complete the description of the card file system. Following the coding of variables from a given document onto the coding instrument, the corresponding reference card was then transferred to the Stage III file. This third stage card file served three purposes. First, it not only provided the alphabetized list of references for inclusion in the thesis, but also simplified the process of distinguishing between different publications by the same author in the same year. Second, the Stage III file effectively separated those documents whose variables had been coded from those whose variables had not. This separation permitted quick identification and retrieval of previously coded documents for re-examination and possible re-coding of some variables. Finally, the card transfer procedure provided the final check that all documents had, in fact, been coded on the instrument. The removal of all cards from the Stage II file (i.e., Stage II file empty) indicated that the first phase of data collection had been completed. It should be noted that throughout the processes of data identification, collection, and organization, no attempt whatever was made to decide which of the accumulated

documents would remain in the final sample. Decisions regarding exclusion were left until the second phase of data collection.

DATA COLLECTION: PHASE II

The first phase of the data collection process resulted in an accumulation of potentially useful documents. The second phase was the extraction and coding of information from each individual document. In order to record the required information from each document, a coding instrument — specific to the purpose of the research — must be developed and tested. This section reports not only the evaluation and piloting of the instrument but also the preparation and revision of the code book without which the instrument is useless. The section concludes with a brief description of the coding process.

The Coding Instrument

The coding process is a method of classifying both the characteristics and the findings of a set of studies. To be useful in the later examination of the influence of study characteristics on the aggregated findings, the classification categories must be broad enough to permit the grouping of similarities yet narrow enough to prevent the blurring of differences. Too much or too little breadth

will ultimately fail to distinguish the influence of a particular variable.

A major difficulty in coding is the need to decide whether or not particular distinctions may become important in the consideration of covariance. If the distinctions are not coded, they cannot be easily retrieved. If they are coded in too much detail, the size of any sub-sample becomes too small to permit meaningful analysis. If the detail is not coded, the sub-sample sizes will be larger but the discriminatory power of the variable would be greatly reduced.

In addition to the difficulties which arise regarding category breadth and detail, there is a further — and more basic — problem in coding. The problem concerns the selection of variables to be coded. As White (1979) had noted, the success of a meta-analysis

> in explaining and summarizing the results of previous research depends largely on whether the proper variables are coded for each study (1979:9).

To some extent, this problem may have been simplified by the singular focus of this meta-analysis. While it was clear that the five variables of Fiedler's Model had to be included on the coding instrument, it was much less clear as to how their inclusion should be handled. Two examples illustrate the point. First, given the possibility that the dependent variable of the Contingency Model can be interpreted as leadership effectiveness (which the Model

purports to predict) or group performance (as leadership effectiveness is operationalized) or the covariance of LPC and performance (moderated by situational favourableness), it was deemed advisable to allow for all three alternatives. In turn, this meant that provision also had to be made to code LPC as an independent variable. Second, since some studies were conducted as laboratory experiments while others were carried out in on-going organizations, it was necessary to reflect these differences in the coding of both task structure and position power. This, in turn, also necessitated the coding of ways in which these two variables were not treated.

Once it had been decided how to code the variables in the Model, the question still remained as to which other study characteristics, apart from the actual findings, should be included. White (1979) has observed that:

> there is no fail-safe technique for making sure that all of the proper variables are included in a meta-analysis (1979:10).

For purposes of this particular meta-analysis, the initial selection of variables was based on a first review of the literature. That original list underwent considerable modification and extension as a result of two pilot runs. The procedures specific to the pilot runs are reported following a discussion of the code book.

The Code Book

Implicit in the preceding discussion was the idea that some variables had several values. A particular study characteristic perhaps could not be coded in binary terms such as "yes" (code value 1) or "no" (code value 2). In fact, only one variable out of the 165 coded for the meta-analysis could be treated dichotomously. Each of the remaining 164 had at least three possible code values. There was clearly a need not only to keep track of the values assigned to each variable but also to establish a method which would help to maintain consistency in the values assigned to any one variable across the studies. In order to do both, a code book was drafted and then revised as new variables or values were added or deleted. It should be noted that, although the list of variables — once finalized — remains unchanged, the list of possible values of each variable may increase during the course of recording the data from the studies. The additional values do not affect the structure of the coding instrument itself. The importance of an explicit, detailed code book for coding reliability is emphasized by Stock et al. (1982).

The 45 page code book written for this study is best described by giving four specific examples which indicate the range of possible values which could be used to code a particular variable. First, in an attempt to determine

whether there was a pattern among journal title, type of article published (e.g., research report, commentary, critique), and the nature of the findings, the journal names were listed and a two-digit code value assigned to each one. The 50 titles on the original listing increased to 53 as documents arrived from the Interlibrary Loan Division. To account for documents which were not journal publications, a different code value was available. Second, in an effort to balance category width and detail (sample size and discriminatory power), the variable "type of subjects" was first divided into leaders and group members. Within each of these two categories, there were five sub-categories (military, university/college, schools, profit- and service-motive organizations) which described the general population from which the subjects had been drawn. Each sub-category had a series of code values which described, in as precise terms as possible, the subjects who participated in the research. For instance, the sub-category "schools" contained 11 descriptive values such as: system administrator (e.g., superintendent, director), elementary principal, secondary teacher, and elementary pupils.

The third example, selected from the variables classified as "research procedures", concerned the score division methods used for both the LPC and Group Atmosphere scales. These two variables were included in an attempt to determine not only the extent of the methodological

variability but also the influence it may or may not have had on the aggregated outcomes. In the case of LPC, there were 16 values each indicating a different procedure for designating subjects as "high LPC" and "low LPC" (e.g., mean or median split, different units of standard deviation, extremes of the score distribution). The Group Atmosphere Scale was coded in terms of six possible values, each reflecting one of the six different methods for determining "good" and "poor" leader-member relations found in the studies.

The final example is given to indicate the way in which quantitative data such as LPC scores and sample sizes (leaders and group members) were coded. In coding study characteristics of this type, two variables were used. The first, referred to as the "prefix variable", provided information about the data themselves. For instance, the prefix variables for high and low LPC scores indicated whether or not they had actually been reported by the author or whether they had been calculated for the present study from other information which had been provided by the author. For example, when the author reported the sample mean and standard deviation and indicated that the cut-off points were units of standard deviation, it was possible to compute the scores used to designate subjects as high or low LPC.

The second type of quantitative data variable,

referred to as the "raw data" variable, was coded in one of two ways. When the data were available (or had been calculated), the actual numbers were recorded on the coding instrument. When the data were not available, for whatever reason, the "raw data" variable was coded using zeros. The decision to use the time-consuming method of recording zeros was made to ensure that subsequent checking, which took place long after the coding, was fully accurate.

In addition to the values which were specific to particular variables, the code book also provided a set of four values which had general applicability to any variable. This set was included either to reflect that the specific code values did not describe what the original researcher had done or to indicate that the specific information required to code a variable was not made available by the primary author. The four general values were: some other method/measure used (in which case, it was written in), scores/information not reported by the original author, cannot tell from what was reported by the primary researcher, and variable not applicable to the primary study (a number of subvalues were used to indicate the reason).

The Coding System: Pilot Runs

While much of the preceding discussion concerning the coding instrument and the code book has reflected the lessons learned from the pilot runs, there are several

points to be made specific to the pilot testing itself. The typical purpose for piloting an instrument is to test its reliability and validity using appropriate statistical techniques. For the coding instrument, however, the purpose of testing was to ensure its substantive adequacy. That is to say, it was essential that the instrument be capable of capturing as much detail as possible while simultaneously permitting classification of the information. In an attempt to achieve this objective, the instrument was subjected to a number of major revisions during the course of pilot testing.

The preliminary version of the coding instrument contained 28 variables listed on two pages. This instrument was soon expanded to three pages, 44 variables, and a few descriptive values. It was subjected to a pilot test using a sample of 12 documents (four dissertations, four journal articles, and four technical reports). This test revealed the need not only for a greater number of variables but also for the addition of values to improve the discrimination. It was at this point that the need for the code book, discussed in the previous section, became apparent.

In addition to writing the code book and revising the list of variables from 44 to 138 (from three to nine pages), another major change was made in the instrument. In an attempt to consider study quality as a covariate, both the original and second versions of the coding instrument had

included a section in which to write a short critical evaluation of the study. It became clear that, to be useful, this critical evaluation had to be quantified. To this end, two scores, each representing a different aspect of study quality, were developed. The TV (threats to validity) score, reflecting the validity of a primary study, was composed of nine criteria drawn mainly from McGaw and White (1981) and Borg and Gall (1983). This score described what Glass et al. (1981) would call the methodological dimensions of the study. In addition, six substantive aspects — reflecting the extent to which a study adhered to Fiedler's method for testing the Model — were appraised. These six criteria together comprised the MA (methodological adherence) score. Each of these 15 criteria was scored on a five point scale in which the higher value indicated a more favourable evaluation of the particular criterion. The scores on all criteria within each aspect of study quality were summed and averaged to yield the TV and MA scores. The 15 criteria and the calculation procedures used in the evaluation of each study are shown in Appendix I (pages 26 and 27 of the Coding Instrument). The second pilot run was conducted using this third version of the instrument.

The sample for this pilot test (n=6) included two dissertations (from the four used previously), one published report, two technical reports, and one convention paper. None of these four studies had been used in the first pilot

run. Although this third version of the coding instrument was found to be more adequate than the the two earlier editions, the list of variables and values required even more expansion and elaboration. The revisions suggested by the second pilot run were incorporated into the fourth and last version of the coding instrument. What had begun as a simple, two-page, 28-variable document was finalized as a complex, 27-page, 165-variable coding instrument deemed to be sufficiently adequate to record both study characteristics and the findings reported by the authors. The final version of the instrument was reproduced so as to allow one for each of the collected documents. A copy of the finalized instrument appears in Appendix I.

The Coding Process

The large number of potentially useful documents made it necessary to establish an orderly approach to coding them. The time span represented by the studies together with the knowledge that some changes had occurred in the Model itself, suggested that coding in chronological order would be sensible. To retain as much coding consistency as possible, it seemed advisable to code one source in its entirety before starting the next. In view of their shorter length, the data from the journal reports were coded first. They were followed by the dissertations/theses and the technical reports in that order. The reason for omitting

book chapters and unpublished papers is discussed in the final section of this chapter.

The last matter to be considered was the existence of both published and unpublished versions of the same study. To avoid duplication of data, it was necessary to code one or the other. Since there is a tendency for greater detail to be reported in the original version (e.g., a dissertation as distinct from the published article(s) based on it), the published paper was deleted from the pool of documents. The actual coding of the studies was a fairly straightforward, albeit very time-consuming, enterprise. For the most part, the problems which emerged were minor and were resolved in consultation with one or more members of the thesis committee. The coding instrument proved to be a satisfactory tool for data collection.

Apart from providing the data to be used in the meta-analysis, the coding process also revealed that some primary studies did not report sufficient information to permit inclusion in the data base. These research reports were placed in a different group as were the expositions, critiques, commentaries, and reviews of research. In all cases, however, the ten demographic variables (e.g., author, year, source and type of study: see Appendix I) were coded for possible future use. Further, the coding process identified a number of studies which either were not applicable (n=22) or contained what Cohen et al. refer to as

"crippling methodological flaws" (1982:239). These studies (n=13) were discussed in detail, and at length, with members of the supervisory committee before being rejected from the document pool. The matter of rejected studies together with the different sorts of documents for which not all variables were coded, and the delimitation of the document pool are all discussed in the final section of this chapter.

DATA REDUCTION

This section reports first the disposition of the 402 potentially useful document titles identified at the beginning of the data collection procedure. Second, it describes the three groups into which different types of documents were placed. Third, it defines three terms each of which refers to a different, but related, unit of analysis. Fourth, it indicates the way in which the retained studies were tentatively partitioned for data analysis.

Document Reduction

The reduction of the 402 potentially useful documents occurred in three successive stages during the course of data collection. The first reduction took place following document identification but preceding document collection. At this time, a total of 54 reports (24 journal articles,

three dissertations/thesis, and 25 technical reports) were found to be inapplicable because they did not use Fiedler's Contingency Model nor any part thereof. They had originally been identified on the basis of Fiedler-related descriptors such as "leadership effectiveness", "group performance", "leadership style", and "situational leadership". Two studies, which did use Fiedler's Model, were deleted because they were written in foreign languages. The document pool was thus reduced to 348 reports of various kinds.

The second, and largest, reduction (n=70) occurred as the documents were being coded. These were deleted for one of five reasons: (1) they were not direct investigations of Fiedler's Contingency Model and were therefore not applicable [n=22; 10 journal articles, four dissertations, one technical report, and seven validation studies of the Leader-Match training programme (Fiedler et al., 1976)]; (2) they never arrived from Interlibrary Loan (n=10); (3) they could not be located (n=7); (4) they were either in both published and unpublished form or the same data appeared in two different publications (n=18); (5) they were rejected (n=13). The reasons underlying rejection are reported at the end of this section. At the conclusion of the coding process, 278 documents remained in the pool.

It had been intended at the outset to include in the meta-analysis the data from the 38 book chapters, the 10 unpublished papers, and the nine convention presentations.

These 57 documents were deleted for two reasons. First, book chapters tended, for the most part, to be expositions of the Contingency Model or reviews of research findings relative to the Model. No chapters were found to have included findings which had not been reported in previously coded studies. The second reason was purely a practical one. The time designated for completion of the coding had already been exceeded by seven or eight months. Since the coding of a primary study — of the sort reported in occasional papers — took not less than three or four hours and since the information in book chapters was repetitive, it was agreed within committee, that no further coding would be done. The deletion of both books and papers and of the developmental studies (those conducted prior to 1965) were the only two delimitations placed on the meta-analysis.

The final total in the pool was 221 documents which included 112 published primary studies, 57 dissertations and theses, five technical reports, 18 reviews of research, 16 critiques, and 13 expositions of the Model. More will be said about the grouping of these documents following the reasons for rejecting some studies.

Thirteen primary studies were rejected from the document pool. All these studies contained either more than two major flaws or a series of minor flaws which, taken together, became major. In either case, the shortcomings were deemed to render the findings sufficiently suspect to

warrant exclusion from the data base. Some studies (n=5) were flawed methodologically. For example, two did not use Fiedler's LPC to measure leadership style; three operationalized the situational variables in non-Fiedlerian terms (e.g., weak manipulation of leader-member relations, task structure measured in terms of lesson content, position power determined on the basis of a single question); two contained sampling problems (selection and attrition); and in one study, the calculation of the means was inaccurate. One study was considered to be conceptually flawed because leadership style was defined in terms of the subordinates' perception of the leader's style using the LPC scale — the leaders themselves did not complete the instrument. A third group of studies was flawed both conceptually and methodologically. For instance, in one study LPC was conceptualized and measured as leader behaviour rather than style. In another, situational favourableness and organizational structure were equated and no attempt made (or, if made, not reported) to correlate the alternate measure with those used in the Model. Yet another substituted teacher performance for task structure and, in so doing, contaminated the measure of effectiveness. It should be emphasized that the decisions to reject these studies were made only after extensive consultation with one or more members of the thesis committee. Had the results of any of these studies been included, the outcomes of the

meta-analysis would, themselves, have been suspect. A list of these 13 studies indicating the specific reasons for rejection appears in Appendix E.

Document Groups

The 221 documents remaining in the pool comprised both primary studies and other types of articles and reports. These other sets of documents were deliberately collected in an attempt to accumulate as much information as possible about Fiedler's Contingency Model. What had not been planned was to separate the primary studies from other types of material. Nor had it been anticipated that some primary studies would report so little detail that the study quality scores (see page 91) could not be calculated. In an attempt to avoid confusion at the data analysis stage, it was necessary to separate the expository articles from the primary studies and also to separate out those studies whose brevity made complete coding impossible. To this end, the first integer of the three digit identity number was used to denote the nature of a particular document. Thus, as each document was read and coded, it was assigned to one of three classification series.

The "300 series" (n=75). The "300 series" documents included all the non-primary studies (i.e., critiques, expositions, and reviews of research). In addition to these

46 articles, the series also embraced a number of primary studies which did not directly investigate the relationships among the key dimensions of Fiedler's Model. For example, one primary study (Potter and Fiedler, 1981) examined the effects of stress, experience, and intelligence on task structure. Another by Beach and Beach (1978) investigated the disproportionate weighting of the situational variables. Besides these published studies, four dissertations were placed in the "300 series". For instance, Rice's (1975) dissertation (described in Chapter 3) which classified LPC-related information was included because, while it was a primary study for its author, it was a review of research for purposes of this present study. Collectively the "300 series" documents (listed in Appendix C) not only provided a wealth of Contingency Model-related information they also indicated the extent to which variables such as stress, motivation, task ability, leader intelligence, and experience have been used in an attempt to extend the Model.

The "400 series" (n=34). This series was composed entirely of primary studies with the exception of one secondary analysis (Kennedy, 1982) and the Strube and Garcia (1981) study discussed in Chapter 3. Of the remaining 32 studies, 16 were technical reports, 13 were journal publications, and three were master's theses. None of these studies provided sufficient detail to permit assignment of

study quality scores. For this reason, these primary investigations could not, for purposes of the present study, be considered comparable with those reporting more complete information. As a result, the studies in this series were not included in the meta-analysis data base. The "400 Series" studies are listed in Appendix D.

The "100 series" (n=112). The "100 series" comprised only primary studies. Without exception, these studies were direct investigations of Fiedler's Model. All of them provided sufficient information about the research to permit a critical appraisal. Three sources of studies are included in this series, namely research reports published in journals (n=58), dissertations/theses (n=49), and technical reports (n=5).

The frequency of each source or type of document within each series is shown in Table 4.3.

TABLE 4.3: FREQUENCY OF DOCUMENT SOURCE AND TYPE WITHIN EACH SERIES

Document	Series			Totals
Source/Type[1]	100	300	400	
Journal (PS)	58	25	13	96
Diss/Th (PS)	49	4	4	57
Technical Report (PS)	5	0	16	21
Review of Research	0	17	1	18
Critique	0	16	0	16
Expositions	0	13	0	13
Totals	112	75	34	221

1. Abbreviations: PS = primary study
Diss = Doctoral dissertation
Th = Master's thesis

Parenthetically, it should be noted that the numbers from 200 through 299 were kept in reserve in case they were needed to code the primary studies. As a result, there is no "200 series". The ultimate disposition of the "100 series" forms part of the substance of the next chapter.

Definition of Terms

The assumption made prior to the actual coding of the primary studies was that each one would include a single sample of subjects who participated in the research. It had not been anticipated that investigators would conduct replications within their own studies (e.g., Hosking, 1978) nor that they would conduct two or more different, but related, studies (e.g., Csoka and Fiedler, 1971) and report

them as a single piece of research. In order to capture all the data reported in every "study within a study", each one had to have its own identity number for purposes of computer analysis. To distinguish between these "multiple studies" and the "single studies", the terms "sample" and "study" were used. A "study", for purposes of this thesis, was defined as all the research reported by the author(s) irrespective of source (journal, dissertation, etc.). On the other hand, a "sample" was used to define each of "multiple studies" within the single research report. For example, Hill (1969a) reported the results of two "samples" (interacting groups in a business firm and coacting groups in a hospital) within a single, published article. To capture all the data from both samples, each was assigned its own identity code number. In a series of replications, Sashkin (1970) first compared the effects of leadership training and no leadership training using different groups of undergraduate students and then attempted to replicate the findings in two other related investigations. All four inquiries were reported within one dissertation. Again each of the four samples was assigned a separate code number. Thus, when examining the number of studies as distinct from samples, the value of n will be smaller. For instance, although there are 112 primary studies, the total number of samples within those studies is 146.

The third term which requires some clarification is

"octant test". This term is defined in keeping with the Contingency Model itself. An "octant test" simply refers to the relationship between the leader's LPC score and the performance of the group (or of the leader) within one of the eight value configurations of the three situational variables.

While it had been expected that the studies would typically test more than one octant, it had not been anticipated that, within the same study, the same octants would be tested more than once. That is to say, if a researcher was investigating Octants I, III, V, and VII, it was assumed that only one result would be reported for each octant. In several studies, however, the author used more than one measure of performance and correlated each one with the leader's LPC score. Again in consultation with one or more members of the thesis committee, decisions — specific to each study — were made as to how many of the dependent measures should be retained. In cases where the measures were clearly different (e.g., quantity vs. quality of performance or process vs. product), both were used. However, when the researcher reported an overall measure — a composite of the criteria used to assess performance — the correlations based on the individual measures were discarded. Thus, in addition to "multiple-sample" studies, there were also "multiple-test" samples. An awareness of these distinctions (studies, samples, octant tests) is

needed to understand the partitioning of the data to be discussed next.

Stratification of Data

The variety both in task group types and in the basis for assessing effectiveness found in the 112 primary studies necessitated some stratification of the data. Before presenting the stratification scheme, two points concerning the Contingency Model should be recalled from the discussion in Chapter 1. First, Fiedler (1967) suggests that the nature of the relationships among the members of a group is a function of the tasks they perform. Interacting and coacting task groups are distinguished by the amount of interdependence or independence required to perform the task. Furthermore, the amount of research on coacting groups is very limited. The second point to be remembered is that Fiedler (1967, 1978) assesses leadership effectiveness, not on the basis of the leader's performance, but rather on the performance of the group. The stratification scheme, shown in Figure 4.1, is based on these two distinctions.

In order to determine the legitimacy of combining the octant tests from the two different task group types, it must first be established that the pattern of LPC-performance correlations for each of them is similar. Moreover, the similarities must be present for all octants in which both groups have been tested (i.e., Octants I, III,

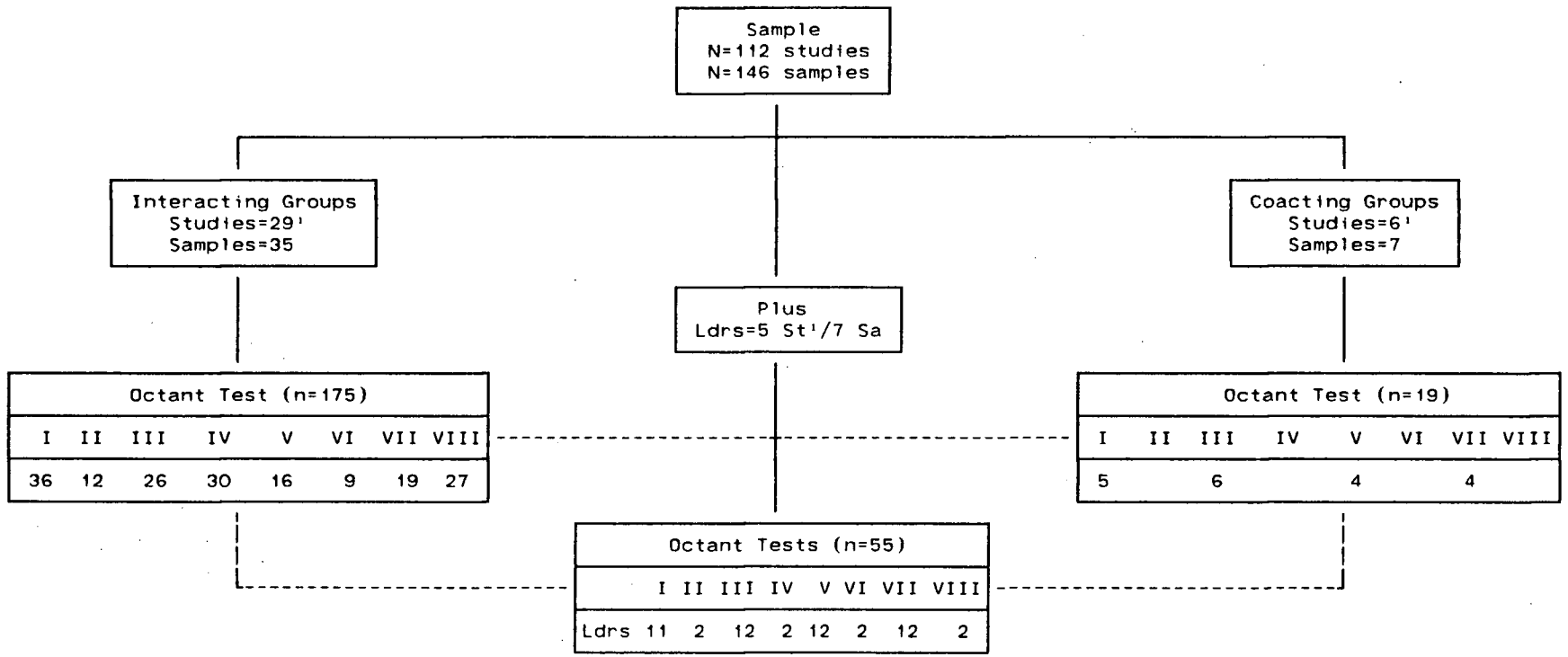


FIGURE 4.1: STRATIFICATION SCHEME

Abbreviations: Ldrs = Leaders only; St = Studies; Sa = Samples

Broken Lines: indicate possible collapsing, given patterns of similarity between the two task group types and between leader performance- and group performance-based octant tests.

1. n=40 (29+6+5) studies rather than 38 because two studies used two different task group types and were therefore counted twice.

V, and VII). If the tests which are based on leader, rather than group, performance are also to be combined with the interacting and coacting groups, a pattern of similarities must also be established. However, before this testing is done, it is first necessary to ensure the comparability of the 112 studies which comprise the "100 series". The determination of that comparability, which constitutes the first phase of the data analysis, is the subject of the next chapter.

Before reporting the assessment of comparability, it is necessary to consider one aspect of the nature of the data, namely whether their status is that of a sample or whether they may be held to constitute a population. The exhaustive search of the literature identified all relevant journal articles and technical reports. These two groups can therefore be regarded as the population of studies within each type of document. Because there are some dissertations which are not listed in any of the available indexes, no definitive statement can be made concerning the number of such studies omitted from the present investigation. The decision was taken to regard the present data as the population of comparable studies produced between 1966 and 1981 in published journals, technical reports, and dissertation indexes. For this reason, the results of the present study are reported using only descriptive statistics.

Chapter 5

THE COMPARABILITY OF STUDIES IN THE DATA BASE

The identification of 112 primary studies as being appropriate for the "100 Series" left unanswered the question of what degree of comparability existed among them and hence to what extent they could all be included in the meta-analysis. This chapter describes the results of assessing their comparability. The assessment was based on an examination of the way in which each study treated the five variables of Fiedler's Model. The end result of this process was the identification of two groups of studies which were labelled "Octant" and "Non-Octant". For reasons which become clear as the chapter progresses, only the "Octant" studies were retained for the meta-analysis.

This chapter discusses first the two groups of studies. The second major section describes the creation of two distinct comparability categories within the "Octant" studies themselves. A third section provides evidence of the reliability of the classification system. The final section presents conclusions and discussion.

"NON-OCTANT" AND "OCTANT" STUDIES

It will be recalled from Chapter 1 that the eight different profiles of situational favourableness in the Contingency Model provide the basis for Fiedler's predictions about leadership effectiveness. It became clear from the examination of the "100 Series" studies that a major distinction could be drawn between those studies which had tested one or more individual octants of Fiedler's Model and those studies which had used the Model in other ways. The latter group of "Non-Octant" studies will be described first, followed by the "Octant" studies.

"Non-Octant" Studies (n=74)

The label "Non-Octant" was used to denote studies that did not report results for octant tests. Initially, 66 studies were placed in this group. Subsequently, eight studies from the "Octant" group were added to the "Non-Octant" category. The reasons for these reclassifications are explained following the section dealing with the groups of "Non-Octant" studies.

The "Non-Octant" studies themselves focussed on a wide variety of topics all of which arose directly from Fiedler's Contingency Model. In spite of the varying directions taken by these studies, it was possible to

subclassify them into four groups. One group consisted of studies focussing in a general way upon the situational favourableness components of Fiedler's Model. A second group paid attention primarily to the LPC measure. An identifiably different group of studies emphasized how the meaning of LPC might be determined. The final group omitted the LPC measure and used only the situational variables of the Model.

The group of studies (n=32) which focussed, for the most part, either on the situational components themselves or on the possible influence of some other variables upon situational favourableness came to be known as the "Model focus" group. Examples of this first group included studies by McNamara (1968) (the applicability of the situational favourableness variables to a school setting); Williams and Hoy (1971) (usefulness of the Model in elementary schools); Stone (1979) (prediction of teacher satisfaction and pupil achievement in elementary schools); and Millard (1981) (applicability of the Model to females). Studies examining the influence of some other variable on situational favourableness were exemplified by Fiedler et al. (1969) (the influence of environmental stress); and Bons (1974) (the effect of change in the job environment).

The second group of studies (n=25) focussed predominently on LPC within the context of the Model. It was that predominance which separated these studies from

those in the previous group. For example, Konar-Goldband et al. (1979) examined three interpretations of the role of LPC with respect to leader-member relations and group performance. A very different approach was used by Schriesheim (1979). He sought to determine the influence of social desirability, as a measure of response bias, on three Fiedlerian self-report, evaluative instruments: the LPC, the GA Scale, and the Position Power Checklist.

A third group of studies (n=14) sought to ascertain the meaning of LPC. Gruenfeld and Arbuthnot (1968) did so by examining some personality correlates of the measure. Shiflett (1974) explored the effect of using a stereotype rather than a real person as the scale referent. Four studies each investigated a different interpretation of LPC: the motivational hierarchy interpretation (Rice and Chemers, 1975); the cognitive complexity interpretation (Arnett, 1978c); the cognitive differentiation interpretation (Hosking, 1978); and the value-attitude interpretation (Rice et al., 1978).

The last group of "Non-Octant" studies (n=3) differed from those in the "Model focus" group in that none of them used the LPC scale. For example, Mitchell (1970) was concerned with the construct validity of the Leader Behavior Description Questionnaire (LBDQ) dimensions of consideration and initiating structure in relation to only one situational variable, namely the GA Scale. Nebeker (1973), on the other

hand, used all three situational components to determine if favourableness would be more parsimoniously conceptualized as environmental uncertainty.

It was noted earlier that eight studies were changed from the "Octant" to the "Non-Octant" category. Of these eight studies, two reported results for pooled octants. Utecht (1970) tested the hypotheses that successful low and high LPC leaders in military settings had held positions in group situations which were, for the most part, favourable to their leadership style (i.e., low LPC's in Octants I, II, III, and VIII; high LPC in Octants IV, V, and VI). In the second study which examined the effectiveness of university faculty, Wyrick (1972) tested four hypotheses by combining the octants according to the degree of situational favourableness (i.e., Octants I, II, III = very favourable; Octants IV, V, VI = moderately favourable). Results yielded by these studies could not be meaningfully combined with those from studies testing individual octants. These two studies were therefore added to the "Model focus" group.

A further four of these eight studies reported results for individual octants but did so using the F ratio as the summary statistic. Each of these studies was characterized by Type IV error (Marascuilo and Levin, 1970) — that is to say, a mixture of classical analysis of variance with simple main effects analysis on the same data. In addition to the Type IV error, there were two other

reasons for excluding these studies from the data base. First, two of the studies did not report sufficient data to permit the transformation of the F statistic to a Pearson r, as suggested by Glass et al. (1981:150). Second, although the correlation could be derived for the other two primary analyses, their inclusion would have introduced "type of study" as another variable. Since the controls applied to experimental or quasi-experimental designs are different from those in correlational designs, the results generated from each may also be systematically different. Given that there were only two studies in the set, meaningful analysis of the influence of study design on outcomes would not have been possible. The decision was therefore made to control for "type of study" by exclusion.

The deletion of these four studies subsequently led to the removal of a fifth individual octant testing study in which a classical multivariate analysis of variance was used. It is not possible to compare the univariate F ratios from the MANOVA with the simple main effect F's from the ANOVA studies. All five of these studies were also placed in the "Model focus" group.

The remaining study, which was re-classified as "Non-Octant: Model focus", used both individual octants and the LPC scale. However, rather than examining the leadership effectiveness of their subjects (college presidents), Van Gundy and Haynes (1978) were interested in

these subjects' perceptions of situational favourableness. The LPC scores of the presidents were correlated with college enrolment but this relationship was not used either as a measure of effectiveness or as a validation test of any octant. The findings were therefore not comparable for purposes of statistical aggregation. All eight of these reclassified studies, together with the initial group of 66, are listed in Appendix B.

It is unfortunate, given the high quality of some of the "Non-Octant" studies, that the lack of both individual octant testing and correlational statistics necessitated their removal from the data base for the meta-analysis. The deletion of these studies should not be interpreted to mean that they made little or no contribution to an increased understanding of the dynamics implied by Fiedler's Model. Their exclusion reflected only the fact that the results yielded by these studies could not be aggregated with those from the "Octant" studies.

Octant Studies (n=38)

These 38 studies, gleaned from the entire "100 Series", constituted the data for the meta-analysis. The "Octant" studies had five characteristics in common. First, all of them had tested one or more individual octants of the Model. Second, a correlation statistic was used to report the results for each octant test. Third, some form of

Fiedler's LPC scale was used to measure leadership style. Fourth, all studies assessed the level of the three situational variables (leader-member relations, task structure, and position power) in an acceptable way. To have been considered "acceptable", one or more of the following conditions prevailed: (1) all three variables were measured either before or before and after task performance; (2) the level of one or two of the situational variables was assumed prior to task performance and the assumption(s) validated post-task either empirically or on the basis of plausible argument; (3) not more than two of the variables were "held constant" — that is to say, a particular variable was considered in the pre-task assessment of situational favourableness but was not measured in any way. In all such studies, a sound case for the decision was presented by the original investigator. Fifth, each of the 38 primary analyses used a measure of leadership effectiveness based on group or leader performance. Pre-existing success, such as achieved military rank, was not regarded as a measure of effectiveness for purposes of comparability in the present study. The extent to which the octant studies exhibited these five characteristics is reported in the next section.

COMPARABILITY CATEGORIES WITHIN THE OCTANT STUDIES

Although all 38 of the octant studies displayed these five characteristics, they did not all do so to the same extent. Some adhered closely to Fiedler's methods with respect to each element; others deviated from Fiedler's procedures in their treatment of one or two of the elements. These deviations were substantial enough to make advisable the creation of two categories of study which were referred to, respectively, as "Comparability Category:Fiedler" (CC:Fiedler) and "Comparability Category:Other" (CC:Other).

Comparability Category: Fiedler

The studies included in CC:Fiedler (n=10) were those which adhered most closely to Fiedler's prescribed methods for testing the Model. All of these studies used either the Group Atmosphere Scale (GAS) or sociometry or both to measure leader-member relations. The task structure and position power variables were either measured or assumed and validated using Fiedler's instruments (Fiedler, 1967; Fiedler and Chemers, 1974; Fiedler, et al., 1976). The only exceptions to the task structure and position power procedures were the four laboratory-based studies in which these factors were defined (task structure) or instructionally manipulated (position power) in accordance

with Fiedler's recommendations. In all the studies, both laboratory and real-life, the degree of situational favourableness — and thus designation of the specific octant(s) being tested — was based on a composite of the GAS, task structure, and position power manipulation results. In nine of the ten studies, leadership effectiveness was determined by correlating leader LPC scores with the measure(s) of group performance. One real-life study (Hill, 1969a) used a direct evaluation of leader rather than group performance.

Comparability Category: Other

The studies classified as "CC:Other" (n=28) adhered less closely to Fiedler's prescribed procedures. While within the group they varied in a number of ways, no single study deviated beyond an acceptable range. Moreover, the deviation was rarely in the instrument used to measure a situational variable but rather in the subjects who were asked to complete the Fiedlerian scales. Several examples are given to illustrate the general nature of the variations found in the measurement of leader-member relations, task structure, and position power.

Leader-member relations. The leader-member relations variable is operationalized, according to the Model, by having the leaders complete the GA scale (Fiedler, 1967:32;

1978:62). This procedure was followed in 17 of the 28 CC:Other studies. Of the 11 studies which did not follow the procedure, only two (Hovey, 1974; Sashkin et al., 1974) used non-Fiedlerian instruments. Six studies used the GA scale but varied in the subjects asked to complete it. For example, in Schneier's (1978) study of emergent leadership, both leaders and group members were asked to assess the level of leader-member relations. A similar procedure was followed by Stemler (1980) in his examination of some behavioural and attitudinal correlates of LPC. By contrast, Blanchard (1978), in her study of the effectiveness of elementary school principals, based the assessment of leader-member relations solely on the teachers' (i.e., group members') perceptions of group atmosphere.

Task structure. The variability encountered in the measurement of task structure was again more a function of the respondents than of instruments used. In seven of the 28 studies, the structure of the group task was assessed by the leaders' superiors in accordance with the method described in Fiedler (1967:28, 143; 1978: 65) and Fiedler and Chemers (1974:58, 66). The percentage of studies adhering to Fiedler's procedures for task structure was substantially lower than that for the leader-member relations dimension (25% and 61%, respectively). Thirteen of the remaining 21 studies operationalized task structure

in six different ways, a fact which shows that considerable variety characterizes researcher's views of not only who should assess the degree of structure but also whose task is being assessed. The following list of the deviant procedures attests to the lack of consistency in the operationalization of this situational variable:

1. group members regarding their own task (e.g., Jacobs, 1975)
2. group members regarding the leader's task (e.g., Bagley, 1972)
3. leaders regarding the task of the group members (e.g., Lanaghan, 1972)
4. leaders regarding their own task (e.g., Csoka, 1975)
5. leaders and group members regarding the leader's task (e.g., Van Gundy, 1975)
6. superiors regarding the leader's task (e.g., Hopfe, 1968)

In addition to these variant methods, three further studies operationalized task structure in terms of leader job training and/or experience.

Position power. The position power component of situational favourableness was the variable used least frequently in the CC:Other group. In ten of the 28 studies, position power was either assumed or held constant. Within the remaining studies, a variety of measurement procedures was again evident. However, unlike the task structure component, the power assessed was invariably that of the

leader. The assessment itself was made by different types of raters across the studies in a pattern similar to that found for task structure dimension. The assessors of leader position power included the group members (e.g., Stemler, 1980), the leaders themselves (e.g., Grunstad, 1972), and both the leaders and group members (e.g., Brown and Smolinsky, 1974).

Summary

The examination of the 112 primary analyses which constituted the "100 Series" revealed that only some studies had tested individual octants of the Contingency Model. This difference led to the creation of the "Octant" and "Non-Octant" groups of studies. The "Non-Octant" set (n=74) was deleted from the data base. All the "Octant" studies (n=38) were characterized by five criteria. Each study specifically:

1. tested one or more individual octants,
2. used a correlation statistic to report the results,
3. used the LPC instrument,
4. determined situational favourableness in an acceptable manner, and
5. measured leadership effectiveness.

The "Octant" studies adhering most and least closely to these criteria were designated, respectively, as "CC:Fiedler" and "CC:Other".

The distribution of studies, samples, and octant tests within each comparability category, partitioned by source is shown in Table 5.1. Relative to the number of

TABLE 5.1: DISTRIBUTION OF STUDIES, SAMPLES, AND OCTANT TESTS WITHIN COMPARABILITY CATEGORIES (BY SOURCE)

Source	CC[1]	Studies	Samples	Tests
Journals	F[2]	4	5	34
	O[3]	6	6	33
Dissertations/Theses	F	6	11	49
	O	20	23	105
Technical Reports	F	0	0	0
	O	2	4	28
Totals		38	49	249

1. CC = Comparability Category
2. F = Fiedler (CC:Fiedler)
3. O = Other (CC:Other)

studies within each category, those in CC:Fiedler yielded a higher proportion of octant tests than did those in CC:Other. However, twice as many octant tests were derived from CC:Other than from CC:Fiedler (n=166 and n=83, respectively). Moreover, the greatest amount of methodological variability came from all unpublished sources. The studies reported in journals showed an almost equal division between more and less adherence to Fiedler's method. A more detailed description of the distribution of

studies, samples, and tests — partitioned in several different ways — is shown in a series of four tables in Appendix G.

RELIABILITY OF CODING

One of the difficulties inherent in a study which relies heavily upon the printed word for its data is the consistency with which that word is interpreted. Since the amount of time required to code an octant study varied from three or four hours to four days, it was not possible to enlist the assistance of other judges to establish a measure of inter-rater reliability. In an attempt to compensate for the absence of such a measure and to lend greater credibility to the outcomes of this meta-analysis, four safeguards concerning coding reliability were used.

The first safeguard involved a series of mechanical checks each of which was used either to prevent the occurrence of coding errors or to identify any errors which had been made during the process of recording information from each study. This was done by requiring an entry for every variable on the coding instrument. The appearance of a blank space signalled a coding inaccuracy. The initial check for blank spaces, conducted at the conclusion of coding each document, involved a page-by-page check of the coding instrument. A second manual check for blanks was

made on the computer print-out of the data entry. The third check for blank spaces was made by instructing the computer to read a blank as "negative one" (-1) on the frequency runs for each code value of every variable. All such blank spaces and "negative ones" were checked against the appropriate document and the necessary corrections made. The final mechanical safeguard involved checking the frequencies of each value recorded for every variable. The purpose of this procedure was to identify out-of-range code values. These inaccuracies were also corrected before any of the analyses were conducted.

The second safeguard to ensure reliable coding was consultation with one or more members of the thesis committee. Such consultations took place when there was doubt regarding the value to be assigned to a particular variable. For purposes of consistent application across studies, a record was kept of the decisions resulting from these consultations.

While the mechanical checks and committee consultations were intended to ensure accuracy and consistency in coding all variables, the other two safeguards were concerned with the consistency of judgments made. It was in evaluating study quality (see pages 90 and 91) that researcher judgment came most into play. The consistency of these evaluative judgments was checked by comparing the early-assigned TV and MA scores with the

later-assigned ones and by comparing the TV and MA scores assigned, respectively, to journals and dissertations.

The results of these analyses are shown in Tables 5.2 and 5.3. These data show only small differences in the

TABLE 5.2: OCTANT STUDIES TV SCORE DATA

Group	$\bar{x}$	SD	n
Early	3.345	0.834	12
Late	3.383	0.570	26
Journals	3.415	0.777	12*
Dissertations	3.351	0.604	26

* includes two technical reports

TABLE 5.3: OCTANT STUDIES MA SCORE DATA

Group	$\bar{x}$	SD	n
Early	3.173	1.046	12
Late	3.104	0.683	26
Journals	3.175	0.994	12*
Dissertations	3.103	0.717	26

* includes two technical reports

values of the means. When the standard deviations are compared, there is more variance between than across the groups within the time and source variables. Moreover, the n's for the groups being compared are identical. There are two reasons which account for this pattern of cell sizes and dispersion. First, because dissertations typically provide much more information than do journal reports, there is greater consistency — and thus less variability — in the evaluative judgments. Second, the "early" coding (n=12) included three dissertations and nine journal articles; the "late" coding (n=26) included 23 dissertations, one journal article, and two technical reports. Apart from resulting in equal cell sizes, the variance in both "early-journal/technical report" and "late-dissertation" reflects the type of study most frequently coded during each time period. Based on the analysis of the TV and MA scores, it was concluded that neither time of coding nor source of document had any influence on the evaluation of study quality.

The MA score was also used to validate the classification of studies as either "CC:Fiedler" or "CC:Other". The rationale underlying the procedure was based on the assumption that the MA scores for studies which adhered more closely to Fiedler's method would be higher than the MA scores for studies which adhered less closely to Fiedler's method. In order to check this assumption, the

mean MA scores for each comparability category were compared. The results are shown in Table 5.4.

TABLE 5.4: MA SCORE DATA FOR CC:FIEDLER AND CC:OTHER

CC	n	$\bar{x}$	SD
CC:Fiedler	10	3.73	0.788
CC:Other	28	2.91	0.699

Based on the five point scale used in the evaluation of methodological adherence, the mean MA score for the CC:Fiedler studies is approximately 20% higher than the mean for the CC:Other studies. This outcome provides evidence which validates the assignment of studies to comparability categories. Lest it be thought that one of these variables was contaminated by the other, it should be noted that the MA score was computed for each study as it was coded. The comparability categories were created long after the coding was completed and their creation was done with no reference to the MA scores.

CONCLUSIONS AND DISCUSSION

The matter of study comparability, extensively reported in this chapter, came to play a more important role

in the study than had been originally anticipated. It was recognized from the beginning that sufficient similarity between and among studies was necessary if their findings were to be aggregated. What was not anticipated at the outset was the enormous methodological variability both across the studies themselves and between the studies and Fiedler's prescribed procedures. It had been assumed, perhaps too naively, that research spawned by Fiedler's Model would carefully follow the testing procedures he suggested. Moreover, it was believed — on the basis of the literature review — that the Model had been widely tested. Neither of these assumptions proved to be the case. That almost 60% (66/112) of the primary investigations in the "100 Series" did not test any octant of the Model suggests that the belief that the Model has been "extensively tested" (Fiedler, 1978:67) is erroneous. A more accurate statement would indicate that the Contingency Model has been extensively applied or widely used — but not extensively tested. That two comparability categories had to be created to accomodate only 38 studies attests to the methodological variability found in the manipulation of all three situational variables — most notably, in the task structure component.

The assessment of comparability among studies, which had originally seemed to be a relatively routine exercise, became almost a study in itself. It is believed that the

stringent criteria by which comparability was established and validated provided a solid foundation for the ensuing phases of the meta-analysis.

Chapter 6

THE EXTENT OF SUPPORT FOR FIEDLER'S MODEL: A DESCRIPTIVE ANALYSIS

The purpose of this analysis was to determine the amount of support the studies yielded for Fiedler's Model in order to ascertain the consistency of the LPC-performance relationship. The analysis was conducted in two phases. The first phase focussed on the Model as a whole; the second examined the octants within the Model. This chapter presents the method, procedures, results, and conclusions for each of the two phases.

PHASE I: THE MODEL

The first phase of the descriptive analysis involved an examination of study authors' conclusions regarding the degree to which their findings provided support for Fiedler's Model. This "author conclusion variable" yielded three possible responses: full, partial, or no support (see p. 26, No. 149 of the Coding Instrument in Appendix I). It is important to note that what was coded was the claim made by the author of the primary study itself, not a judgment made by the present researcher. In addition, it should be remembered that this analysis, like all ensuing ones, is based only on the 38 "Octant" studies described in the preceding chapter.

Procedure

The author conclusion variable was used in two ways. First, the degree of support was determined by simply counting the frequencies for each code value (full, partial, none) of the variable for all 38 studies. Second, in order to ascertain whether there was any particular pattern to the frequency of support, the data were partitioned in terms of the comparability category in which the study had been placed (Fiedler or Other), the task group type used in the research (interacting or coacting), and the setting in which the study was conducted (real-life or laboratory). The results for each analysis are presented in terms of both studies and samples. It will be recalled from the discussion in Chapter 4 that some authors used more than one sample of subjects in their research (the "multiple-sample" studies, n=49).

Frequency of Support Results

The data displayed in Table 6.1 show that the majority of studies claimed only partial support for the Model; a minority yielded full support. There was virtually no difference when the frequencies of full support and non-support were based on samples rather than studies.

TABLE 6.1: AUTHOR CONCLUSIONS REGARDING FREQUENCY OF SUPPORT FOR FIEDLER'S MODEL

Support Frequency[1]	Studies[2] n	%	Samples n	%
Full	6	15.0	8	16.3
Partial	25	62.5	32	65.3
None	9	22.5	9	18.4
Total	40	100.0	49	100.0

1. "Frequency of Support" refers here to the conclusion reached by the author(s) of the study irrespective of the magnitude and direction of the correlation(s) for the octant(s) tested.

2. Studies: n=40 rather than 38 because two studies were counted twice, each time using a different task group type.

Pattern of Support Results

As shown in Table 6.2, the "CC:Fiedler" studies yielded either full or partial support. All the non-supportive results came from "CC:Other". A comparison of study settings across the comparability categories revealed that partial support was reported more frequently by authors of real-life studies than by authors of laboratory studies (64% and 36%). When these same comparisons were made within comparability categories, there was very little difference in the frequencies of partial support from either laboratory or real-life studies within CC:Fiedler (30% and 70%, respectively) and CC:Other (40% and

TABLE 6.2: PARTITIONED FREQUENCY OF SUPPORT FOR FIEDLER'S MODEL BASED ON AUTHOR CONCLUSIONS

Comparability Category	Study Setting	Task Group Type[2]	Support Frequency[1] Full	Partial	None	Totals Studies
Fiedler	Real Life	1	1	4	0	5
		2	0	3	0	3
	Lab	4	1	3	0	4
Other	Real Life	1	1	1	5	7
		4	0	4	1	5
		2	0	2	1	3
		8	2	2	0	4
	Lab	1	0	0	1	1
		4	1	6	0	7
		8	0	0	1	1
Totals			6	25	9	40[3]

1. "Frequency of Support" refers here to the conclusion reached by the author(s) of the study irrespective of the magnitude and direction of the correlation(s) for the octant(s) tested.

2. Task Group Types: 1 = stated interacting; 2 = coacting; 4 = inferred interacting; 8 = nonacting (see also footnote on page 142).

3. Studies: n=40 rather than 38 because two studies were counted twice, each time using a different task group type.

60%, respectively). No meaningful comparison of the task group types was possible because of the very small cell sizes created by the stratification. Only two discernible patterns emerged from the partitioned data. First, support

is concluded more frequently by the authors of studies classified as CC:Fiedler. Second, supportive conclusions occurred more often in real life than in laboratory studies, regardless of comparability category.

Summary and Conclusions

Phase I of this descriptive analysis used the authors' stated conclusion to examine the frequency of support for Fiedler's Contingency Model. The author conclusion variable was analyzed for all studies together and for the studies partitioned by comparability category, study setting, and task group type. In summary, the results showed that:

1. across all studies, the conclusion of partial support was more frequent than that of either full support or non-support;
2. support was concluded more frequently in studies categorized as "Fiedler" than in those categorized as "Other";
3. studies conducted in real-life settings tended to yield supportive conclusions more frequently than did those in laboratory settings.

On the basis of these results, it could be concluded that the Model lacks firm support. However, such a conclusion — apart from being premature at this stage of the analysis — may be misleading. Fiedler's Model is composed of eight separate octants each representing a different combination of the three situational variables. Only when a study tests all eight octants, can the author

legitimately draw Model-relevant conclusions. Studies which test fewer than eight octants can lead only to octant-relevant conclusions. Given the unique character of each octant, these conclusions should not be used to make inferences about the Model as a whole. It is precisely because most of the 49 samples in the meta-analysis data base tested only two (n=17 samples) or four octants (n=16 samples) that the above results should not be used as the foundation for Model-relevant conclusions. Thus, the most defensible conclusion to emerge from the analysis of the author conclusion variable relates not to a declaration of support or lack of support for the Model but rather to the need for an examination of the findings for each octant as a separate unit of analysis. This conclusion led to the second phase of the descriptive analysis.

PHASE II: THE OCTANTS

When comparisons were made between authors' conclusions regarding frequency of support and the correlation(s) upon which those conclusions were based, it was found that different authors reached different conclusions on the basis of similar data. Take, for example, a hypothetical — but not atypical — study which examined two octants using two measures of effectiveness and found that three out of four LPC-performance correlations

were in the direction predicted by Fiedler but none was statistically significant at the stated level. On the basis of these results, the researcher could legitimately reach any one of three possible conclusions: full support, partial support, or no support for the Model. The choice among the three conclusions could be influenced by at least three factors: the Contingency Model itself, the way in which the research questions or hypotheses had been stated, and the epistemological stance assumed by the author. For instance, one researcher, given these hypothetical data, may conclude "no support" on the basis of nonsignificant findings. Another may declare "partial support" because of the direction of the three correlations. A third investigator, having hypothesized agreement with one octant (e.g., III) but disagreement with the other (e.g., VII), may decide in favour of a "full support" conclusion. The existence of these conflicting conclusions, together with the results of the author conclusion variable analysis, made it clear that octant-based conclusions would be possible only by separating study findings from their authors' interpretations.

Method

In an attempt to introduce consistency across the studies, an examination of the signs of the reported correlations was conducted. The question arose as to which

correlation should be used. Fiedler's own suggestion for correlating the LPC score with the performance measure, was that the Spearman rho be used. Although he has not specifically stated since then that the Pearson r could or should be used, it was the one reported in over half the octant tests which formed the the data base for this meta-analysis. The Pearson r was the preferred statistic for 137 tests while the Spearman rho was chosen for 132 tests.

It will be noted that the sum of these two frequencies is greater than the number of octant tests (n=249). This apparent discrepancy exists because four studies, representing 20 octant tests, reported both correlations for each test (i.e., Butterfield, 1968; Brown and Smolinsky, 1974; Van Gundy, 1975; Gilbert, 1977). Because the retention of these "multiple statistic" tests had the potential to contaminate all ensuing analyses, it was necessary to separate the values of Pearson r from those of Spearman rho. To this end, four separate correlation listings were made. The first listing, "r + rho", contained the values of r from studies reporting only r, the values of rho from studies reporting only rho, and the values of r from studies reporting r and rho. The second listing, "rho + r", included the values of rho from studies reporting only rho, the values of r from studies reporting only r, and the values of rho from studies reporting rho and r. The third

listing, "r only", contained only the values of r. The fourth listing, "rho only", contained only the values of rho. The number of correlations in each listing is shown in Table 6.3.

TABLE 6.3 NUMBER OF REPORTED CORRELATIONS IN EACH LISTING (BY OCTANT)

CORR'N LISTING		OCTANT I	II	III	IV	V	VI	VII	VIII	ROW TOTAL
r+rho	n	52	14	44	32	32	11	35	29	249
rho+r	n	52	14	44	32	32	11	35	29	249
rho	n	29	11	22	12	22	9	17	10	132
r	n	28	5	27	21	14	3	20	19	137
Fiedler	n	8	3	12	10	6	0[1]	12	12	63

1. There were no antecedent correlations in Octant VI.

The choice of which listing to retain was made for by three reasons, the first of which was the number of within-octant correlations in each listing. The data in Table 6.3 show that the use of either "rho only" or "r only" would have substantially reduced the data base for the meta-analysis. These two listings were therefore given no further consideration. The second and third reasons, therefore, concern the selection of one of the two combined listings ("rho + r" or "r + rho"). That a choice had to made between them was caused not only by the multiple

statistical tests but also by the emergence of conflicting signs between the r and rho correlations for the same octant test. Of the four studies in which this occurred, only one author provided any explanation for the reversal in direction. Gilbert noted that:

> Each octant includes one or two spurious points away from the visual line of best fit. These spurious points had the effect of reducing the Pearson product-moment correlation while rho seemed less affected. The effect of these extreme cases was especially evident in Octant VII where one point ... produced a negative Pearson product-moment correlation while the rank-order correlation was positive (1979:102).

It can be only speculated that the reversal of direction found in the other three studies may also have been the result of extreme values. Gilbert's observation, together with the lack of any other plausible explanation for the conflict in signs, suggested that the values of rho rather than r be used. The third reason for choosing the "rho + r" listing rests with the Contingency Model itself. As noted earlier, Fiedler recommends the use of the Spearman rank-order correlation. Of the 249 correlations on the "rho + r" listing, 132 are Spearman rho and 117 are Pearson r (112 and 137, respectively, on the other combined listing). Collectively, these three reasons constituted the rationale for the use of the "rho + r" listing. All ensuing analyses in this chapter and in Chapters 7, 8, and 9 are based only

on the "rho + r" listing of reported correlations[1].

The analysis of the correlation signs within the "rho + r" listing was dichotomous in nature, assigning support or non-support for the octant(s) tested on the basis of congruency with the direction predicted by Fiedler. Using "directional congruency" as the criterion, the frequency of support and non-support was analyzed in two ways: first, for all octant tests together; second, for the octant tests partitioned into comparability categories and task group types. The results of each analysis are reported in that order.

Frequency of Support for all Octant Tests

It was stated earlier (p. 132) that any conclusion concerning support for the Model as a whole is valid only when all eight octants have been tested. Because the application of the directional congruency criterion to each of the 249 LPC-performance correlations produced results for every octant, it became possible to assess support for both individual octants and the whole Model.

The data shown in Table 6.4 indicate that 54% (135/249) of the correlations had signs in the direction predicted by Fiedler; the remaining 114 (46%) were in the

---

[1] The four listings of reported correlations, and a list of various groupings of the correlations, are shown in Appendix H.

TABLE 6.4: FREQUENCY OF SUPPORT[1] FOR OCTANTS

	Octant								Total
	I	II	III	IV	V	VI	VII	VIII	
Support	36	7	22	17	21	6	13	13	135
Non-support	16	7	22	15	11	5	22	16	114
Total	52	14	44	32	32	11	35	29	249

1. "Support" is defined on an octant basis as directional congruence between the correlation reported by the researcher, on the one hand, and that predicted by Fiedler, on the other (i.e., "support" = sign in predicted direction; "no support" = sign in direction opposite to prediction); statistical significance of reported correlation was not considered.

opposite direction from Fiedler's predictions. This finding again highlights the misleading nature of conclusions which declare support for the Model. While the sign analysis revealed that just over half the tests were supportive across all octants, a different conclusion was suggested when the results were examined within individual octants. In five out of the eight octants (II, III, IV, VI, VIII), there was either no difference or a difference of three or less between the frequency of supportive and non-supportive tests. Of the remaining three, Octants I and V emerged with some support (69% and 66%, respectively) while Octant VII showed 63% non-support.

These within-octant results do not lead to a conclusion which suggests support for the Model. The best

that can be said, at least at this stage of the analysis, is (1) that Octants I and V seem to be supported; (2) that Octant VII remains in doubt because Fiedler's predicted correlation of 0.05 is so close to zero that it could become negative simply as a result of sampling fluctuation; and (3) that Octants II, III, IV, VI, and VIII are neither supported nor unsupported. The possibility that these within-octant outcomes could change is explored in the next sections dealing with comparability categories, task group types, and study quality.

Frequency of Support by Comparability Category

The purpose of this analysis was to ascertain the frequency of support within the comparability categories both for individual octants and all octants together. The results of the analysis are shown in Table 6.5. There are two important points to be made regarding both the within-octant and across-octant frequencies. First, in the "Fiedler" studies, the frequency of support in all but one of the octants is greater than the frequency of non-support. Across the octants, support outweighs non-support by a margin of 57 to 26. Second, and by contrast, in the "Other" category of the studies non-support is more frequent than support for five octants (II, III, IV, VII, and VIII). Overall, 78 tests were supportive while 88 were non-supportive in this category.

TABLE 6.5: FREQUENCY OF SUPPORT[1] FOR OCTANTS WITHIN COMPARABILITY CATEGORIES

CC	FS[2]	OCTANT								TOTALS	% of CC	% of TOTAL
		I	II	III	IV	V	VI	VII	VIII			
F	S	20	3	7	8	6	1	7	5	57	69	23
	NS	4	1	6	3	1	3	4	4	26	31	11
O	S	16	4	15	9	15	5	6	8	78	47	31
	NS	12	6	16	12	10	2	18	12	88	53	35
TOTAL		52	14	44	32	32	11	35	29	249		100

1. "Support" is defined on an octant basis as directional congruence between the correlation reported by the researcher, on the one hand, and that predicted by Fiedler, on the other (i.e., "support" = sign in predicted direction; "no support" = sign in direction opposite to prediction); statistical significance of reported correlation was not considered.

2. Abbreviations: FS = Frequency of support; S = Support
NS = Non-support; F = CC:Fiedler;
O = CC:Other

Frequency of Support by Task Group Types

The purpose of this analysis was to determine whether there were any differences in the frequency of support among

the four task group types[2]. The data displayed in Table 6.6 show that there were four octants (III, IV, VII, and VIII) containing some cells in which the non-supportive tests outnumbered the supportive ones. In Octants IV and VIII, the non-support came from task groups which were stated by the author to be interacting (task group type 1). In Octant VII, part of the incongruency also came from task group type 1 but the majority was derived from "nonacting" groups (task group type 8: those groups for which leadership effectiveness was assessed on the basis of leader rather than group performance). In Octant IV, the higher non-support frequency came from task groups stated by the primary author to be interacting (task group type 1).

It is interesting to note, on the one hand, that task groups which were inferred to be interacting showed more support than non-support across all octants (61% and 39%, respectively). On the other hand, more non-support (55%) than support (45%) emerged from the 101 tests within the set of studies whose authors declared the task group to be interacting. These percentages were almost identical with

---

[2] Classification of Task Group Types: TGT1 = stated by the primary author to be interacting; TGT4 = inferred by the present researcher to be interacting; TGT2 = stated by the primary author to be coacting; TGT8 = groups for which leadership effectiveness was assessed on the basis of leader rather than group performance.

TABLE 6.6: FREQUENCY OF SUPPORT FOR OCTANTS WITHIN TASK GROUP TYPES

CC	TGT[1]	FS	Octant I	II	III	IV	V	VI	VII	VIII	Total
	1	S	14	4	8	7	4	2	4	2	45
	1	NS	11	4	4	11	4	1	7	14	56
	4	S	7	2	3	8	6	4	5	10	45
	4	NS	4	2	11	4	2	2	3	1	29
F+O	2	S	5	0	3	0	4	0	3	0	15
	2	NS	0	0	3	0	0	0	1	0	4
	8	S	10	1	8	2	7	0	1	1	30
	8	NS	1	1	4	0	5	2	11	1	25
	All	S	36	7	22	17	21	6	13	13	135
	All	NS	16	7	22	15	11	5	22	16	114

1. Abbreviations: CC = Comparability category
TGT = Task group type
FS = Frequency of support
F+O = Combined comparability categories
S = Support; NS = Non-support

those for all octant tests together (support=54%, non-support=46%, n=249). Moreover, when the within-task group type frequencies were compared with those in the comparability categories, the same four octants (III, IV, VII, VIII) contained cells which showed more non-support than support (see Table 6.5). It would appear from a comparison of the four task groups types that the majority of non-supportive tests came from task groups stated to be interacting within the studies classified as CC:Other.

By contrast, the tests within the other four octants (I, II, V, and VI) were generally more supportive than

non-supportive. Of all the supportive tests in Octant I, the two highest percentages of support came from task groups stated to be interacting (39%) and from "nonacting" groups (28%). Of all the supportive tests in Octant V, however, the "nonacting" groups yielded the greatest proportion of support (33%), followed by groups inferred to be interacting (29%). Because both Octants II and VI had "missing data"[3], they were not included in this within octant analysis. The analysis has revealed that differences exist among the task group types with respect to the amount of support generated for the octants.

Summary of Results

Phase II of the descriptive analysis examined the frequency of support for Fiedler's Contingency Model in terms of the directional congruency of 249 octant tests. In summary, the results showed that:

1. supportive results were more frequent than non-supportive results (support=135, non-support=114)

2. within CC:Fiedler, supportive results were more frequent than non-supportive results (support=57, non-support=26)

3. within CC:Other, non-supportive results were more frequent than supportive results (non-support=88, support=78)

---

[3] There were no reported tests for coacting groups (TGT 2) in either octant nor for nonacting groups (TGT 8) in Octant VI.

4. supportive results were more frequent than non-supportive results within coacting groups (support=15, non-support=4), inferred interacting groups (support=45, non-support=29), and nonacting groups (support=30, non-support=25)

5. within stated interacting groups, non-supportive results were more frequent than supportive results (non-support=56, support=45)

It would be premature at this stage to conclude that Fiedler's Model is, or is not, supported. It must be remembered that this analysis was based solely on the direction of the reported correlations. No firm conclusion regarding the efficacy of Fiedler's Model can be made without considering the magnitude of the correlations.

CONCLUSIONS

There are two important points to emerge from the descriptive analysis reported in this chapter. The first point concerns the basis upon which the Model is deemed to be supported or not supported. The analysis in this chapter has made it apparent that neither an author's conclusion nor a study testing fewer than eight octants should be used as the basis upon which to assess the support for the Model.

The second point arises directly from the previous one. Because each test result is specific to one of eight unique combinations of leader-member relations, task structure, and position power, the 249 correlations which comprise the data base cannot be aggregated as if those

eight octants represented the same level of favourableness. A sound basis for any decision concerning support for the Model requires not only that the reported correlations be aggregated within octants but also that the magnitude of each correlation be taken into account. This potentially more powerful analysis forms the substance of the next chapter.

Chapter 7

THE COMPARISON OF OBTAINED AND PREDICTED VALUES: AN ANALYSIS OF MEDIAN CORRELATIONS

This second stage of the meta-analysis differed quite markedly from the descriptive phase reported in the second part of the preceding chapter. Rather than relying solely on correlation signs, this analysis took into account both the magnitude and direction of the reported correlations. By aggregating all reported values for a given octant to yield one composite value, it was possible to compare Fiedler's predicted values with the values obtained from the partitioned data sets.

The common metric selected for this analysis was a median correlation[1]. There are two reasons underlying this choice. First, as was discussed in Chapter 1, Fiedler's predictions for each octant were the median values derived from the LPC-performance correlations reported in the antecedent studies. Second, Fiedler has subsequently used these antecedent correlations to determine the extent of agreement with those values derived from the validation studies (Fiedler and Chemers, 1974:82; Fiedler, 1978:69).

---

[1] Although the mean correlation would have been a more powerful method, its use would have precluded a direct comparison with Fiedler's median correlations.

This chapter is divided into four sections. The first section compares the obtained median correlations with those predicted by Fiedler. The second section presents the major findings from the analysis of the two stratification variables (comparability categories and task group types). The third section reports the results yielded by the consideration of covariance in terms of seven selected study characteristics. The details specific to each of these analyses are explained in conjunction with the presentation of the findings. The last section summarizes the results.

COMPARISON OF THE OBTAINED AND PREDICTED MEDIAN CORRELATIONS

The purpose of this comparison was to show the extent of agreement between the median correlations predicted by Fiedler and those derived from the "rho + r" listing of reported correlations (see Chapter 6, pp. 134-138). These obtained median values, together with Fiedler's predictions, are shown in Table 7.1 and depicted graphically in Figure 7.1[2].

---

[2] A comparison of median correlations from all four correlation listings is reported in Appendix F (pp. 349-350).

TABLE 7.1: OBTAINED MEDIAN CORRELATIONS FROM THE "rho + r" LISTING

CORR'N LISTING	OCTANT I	II	III	IV	V	VI	VII	VIII
rho+r	-.165	-.010	-.050	.055	.245	.070	-.038	.090
n	52	14	44	32	32	11	35	29
Fiedler	-.52	-.58	-.33	.47	.42	(+)[1]	.05	-.43
n	8	3	12	10	6	0[2]	12	12

1. Fiedler (1967) predicted only direction for Octant VI.

2. There were no antecedent correlations in Octant VI.

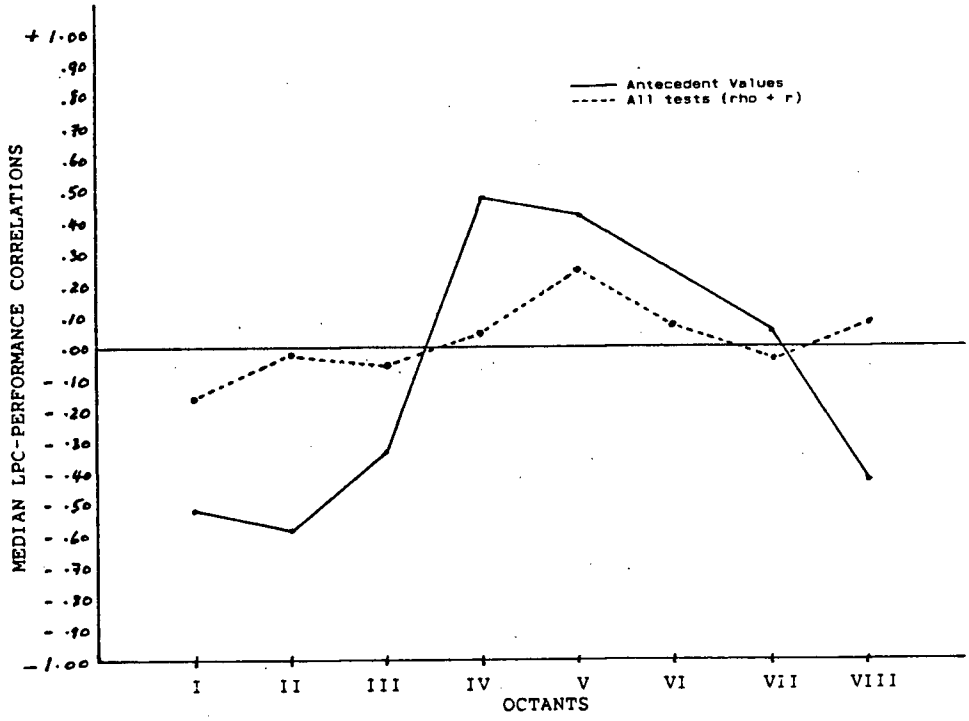


FIGURE 7.1: EXTERNAL CONGRUENCY: All Octant Tests (rho + r listing)

The term "external congruency" denotes the correspondence between the obtained medians and those predicted by Fiedler (the antecedent values). The visual display clearly reveals a lack of external congruency. Although six of the eight obtained medians correspond directionally with the antecedent values, there appears to be a considerable difference in the magnitude in five of the eight medians — specifically, Octants I, II, III, IV, and VIII. These differences have produced a curve which is quite different in shape from the one yielded by Fiedler's predicted value for each octant.

RESULTS FOR THE STRATIFICATION VARIABLES

This section reports the results yielded by a comparison of the median correlations in the two comparability categories and in the four task group types. To derive the median values, the 249 reported correlations from the "rho + r" listing were first partitioned into the "Fiedler" and "Other" comparability categories (CC:Fiedler and CC:Other) and, second, into the task group type (stated or inferred interacting, coacting, or nonacting) used in the primary study. The results from the analysis of each stratification variable are presented in the order stated above.

Comparison of Median Correlations by Comparability Category

The median correlations for each comparability category are displayed in Table 7.2. In "CC:Fiedler", seven

TABLE 7.2: COMPARISON OF MEDIAN CORRELATIONS FOR COMPARABILITY CATEGORIES

CORR'N LISTING	CC	OCTANT I	II	III	IV	V	VI	VII	VIII
rho + r	F	-.335	-.275	-.050	.240	.280	-.370	.300	-.097
rho + r	O	-.115	.120	-.050	-.080	.220	.100	-.085	.115
Fiedler[1]		-.52	-.58	-.33	.47	.42	(+)[2]	.05	-.43
rho + r	F	24	4	13	11	7	4	11	9
n's	O	28	10	31	21	25	7	24	20
Fiedler	n	8	3	12	10	6	0	12	12

1. The median correlations predicted by Fiedler's Contingency Model.

2. Fiedler (1967) predicted only direction for Octant VI.

out of eight octants yielded median correlations which were directionally congruent with Fiedler's predictions. In "CC:Other", directionally congruent medians emerged in only four octants (I, III, V, and VI). Of the seven directionally congruent median correlations in CC:Fiedler, six are of greater magnitude than the corresponding median values in CC:Other. For example, in Octant I, the CC:Fiedler median is -0.335; the value for CC:Other is -0.115. When the directionally congruent median correlations in both CC:Fiedler and CC:Other are compared with the magnitude of Fiedler's predicted values, it is the CC:Fiedler medians which tend to approximate more closely the size of the antecedent values (see third row of Table 7.2). A graphic comparison of the three sets of results is shown in Figure 7.2.

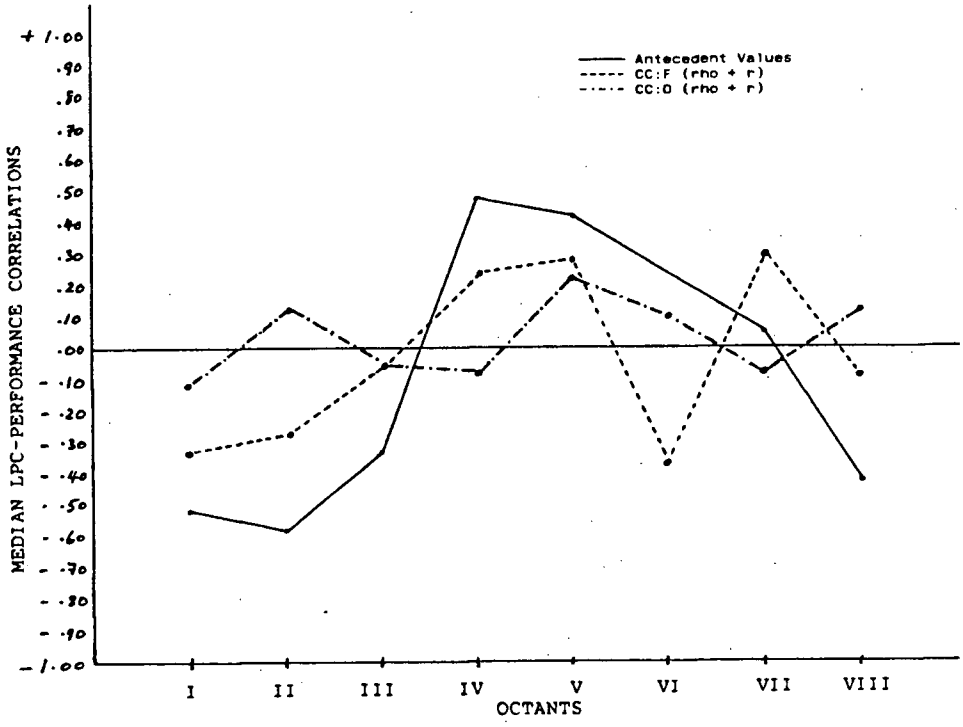


FIGURE 7.2: INTERNAL AND EXTERNAL CONGRUENCY: CC:FIEDLER and CC:OTHER

This visual display clearly indicates the lack of correspondence between the two comparability categories themselves. This disparity is referred to as "internal congruency" on the figure. The "CC:Fiedler" curve, however, reveals more external congruency than does the "CC:Other" curve. In other words, the median correlations generated from the CC:Fiedler studies correspond more closely to Fiedler's predicted values than do the medians from the CC:Other studies.

There are two important observations to be made about these results. First, the greater and lesser correspondence of studies classified, respectively, as CC:Fiedler and CC:Other confirms the validation of the comparability category designation reported in Chapter 5. Second, since the CC:Fiedler studies not only followed more closely the procedures prescribed by Fiedler but also yielded the stronger results, it appears as if there may be a positive relationship between methodological adherence and results which are supportive of the Contingency Model.

Comparison of Median Correlations by Task Group Type

The median correlations for each task group type are shown in Table 7.3. There are two important points to be noted about these data. First, none of the medians for the four task groups is directionally congruent with Fiedler's predictions across all eight octants. The direction of the

TABLE 7.3: COMPARISON OF MEDIAN CORRELATIONS[1] FOR THE TASK GROUP TYPES

CORR'N LISTING	TGT	OCTANT I	II	III	IV	V	VI	VII	VIII
rho+r	1	-.040	-.010	-.175	-.065	-.345	.070	-.140	.165
	4	-.120	-.275	.145	.125	.265	.115	.090	-.332
	8	-.310	-.180	-.555	.165	.050	-.425	-.365	-.285
	2	-.320		-.040		.505		.260	
Fiedler[2]		-.52	-.58	-.33	.47	.42	(+)[3]	.05	-.43
rho+r	n 1	25	8	12	18	8	3	11	16
	n 4	11	4	14	12	8	6	8	11
	n 8	11	2	12	2	12	2	12	2
	n 2	5	0[4]	6	0[4]	4	0[4]	4	0[4]
Fiedler	n	8	3	12	10	6	0	12	12

Abbreviations: TGT 1 = stated interacting
TGT 4 = inferred interacting
TGT 8 = nonacting
TGT 2 = coacting

1. The n for each within octant median is shown in the bottom rows of the table.

2. The median correlations predicted by Fiedler's Contingency Model.

3. Fiedler (1967) predicted only direction for Octant VI.

4. There were no reported correlations in Octants II, IV, VI, and VIII for TGT 2.

median correlations for groups whose interacting nature had to be inferred (TGT 4) corresponds with that predicted in seven out of eight octants. Six of the eight medians are directionally congruent for nonacting groups (TGT 8). Only half of the medians for groups stated to be interacting (TGT 1) are in the direction predicted by Fiedler.

The second point concerns the relative magnitude of the obtained and predicted medians. Although 21 of the 28 median correlations shown in Table 7.3 are directionally congruent, only eight are both in the predicted direction and approximate the predicted magnitude. These eight include: (1) Octant III in the stated interacting groups; (2) Octants V, VII, and VIII in the inferred interacting groups; (3) Octants I and VIII in the nonacting groups; and (4) Octants I and V in the coacting groups. Looking at this result on a within-octant basis, only three octants (I, V, and VIII) yield median correlations which meet the dual criteria of direction and magnitude for more than one task group type. There are no octants in which all four task groups are congruent with respect to both criteria. Collectively, these results do not offer strong support for the Contingency Model[3].

---

[3] A detailed analysis of the results for the "r only" and "rho only" listings, together with a tabular display of the "r + rho" results, appear in Appendix F (pp. 351-354).

The relationship among the task group types themselves (internal congruency) and between Fiedler's predicted curve and the curves generated for each task group (external congruency) are shown in Figure 7.3.

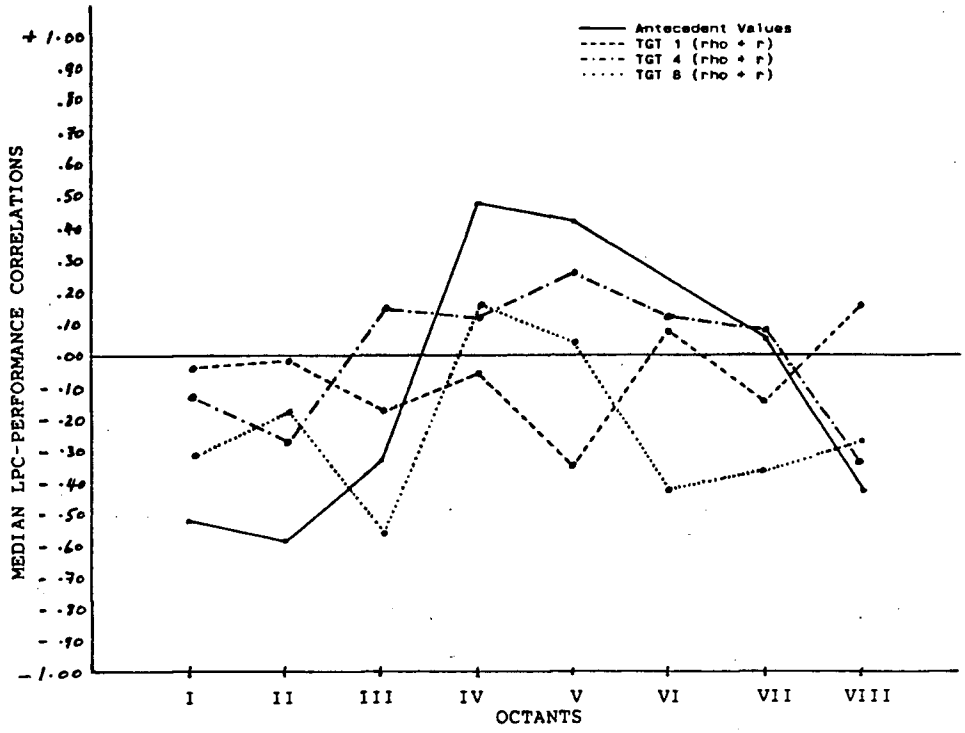


FIGURE 7.3: INTERNAL AND EXTERNAL CONGRUENCY:
Task Group Types

There is no curve for coacting groups (TGT2) because of the empty cells in four octants. The curves yielded by the median correlations in the remaining three task groups show very clearly the lack of internal congruency. The only two octants which appear to be congruent both internally and externally are Octants I and II. While the lack of correspondence among the task group types is interesting and potentially useful, a more important outcome to emerge

concerns the nonacting groups (TGT 8). Typically, the LPC-performance correlations for these tests were based on a "direct" evaluation of the leader's effectiveness rather than an "indirect" appraisal based on the performance of either the group members or the group members and the leader. This latter procedure is the one prescribed by Fiedler. It would appear, at least on the basis of the graphic representation, that the predictive validity of the Contingency Model might be improved by using a direct assessment of leader performance.

RESULTS FOR THE STUDY CHARACTERISTIC VARIABLES

The octant tests were also partitioned on the basis of seven selected study characteristics. In meta-analytic terms, these characteristics are regarded as covariates and are analyzed to determine what influence, if any, they may have had on study results (Glass, 1978; White, 1979). Before discussing the study characteristics, it is necessary to explain why these seven factors were designated as covariates while comparability category and task group type were not. It is clear that both of them were "covariates" in the sense that each one was examined in terms of its possible effect on the amount of support generated for the Model. The distinction between the sets of variables is an analytical one and lies not in kind but in usage.

The two stratification variables were each used for a different reason. The reason for differentiating between the task group types arose from the Fiedler's taxonomy of groups which suggests that the nature of the relationships among group members varies as a consequence of the task the group is required to perform (1967:17). The comparability categories evolved in reponse to the methodological variability among the studies. The purpose for retaining them was linked to the credibility of the aggregated findings. In other words, these variables were included primarily for purposes other than accounting for variablility across study results. That purpose was served specifically by the seven selected study characteristics. These variables are as follows:

1. basis for assessing leadership effectiveness,
2. study setting,
3. organizational setting,
4. academic discipline,
5. Fiedler associate,
6. source of document, and
7. study quality.

As White has observed, "there is no fail-safe technique for making sure that all of the proper variables are included in a meta-analysis" (1979:10). Nor does there appear to be an established procedure to guide the selection of variables which may have influenced the outcomes of certain studies in the data base. The ultimate choices are perhaps more a product of evolving impression than of fixed practice. Of the seven which were chosen as covariates in

the present study, two were prompted by the Fiedler literature (study and organizational settings), three by the meta-analysis literature (source of document, associate of Fiedler, and study quality) and two by the aggregated findings themselves (basis for assessing leadership effectiveness and academic discipline of primary author). Each characteristic will be described, together with the specific reason for its selection, as the results are presented. A cautionary note must be attached to the ensuing results. The octant tests included in the analysis of any one of the variables are representative of both comparability categories and therefore of all task group types.

The Basis for Assessing Leadership Effectiveness

This covariate was suggested by the aggregated findings — specifically, the median correlation analysis of task group types. The visual data display in Figure 7.3 (p.156) showed that the obtained values for nonacting groups corresponded most closely with Fiedler's predicted medians. This observation raised a question concerning the relationship between the method of determining leadership effectiveness and congruency with the predicted curve. According to the Model, effectiveness should be assessed on the basis of group performance yet about one third of the octant tests (n=87/249) used leader performance instead.

The obtained values derived from both methods of assessing leadership effectiveness are shown in Table 7.4. Apart from

TABLE 7.4: MEDIAN CORRELATIONS FOR GROUP AND LEADER PERFORMANCE AS BASIS FOR ASSESSING LEADERSHIP EFFECTIVENESS (LE)

Basis For LE		OCTANT I	II	III	IV	V	VI	VII	VIII
Grp Perf		-.132	-.111	-.035	-.010	.245	.115	.080	.095
Ldr Perf		-.172	.140	-.405	.090	.230	-.200	-.145	-.090
Fiedler[1]		-.52	-.58	-.33	.47	.42	(+)[2]	.05	-.43
Grp Perf	n	35	7	28	27	16	6	19	24
Ldr Perf	n	17	7	16	5	16	6	16	5
Fiedler	n	8	3	12	10	6	0	12	12

1. The median correlations predicted by Fiedler's Contingency Model.

2. Fiedler (1967) predicted only direction for Octant VI.

the very low correlations (i.e., <0.10) in six cells, there are two important points to be made about these data. First, there is directional agreement between the median correlations for group performance and leader performance in only three octants (I, III, and V). In other words, the tests in more than half the octants indicate that either high or low LPC leaders may be shown as effective under the same condition of situational control depending on the basis used to assess effectiveness. In the five octants showing

opposite signs, three from group performance (II, VI, and VII) and two from leader performance (VI and VII) are in the direction predicted by Fiedler. Given that directional congruency is present in six octants for group performance and five for leader performance, it could be concluded that the basis upon which leadership effectiveness is assessed makes very little difference to the outcomes.

The second point concerns the magnitude of the median correlations. Considering only the directionally congruent pairs with values >0.10, just two octants (I and V) have yielded medians which are reasonably similar. However, a comparison of these same obtained values for group performance and leader performance with those predicted by Fiedler indicates that all four of them are considerably different from the antecedent values. When both magnitude and direction are considered, only the leader performance-based result in Octant III (-0.405) is close to prediction (-0.33). The extent of internal and external congruency yielded by these data is displayed in Figure 7.4.

It is readily apparent from this figure that the curve yielded by the leader performance-based assessment of leadership effectiveness bears closer resemblance to the predicted one than does the group performance based curve. Taken together, the leader performance-based curve and that of nonacting groups (see TGT 8 in Figure 7.3, p.156) add

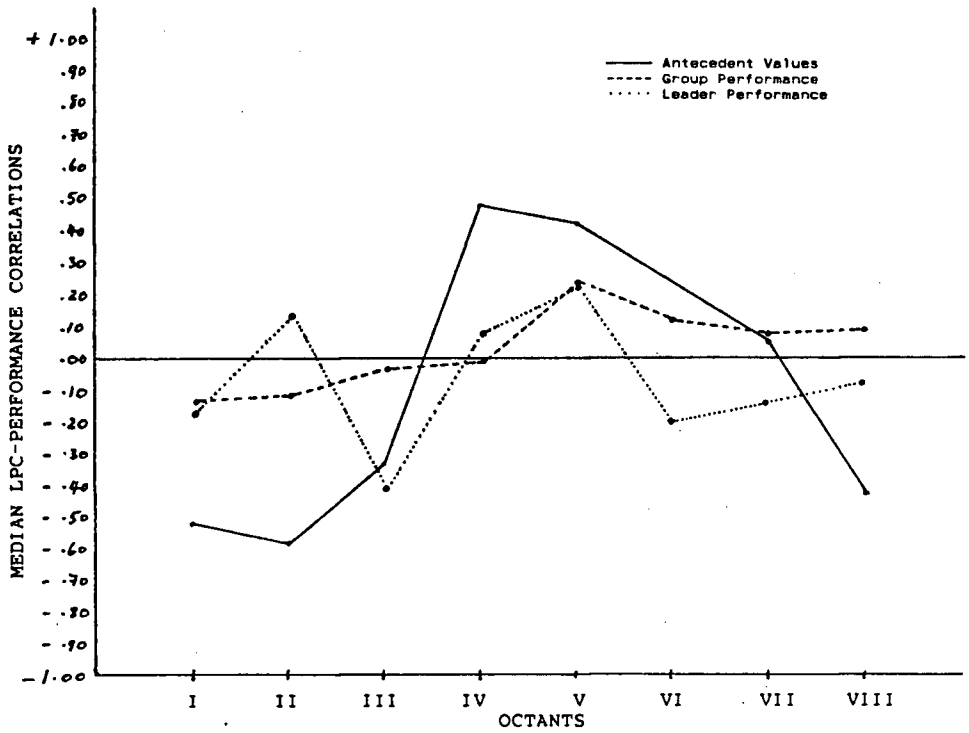


FIGURE 7.4: INTERNAL AND EXTERNAL CONGRUENCY:
Group and Leader Performance

some credibility to the observation that effectiveness may be predicted more reliably if leader performance were evaluated directly rather than indirectly. These findings are consistent with the results of Rice's (1975) study (see Chapter 3, p. 65) which showed a consistently stronger

relationship between LPC and leader performance than between LPC and group performance.

Study Setting

This second covariate was selected for analysis on the basis that short-lived, ad hoc groups assembled for leadership studies are not the same as groups which exist in real-life, ongoing organizations. Fiedler has noted that

> ... the Model can be more meaningfully evaluated as a predictor if natural groups and ad hoc groups in laboratory experiments are considered separately, as well as together (Fiedler, 1971d:132).

This statement suggested that real-life (i.e., field) and laboratory studies should be examined separately.

The data-base contains 187 real-life tests and 62 laboratory tests. The median correlations yielded by the aggregated findings are shown in Table 7.5. An inspection of the directional congruency for each setting casts some doubt on the predictive capability of the Model. Within the real-life setting, only four octants (I, II, III, and V) have signs which concur with the directional predictions. The laboratory tests show five congruent octants (I, IV, VI, VII, and VIII). The two settings together show agreement in signs only for Octant I. For seven out of eight octants, it would appear that the leadership style associated with effectiveness may not be contingent upon the situational favourableness but rather upon the setting in which the research was conducted. For example, when Octant VIII is

TABLE 7.5: MEDIAN CORRELATIONS FOR REAL LIFE AND LABORATORY SETTINGS

Study	OCTANT							
Setting	I	II	III	IV	V	VI	VII	VIII
Real Life	-.167	-.020	-.097	-.065	.245	-.200	-.140	.155
Lab	-.130	.110	.027	.265	-.140	.115	.085	-.332
Fiedler[1]	-.52	-.58	-.33	.47	.42	(+)[2]	.05	-.43
Real Life n	43	9	33	22	26	5	29	20
Lab n	9	5	11	10	6	6	6	9
Fiedler n	8	3	12	10	6	0	12	12

1. The medians correlations predicted by Fiedler's Contingency Model.

2. Fiedler (1967) predicted only direction for Octant VI.

tested in the laboratory, the more effective leaders have low LPC scores but, in ongoing organizations, they have high scores on the LPC scale. Only under Octant I conditions might a low LPC leader be effective in both settings. It is also possible, given the weakness of three median correlations in real-life Octants II, III, and IV and in laboratory Octants III and VII, that either leadership style — high or low LPC — may be effective. The data for the two study settings are displayed graphically in Figure 7.5.

A comparison of the laboratory and antecedent curves suggests, with the exception of Octants VII and VIII, that either laboratory settings are too artificial to yield

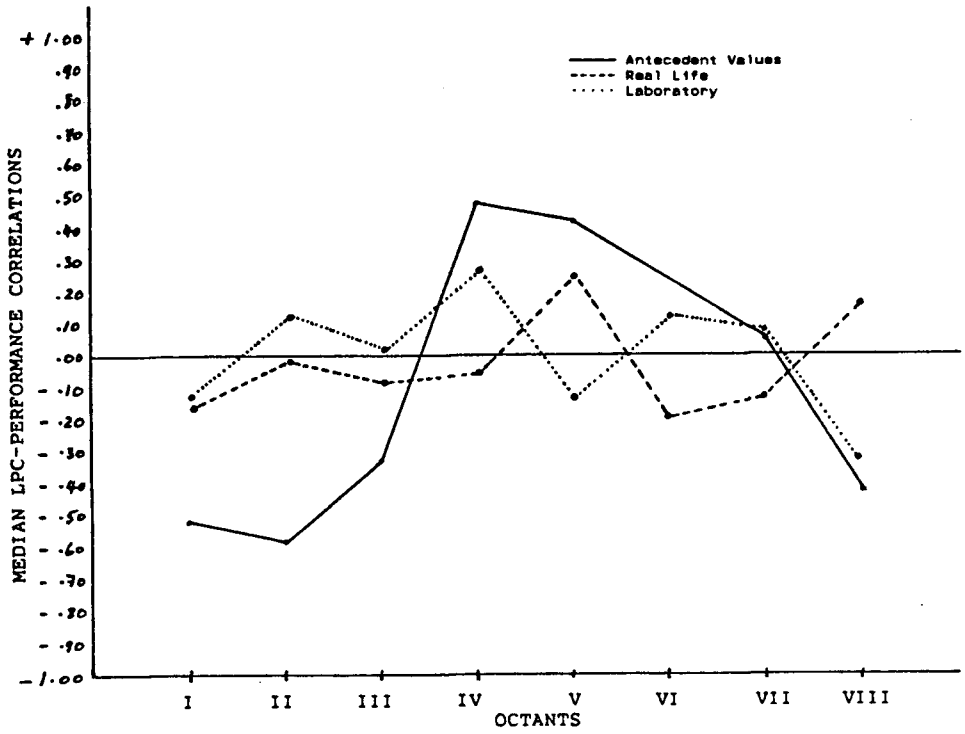


FIGURE 7.5: INTERNAL AND EXTERNAL CONGRUENCY:
Real Life and Laboratory

meaningful results or there are other unspecified factors which mitigate against successful validation attempts. The results yielded by the real-life tests are by no means conclusive. In terms of external congruency, there is some similarity between the real-life and antecedent curves but not enough to permit use of the Model to predict which leadership style will be effective in what combination of situational variables — except when the organizational conditions match those specified for Octant V and, perhaps, Octant I.

Organizational Setting

This third covariate was chosen for analysis in response to one of of the criticisms of Fiedler's Model. A number of researchers have charged that the Model lacks generalizability (e.g., Schriesheim and Kerr, 1977:13). More specifically, Vecchio (1977) has suggested — in connection with his attempt to explain the successful validation study by Chemers and Skrzypek (1972), using West Point cadets — that the Contingency Model is population specific and "task-dependent" (1977:207). If he is correct in his observation, the aggregated findings for the military studies ought to approximate Fiedler's curve. By contrast, studies conducted in non-military settings should reveal less external congruency.

When "organizational setting" was coded for the meta-analysis, the variable had a total of 11 possible values. These were subsequently reduced to five on the basis of their similarity (e.g., business and industry were combined as "profit-motive organizations"; elementary and secondary schools as "schools"). The religious setting was dropped entirely because there was only one study in the subset. The two remaining settings were hospitals and universities. The data to be used in the presentation of results are shown in Table 7.6. In terms of directional congruency, the military population had seven signs in the

TABLE 7.6: MEDIAN CORRELATIONS FOR ORGANIZATIONAL SETTINGS

		OCTANT							
Setting[1]		I	II	III	IV	V	VI	VII	VIII
Military		-.490	-.460	-.570	.350	.480	.130	-.460	-.565
Univ/ROTC		-.085	-.385	.030	.205	-.330	-.370	.170	-.395
Schools		-.020	.250	-.045	-.065	.050	.070	-.080	.170
PMO		-.470	-.100	-.035	-.080	.210	-.240	.200	-.330
Hospitals		-.315		.080		.460		.170	
Fiedler[2]		-.52	-.58	-.33	.47	.42	(+)[3]	.05	-.43
Military	n	8	2	9	1	8	1	10	2
Univ/ROTC	n	12	6	13	10	9	6	5	8
Schools	n	11	5	14	18	8	3	12	17
PMO	n	7	1	6	3	5	1	6	2
Hospitals	n	14	0	2	0	2	0	2	0
Fiedler	n	8	3	12	10	6	0	12	12

1. Abbreviations: Univ/ROTC = University including ROTC cadets, Schools = Elementary and secondary schools PMO = Profit-motive organizations

2. The medians correlations predicted by Fiedler's Contingency Model.

3. Fiedler (1967) predicted only direction for Octant VI.

predicted direction (only Octant VII is opposite) followed by the profit-motive organizations which produced directional support in six (all octants except IV and VI). However, it must be noted that the number of reported correlations upon which the medians are based is less than three in four of the military octants, three of the profit-motive organization octants, and three of the hospital octants. The presence of these small cell sizes considerably reduces the confidence which might otherwise be

placed in directional support for the Model as a whole. Of equal interest is the number of correlations in the non-support cells. In Octant VII for the military studies there are ten reported correlations, a number within the same range as that found in cells yielding directional support.

Of the remaining three settings, only two — university/ROTC and schools — had tests in all octants. Although the two settings yielded directional support in four octants, only Octant I was supported by both of them. The relationship between cell size and support or non-support shows that smaller cell sizes are associated with support while larger ones are associated with non-support. For example, in schools, n's of 18, 12, and 17 (Octants IV, VII, and VIII, respectively) are in the non-support cells while n's of 11, 8, and 3 (Octants I, V, and VI, respectively) are in the support cells.

While the previous observations have been based on an across-octant, within-organization sign analysis, it is of some interest to inspect the results within octants, and across organizational settings. The procedure places the emphasis on octant support rather than support derived from any particular organization. The column inspection shows that only Octant I is totally supported by all median correlations. However, when the predicted magnitude of -0.52 is considered, this support for Octant I decreases

markedly with the deletion of two values (-0.085 and -0.020).

In order to evaluate the amount of internal and external congruency, the medians have been displayed graphically in Figure 7.6.

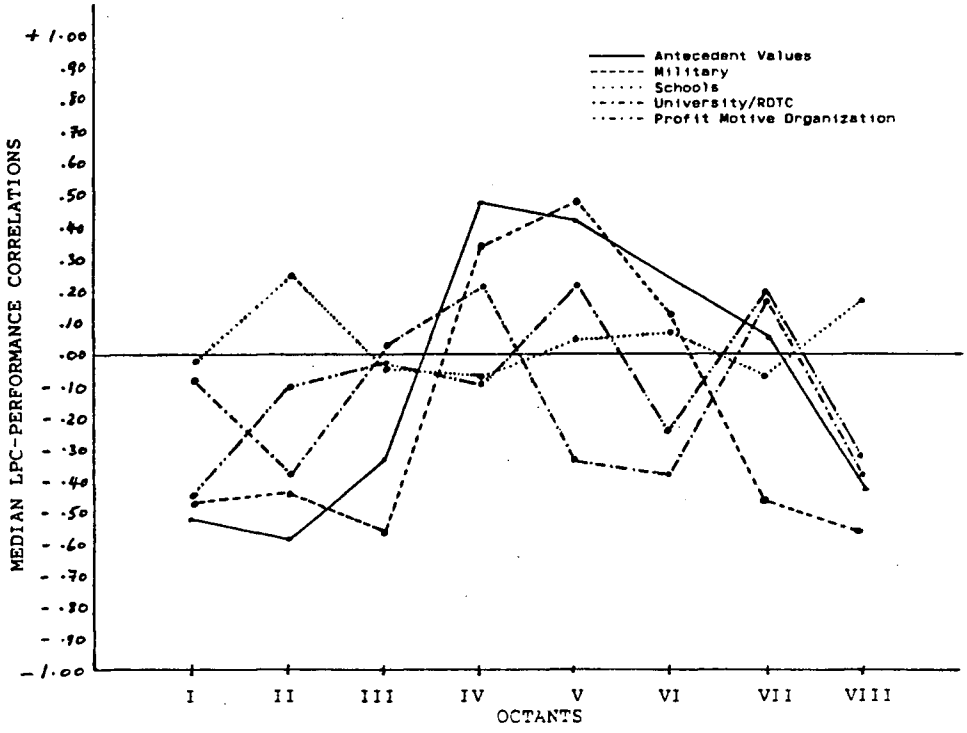


FIGURE 7.6: INTERNAL AND EXTERNAL CONGRUENCY:
Organizational Setting

Because the hospitals had only four median correlations, their results were not plotted. It is readily apparent that military and school settings show the most and least congruency, respectively, with Fiedler's curve. In the case of school-based studies, it would appear — with the possible exception of Octants II and VIII — that LPC score of educational leaders has little to do with their

effectiveness. Moreover, and contrary to the predictions of the Model, it seems that high LPC leaders (positive correlation) are the more effective when the degree of control and influence is at or near the extremes of the situational favourableness continuum. It is difficult to avoid the conclusion that the Contingency Model is not appropriate for predicting leadership effectiveness in school settings.

By contrast, the correspondence yielded by the military studies not only suggests that the Model may be useful in this setting but also lends considerable support to the population specificity dimension of Vecchio's criticism. These data do not speak to his observation of task-dependency.

Academic Discipline

This fourth covariate, suggested by the aggregated findings themselves, was intended to serve two purposes. The first purpose, as in the previous three analyses, was to ascertain the level of internal congruency among the disciplines and their external congruency with the predicted curve. The second purpose was to determine the extent of agreement between the results yielded by the organizational settings and those generated by authors from different academic disciplines. Although not exclusively so, the Model was typically tested by psychologists in military

settings (n=5), by educators in schools (n=9), and by commerce people in business and industrial settings (n=10). The aggregated results for each academic discipline are shown in Table 7.7. Two explanatory notes are required

TABLE 7.7: MEDIAN CORRELATIONS FOR ACADEMIC DISCIPLINES

Academic		OCTANT						
Discipline[1]		I	II	III	IV	V	VI	VII	VIII
Psychology		-.430	-.285	-.190	.130	.280	.130	-.200	-.365
Ed Admin		.055	.150	-.050	-.050	-.165	-.180	-.090	.170
Ed/Ed Ad		-.050	.150	.035	-.050	-.145	-.180	-.045	.170
Commerce		-.315	-.525	-.210	.260	.310	-.070	.185	-.375
Fiedler[2]		-.52	-.58	-.33	.47	.42	(+)[3]	.05	-.43
Psychology	n	11	14	17	6	11	5	14	6
Ed Admin	n	13	6	13	19	10	4	9	17
Ed/Ed Ad	n	29	6	18	19	14	4	13	17
Commerce	n	12	4	9	7	7	2	8	6
Fiedler	n	8	3	12	10	6	0	12	12

1. Abbreviations: Ed Ad(min) = Educational Administration
   Ed/Ed Ad = Education and Educational Administration
2. The median correlations predicted by Fiedler's Contingency Model.
3. Fiedler (1967) predicted only direction for Octant VI.

before discussing the results. First, the "education" studies are subdivided into (1) educational administration and (2) any area other than educational administration (e.g., higher education and physical education). Because

the second group of education studies yielded results in only four octants (I, III, V, and VII), they have been combined with educational administration for purposes of this analysis. Second, "psychology" includes one study from sociology.

When the aggregated results are analyzed in terms of directional congruency, some support for the Model as a whole is yielded by psychology and commerce. Each showed directional agreement in seven out of eight octants. The outcomes of both the education and educational administration are in sharp contrast not only with those from psychology and commerce but also with the predictions of the Model. The sign analysis of educational administration shows only Octant III to be in the predicted direction. When the educational administration and education studies are considered together, there is still only one octant directionally congruent but it is now Octant I. When magnitude and direction are used jointly to assess support, the psychology and commerce medians are fairly close to the antecedent values in all octants except VI and VII. The internal and external congruencies yielded by psychology, commerce, and the total sample of education studies are illustrated in Figure 7.7.

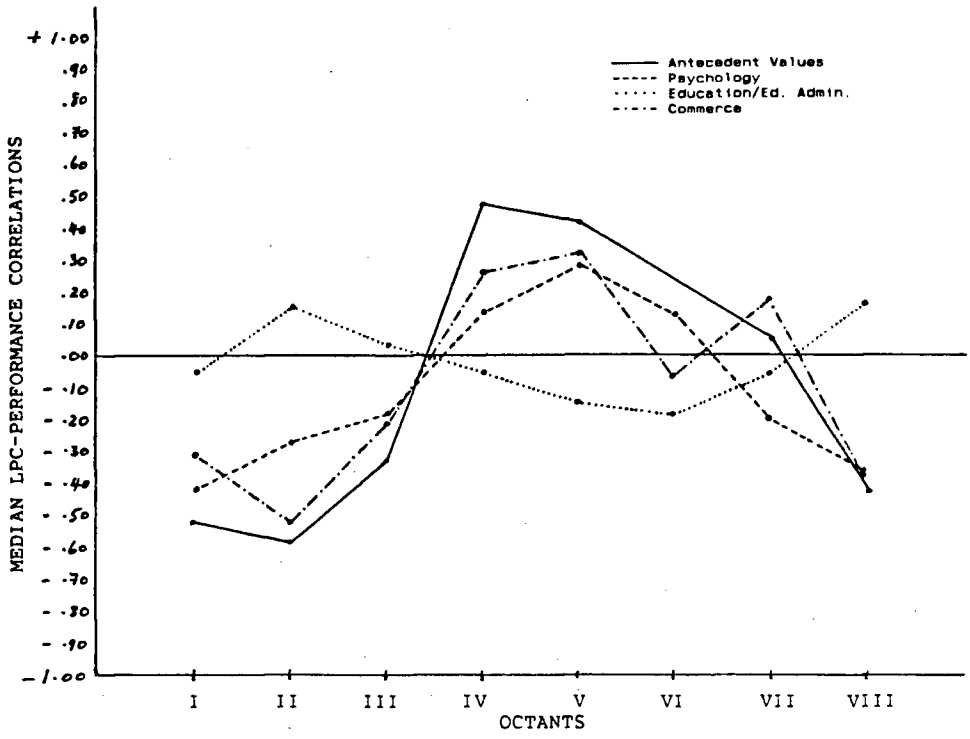


FIGURE 7.7: INTERNAL AND EXTERNAL CONGRUENCY:
Academic Disciplines

The most striking feature in Figure 7.7 is the incongruency of the education studies relative to the other two disciplines and to the antecedent curve. In Octants I, II, III, V, and VIII, the values from psychology, commerce, and Fiedler almost converge on the same points. By contrast, the shape of the curve generated from the entire sample of education studies reveals an almost total lack of both internal and external congruency.

When the octant tests are partitioned by academic discipline, the outcomes which emerge correspond quite closely with those generated by the organizational settings.

The comparisons between disciplines and organizations relative to each other and to Fiedler's curve are shown in a series of graphic representations. Figure 7.8 compares psychology with the military; Figure 7.9 compares commerce with profit-motive organizations (i.e., business and industry); Figure 7.10 compares educational administration with schools.

There are two points to be made regarding the displays in these three graphics. The first point concerns the congruence between Fiedler's Model and certain subsets of the data. As shown in Figure 7.8, the Model has predictive validity in all octants except VII, when the testing is done by psychologists and the subjects are active duty personnel or military cadets. The same observation applies, although not as strongly, to the visual display in Figure 7.9 which shows the results when researchers from commerce backgrounds test the Model using subjects in business (e.g., insurance company executives) or in industry (e.g., shop foremen).

The second point relates to Figure 7.10. It was noted, in connection with the organizational setting covariate, that leader LPC score seemed to have little to do with the effectiveness of educational leaders. Another way of interpreting the uniformly low median correlations is to suggest that any one or more of the three situational variables (leader-member relations, task structure, and

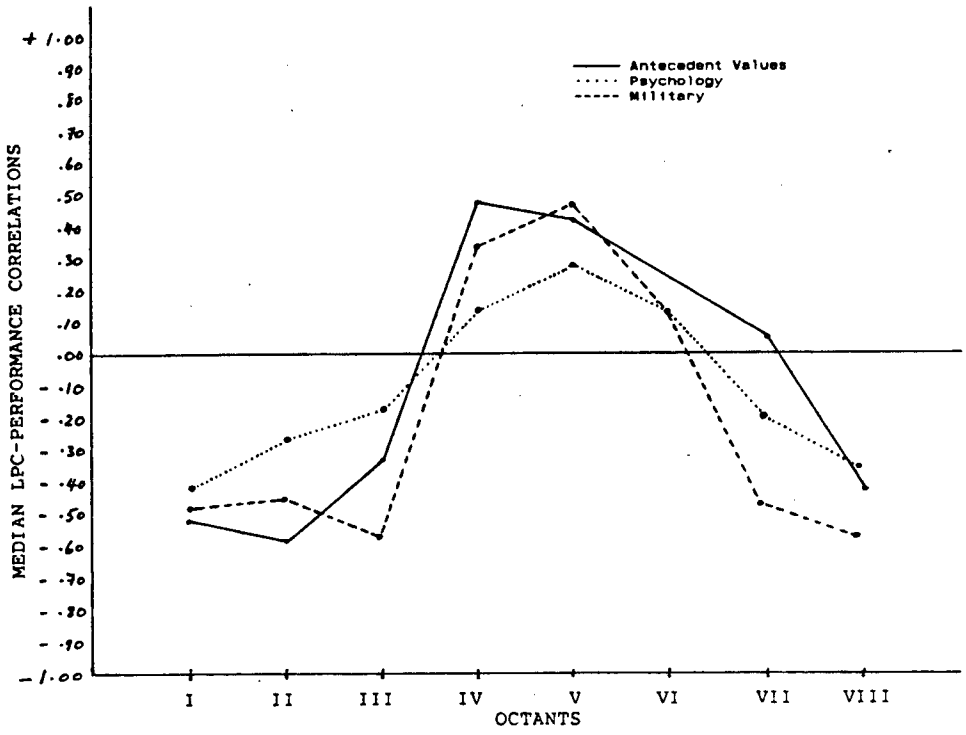


FIGURE 7.8: INTERNAL CONGRUENCY: Psychology and Military

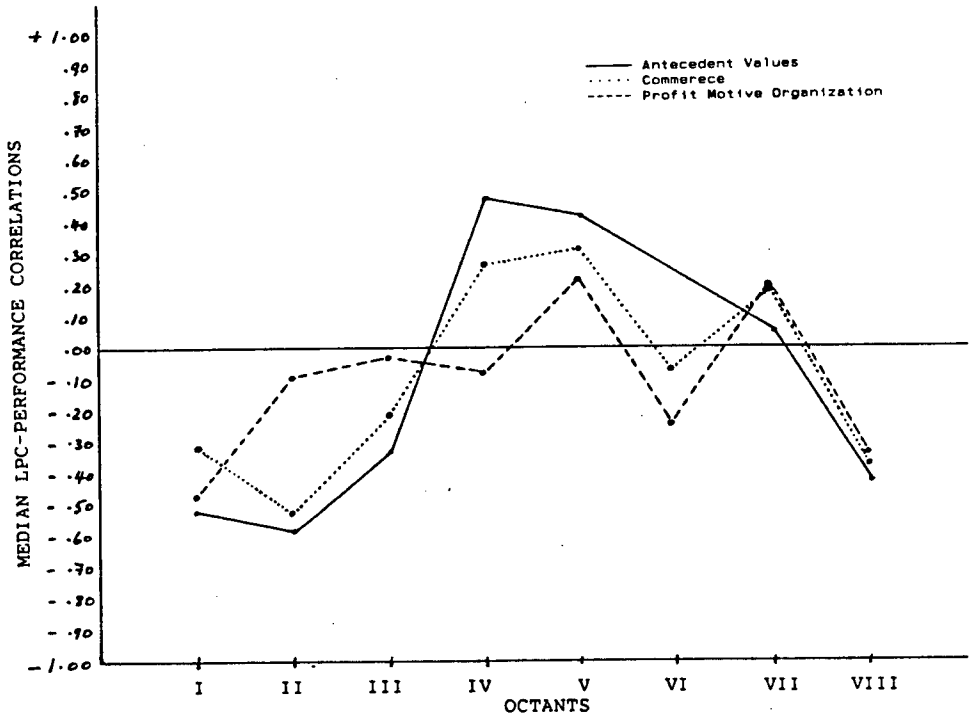


FIGURE 7.9: INTERNAL CONGRUENCY: Commerce and Profit Motive Organizations

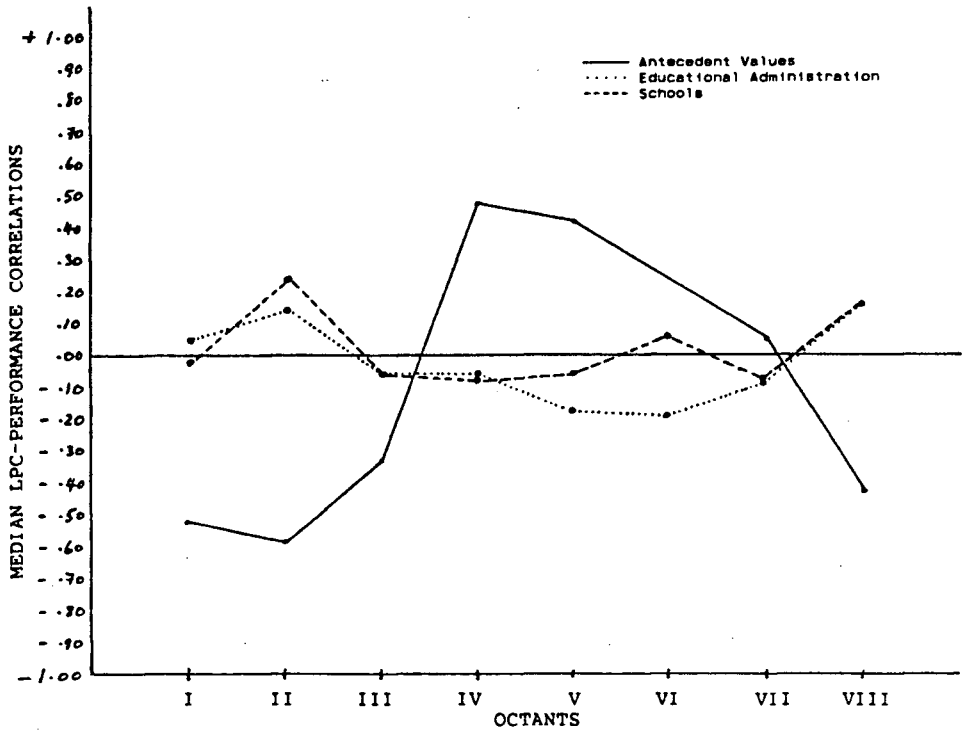


FIGURE 7.10: INTERNAL CONGRUENCY: Educational Administration and Schools

position power) may have less relevance for education than for military and business organizations. It may be that in educational settings there are some unidentified contextual variables which function either to moderate or nullify the relationship between situational favourableness and leadership style.

Fiedler Associate

This fifth covariate was chosen in an attempt to corroborate a recent observation made by Strube and Garcia. They note that:

> The obtained test, $t(143)=1.43$, $p<.08$, indicates that studies conducted by researchers affiliated with Fiedler were somewhat, though not significantly, more supportive of the Model than those conducted by independent researchers (1981:317).

This variable was coded in terms of three possible values: no, yes, and indirectly. The first requires no explanation. The value "yes" referred to the following associations with Fiedler: a co-author, a graduate student, or a research associate. The value "indirectly" included those who were associated through an associate. While such a category may seem "hair-splitting" to some, it was included to identify the growth in Fiedler's "family tree" (i.e., the research network). Thus, in keeping with the analogy, the "grandchildren" (e.g., Rice who was a student of Chemers who was a student of Fiedler himself), and the "great grandchildren" (e.g., Seaman who was a student of Rice) were "indirect" associates. The category also included two or three dissertations in which Fiedler was acknowledged for "helpful suggestions" with data analysis. Because of "missing data" for Octants IV, VI, and VIII, the "indirect" and "yes" group were combined for purposes of analysis. The resulting data are shown in Table 7.8.

There is a substantial difference between the values in the two extreme categories. In terms of magnitude or direction or both, the "no" category shows little congruency with the "yes" group. With the exception of Octant VII, the studies conducted by Fiedler's associates show considerably

TABLE 7.8: MEDIAN CORRELATIONS FOR FIEDLER ASSOCIATES[1]

		OCTANT							
Value		I	II	III	IV	V	VI	VII	VIII
Yes		-.500	-.320	-.570	.200	.490	.300	-.360	-.615
Yes+Ind		-.430	-.460	-.110	.200	.385	.300	-.205	-.615
No		-.050	.120	-.047	-.005	.045	-.235	.180	.120
Fiedler[2]		-.52	-.58	-.33	.47	.42	(+)[3]	.05	-.43
Yes	n	11	1	11	2	11	3	12	4
Yes+Ind	n	17	2	19	2	16	3	18	4
No	n	35	12	25	30	16	8	17	25
Fiedler	n	8	3	12	10	6	0	12	12

1. Abbreviations: Yes+Ind = Yes + Indirectly

2. The median correlations predicted by Fiedler's Contingency Model.

3. Fiedler (1967) predicted only direction for Octant VI.

greater correspondence with the predicted values than do those conducted by independent researchers. When both groups of associates are considered together (yes + indirectly), the obtained values in Octants I, III, V, and VII are consistently lower than those for the direct associates but higher than those for non-associates. Collectively, these data suggest that studies conducted by non-associates are less supportive of the Model than are those conducted by associates. The internal and external congruency is represented graphically in Figure 7.11.

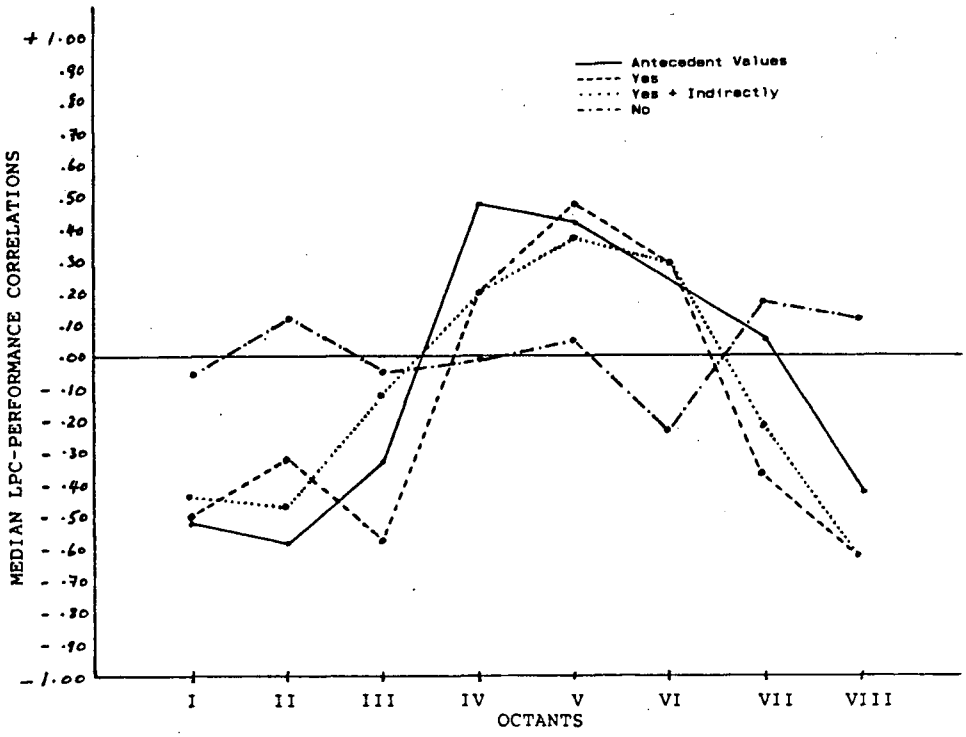


FIGURE 7.11: INTERNAL AND EXTERNAL CONGRUENCY: Fiedler Associates

The curves yielded by either group of associates are quite different in shape from the curve yielded by non-associates. Moreover, the associates' curves resemble Fiedler's predicted curve more closely than does the curve produced by non-associates. The lack of internal congruency between the associate and non-associate curves together with the presence of greater external congruency for the associates' curves lend support to the difference found by Strube and Garcia (1981).

Source of Document

The sixth "covariate" was prompted by the meta-analysis literature which points rather strongly to the existence of publication bias. More specifically, Glass et al. note that:

> The bias in the journal literature relative to the bias in the dissertation literature is not inconsiderable .... findings reported in journals are .... more disposed toward the favored hypotheses of the investigators than findings reported in theses or dissertations (1981:67).

The meta-analysis reported here included research reports from three sources which were coded as journal, dissertation or thesis, and technical report. Because the number of technical reports was very small (n=2), they were not included in this particular analysis. The exclusion of this source of data was intended not as an attempt to increase or decrease the bias effect but rather to retain the purity of the other two groups. If the evidence reported by Glass et al. (1981) is applicable to the present study, then the journal reports should yield the greater correspondence with Fiedler's curve. The within-octant median correlations for journals and dissertations are shown in Table 7.9.

When the results were analyzed in terms of directional congruency, there appeared to be little difference between the two sources. Four octants were supported by the published research; five were supported by

TABLE 7.9: MEDIAN CORRELATIONS FOR SOURCE OF DOCUMENT

		OCTANT							
Source		I	II	III	IV	V	VI	VII	VIII
Journal		-.320	-.320	.020	-.330	.280	-.240	.430	.125
Dissertation		-.115	.140	.027	.140	.050	.070	-.045	-.090
Fiedler[1]		-.52	-.58	-.33	.47	.42	(+)[2]	.05	-.43
Journal	n	5	5	9	15	5	5	7	16
Dissertation	n	42	9	29	17	21	5	19	12
Fiedler	n	8	3	12	10	6	0	12	12

1. The median correlations predicted by Fiedler's Contingency Model.

2. Fiedler (1967) predicted only direction for Octant VI.

the unpublished research. Of the four unsupported octants (III, IV, VI, and VIII) from the journal articles and the three unsupported octants (II, III, and VII) from the dissertations, only Octant III is common to both sources. Moreover, only two octants (I and V) are supported by both journals and dissertations.

When magnitude and direction were treated jointly to assess octant support, a slightly different pattern emerged. A comparison of the obtained and antecedent values for the congruent octants from both sources indicated that there was a greater degree of support generated by the journals than by the dissertations. As indicated in Table 7.9, the magnitude of the journal medians for Octants I and V were closer to the strength of the predicted correlations than

were those from the same octants in the dissertation research. The internal and external congruency is shown in Figure 7.12.

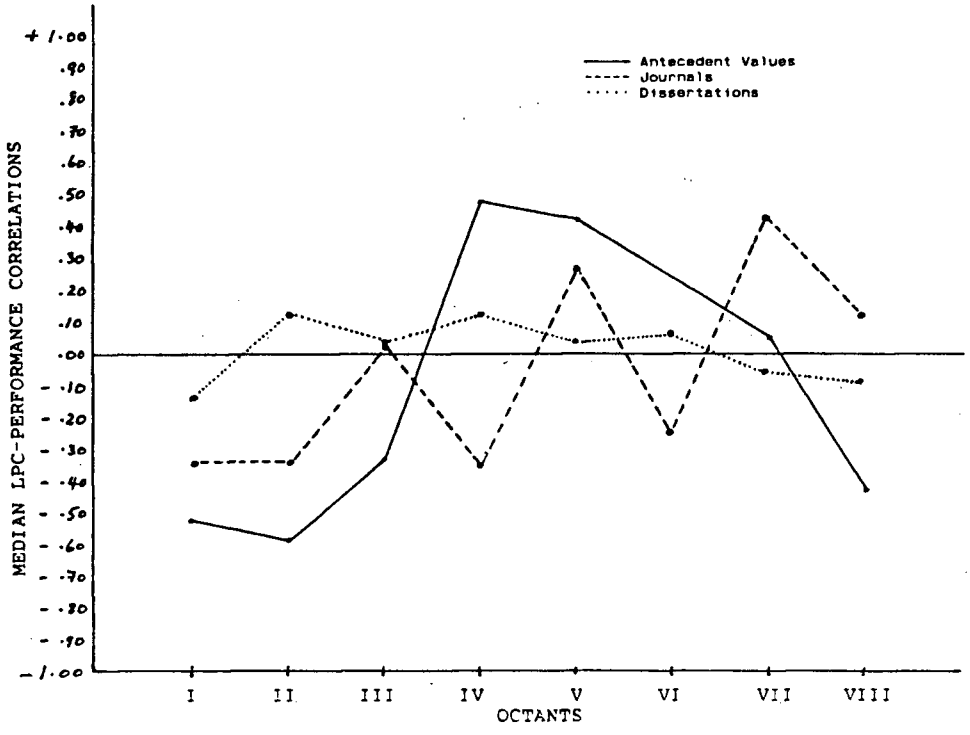


FIGURE 7.12: INTERNAL AND EXTERNAL CONGRUENCY: Source of Document

It is clear from the visual representation that neither source yielded more than partial support for the Model as a whole. However, consistent with the observation by Glass et al. (1981), the journal results did reveal a greater degree of external congruency than did the dissertation findings.

Study Quality

The selection of this seventh covariate was also prompted by the meta-analysis literature. Glass et al. (1981) report that quality has influenced the outcomes of some studies but not of others. The two measures of quality used in the present study permitted an examination of whether or not study quality made any difference to the support found for Fiedler's Model. The reported correlations were grouped within octants according to a mean split on the TV (threats to validity) score ($\bar{x}$=3.343, n=249) and on the MA (methodological adherence) score ($\bar{x}$=3.173, n=249) assigned to each study (see Chapter 4, pp. 90 and 91). The resulting median correlations are shown in Table 7.10.

TV scores. The median correlations derived from the more valid studies (higher TV scores) are directionally congruent in seven of the eight-octants. The less valid studies (lower TV scores) are directionally congruent in only three octants (I, V, and VII). The strength of the "high TV" medians is generally greater than that of the "low TV" medians. The median correlations yielded by studies judged to be more valid are closer in magnitude to the values predicted by Fiedler than are the medians yielded by

TABLE 7.10: MEDIAN CORRELATIONS FOR STUDY QUALITY BASED ON HIGH AND LOW TV AND MA SCORES USING MEAN SPLITS[1]

Study Quality Score	OCTANT I	II	III	IV	V	VI	VII	VIII
High TV	-.320	-.295	-.180	.245	.315	.115	-.150	-.100
Low TV	-.040	.150	.027	-.325	.045	-.200	.015	.150
High MA	-.317	-.320	-.097	.140	.210	.215	.075	-.100
Low MA	-.120	.140	.000	-.077	.250	-.200	-.042	.115
Fiedler[2]	-.52	-.58	-.33	.47	.42	(+)[3]	.05	-.43
High TV n	25	6	23	14	20	6	21	10
Low TV n	27	8	21	18	12	5	14	19
High MA n	27	3	15	9	11	4	12	7
Low MA n	25	11	29	23	21	7	23	22
Fiedler n	8	3	12	10	6	0	12	12

1. Mean scores: TV = 3.343 (n=249); MA = 3.173 (n=249).

2. The median correlations predicted by Fiedler's Contingency Model.

3. Fiedler (1967) predicted only direction for Octant VI.

studies judged to be less valid. The extent to which the more and less valid studies are internally and externally congruent is shown in Figure 7.13.

With the possible exception of Octants I and V, there is a general lack of internal congruency between the two curves based on the obtained medians. In terms of external congruency, the "high TV" curve corresponds much more closely with Fiedler's predicted curve than does the "low TV" curve. The finding of greater external congruency for the more valid studies lends support to the argument that

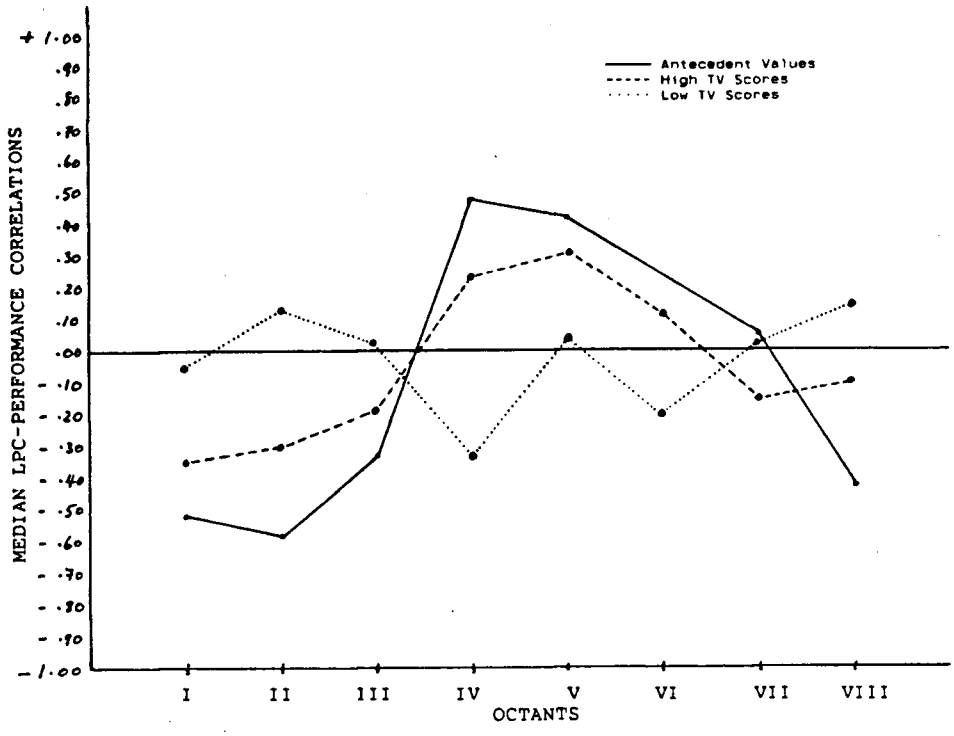


FIGURE 7.13: INTERNAL AND EXTERNAL CONGRUENCY:
High and Low TV Scores

meta-analytic outcomes do vary according to the quality of the studies comprising the data base.

MA scores. The median correlations derived from studies which adhered more closely to Fiedler's method (higher MA scores) are directionally congruent in all eight octants. By contrast, studies which adhered less closely (lower MA scores) are directionally congruent only in Octants I and V. The magnitude of the "high MA" medians generally exceeds that of the "low MA" medians. The extent of correspondence between the two sets of obtained medians and between each set of medians and Fiedler's predicted values is shown in Figure 7.14.

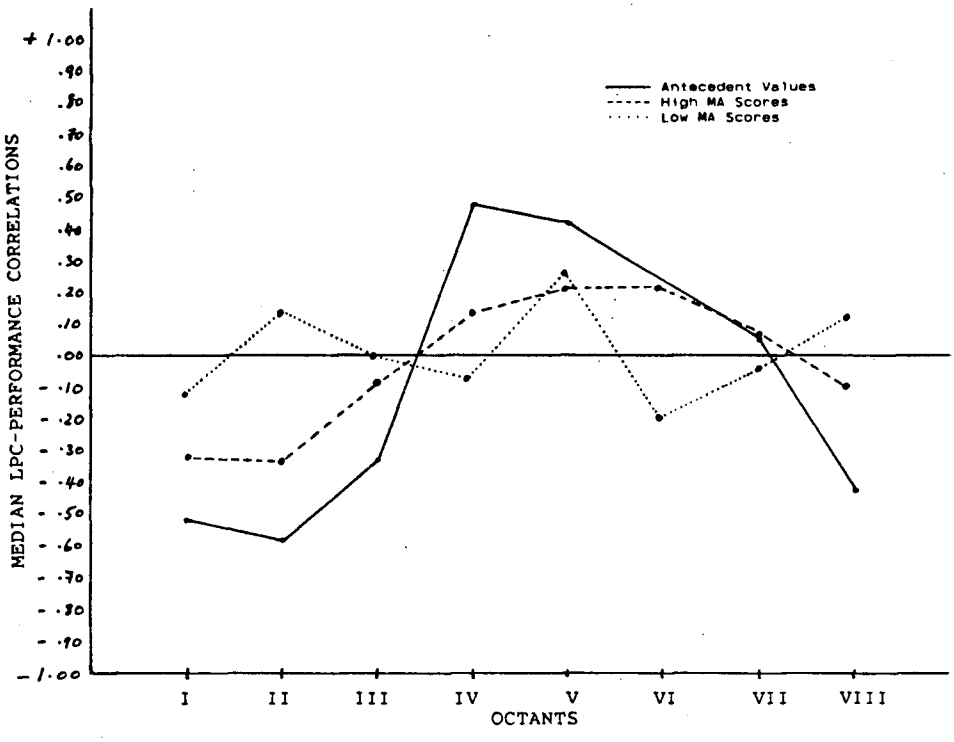


FIGURE 7.14: INTERNAL AND EXTERNAL CONGRUENCY:
High and Low MA Scores

The two curves based on the obtained medians are, with the exception of Octants I and V, clearly lacking in internal congruency. The extent of external congruency for the "high MA" curve and the lack of external congruency for the "low MA" curve indicate that closer adherence to Fiedler's method generates results which are more supportive of the Contingency Model. This finding is consistent with the criticism that the Model is predictive only when the prescribed method is followed (e.g., Ashour, 1973a; Barrow, 1977; Schriesheim and Kerr, 1977).

SUMMARY OF RESULTS

Two analyses were conducted using the median correlations derived from the "rho + r" listing. The first part of the summary reports the results for the analysis of the two stratification variables. The second part of the summary reports the findings derived from the examination of the seven selected study characteristics.

Results for the Stratification Variables

1. Comparability Categories (Table 7.2 and Figure 7.2)

    1.1 the median correlations from studies classified as CC:Fiedler corresponded more closely with Fiedler's predicted medians both in direction and magnitude than did those from studies classified as CC:Other;

    1.2 the median correlations from studies classified as CC:Other showed directional correspondence in only four octants; the magnitude of all obtained values was lower than that predicted by Fiedler.

2. Task Group Types (Table 7.3 and Figure 7.3)

    2.1 the median correlations from inferred interacting task groups (TGT 4) corresponded more closely with Fiedler's predicted medians than did those from groups stated to be interacting (TGT 1);

    2.2 the median correlations from nonacting task groups (TGT 8) corresponded more closely with Fiedler's predicted medians than did those from either TGT 1 or TGT 4;

2.3 the median correlations from all task group types together corresponded in both direction and magnitude with Fiedler's predicted medians in only two Octants.

Results for the Study Characteristics Variables

1. Basis for Assessing Leadership Effectiveness (Table 7.4 and Figure 7.4)

   1.1 leader performance and group performance yielded quite disparate within-octant median correlations;

   1.2 leader performance median correlations corresponded more closely with Fiedler's predicted medians than did group performance median correlations.

2. Study Setting (Table 7.5 and Figure 7.5)

   2.1 real-life and laboratory studies yielded median correlations which were unlike each other in both direction and magnitude;

   2.2 neither real life nor laboratory settings yielded conclusive support for all octants; only Octant I was supported by both settings;

   2.3 real-life median correlations corresponded more closely with Fiedler's predicted medians than did those from laboratory settings.

3. Organizational Setting (Table 7.6 and Figure 7.6)

   3.1 military median correlations corresponded most closely with Fiedler's predicted medians;

   3.2 school-based median correlations corresponded least closely with Fiedler's predicted medians;

3.3 the median correlations from all organizations corresponded with Fiedler's predicted medians only in Octant I.

4. Academic Discipline (Table 7.7 and Figure 7.7)

4.1 psychology and commerce yielded median correlations which corresponded quite closely with Fiedler's predicted medians;

4.2 education generated median correlations which showed very little correspondence with those from either psychology or commerce;

4.3 education yielded median correlations which showed virtually no correspondence with Fiedler's predicted medians;

4.4 military studies conducted by psychologists yielded median correlations which corresponded very closely with those predicted by Fiedler (see Figure 7.8);

4.5 studies conducted in business organizations by researchers from commerce yielded median correlations which corresponded closely with those predicted by Fiedler (see Table 7.9);

4.6 school studies conducted by educational administrators yielded median correlations which showed an almost total lack of correspondence with Fiedler's predicted medians (see Figure 7.10).

5. Fiedler Associate (Table 7.8 and Figure 7.11)

5.1 studies conducted by Fiedler and/or his associates produced median correlations which showed considerable correspondence with the predicted values;

5.2 regardless of author association with Fiedler, the median correlations in Octants I, III, and V were directionally congruent with the Fiedler's predicted medians;

5.3 studies conducted by non-associates yielded median correlations which lacked correspondence both with Fiedler's predicted medians and with those generated by associates.

6. Source of Document (Table 7.9 and Figure 7.12)

6.1 the median correlations yielded by journal reports (published research) and dissertations (unpublished research) were directionally congruent with each other only in Octants I, III, and V;

6.2 the median correlations yielded by journal reports (published research) corresponded more closely with Fiedler's predicted medians than did those from dissertations (unpublished research).

7. Study Quality (Table 7.10 and Figures 7.13 and 7.14).

7.1 the median correlations yielded by both high TV and high MA scores corresponded more closely with Fiedler's predicted medians than did those yielded by low TV and low MA scores;

7.2 the median correlations yielded by high and low TV scores and by high and low MA scores were directionally congruent with each other and with Fiedler's predictions only in Octants I and V.

To provide a complete picture of the obtained within-octant results, the median correlations yielded by the stratification variables and the seven study characteristics are summarized in Table 7.11.

To avoid possible misinterpretation of these data, three points need to be made about this summary table. First, the within-octant medians are derived from the same

TABLE 7.11: SUMMARY OF MEDIAN CORRELATIONS WHEN OCTANT TESTS ARE GROUPED BY STRATIFICATION AND STUDY CHARACTERISTICS VARIABLES

		OCTANT							
		I	II	III	IV	V	VI	VII	VIII
VARIABLE	Predicted Medians[1] ⟶	-.52	-.58	-.33	.47	.42	(+)[2]	.05	-.43
Comparability	Fiedler	-.335	-.275	-.050	.240	.280	-.370	.300	-.097
Category	Other	-.115	.120	-.050	-.080	.220	.100	-.085	.115
	Stated Interacting	-.040	-.010	-.175	-.065	-.345	.070	-.140	.165
Task Group	Inferred Interacting	-.120	-.275	.145	.125	.265	.115	.090	-.332
Type	Coacting	-.320		-.040		.505		.260	
	Nonacting	-.310	-.180	-.555	.165	.050	-.425	-.365	-.285
Leadership	Group Performance	-.132	-.111	-.035	-.010	.245	.115	.080	.095
Effectiveness	Leader Performance	-.172	.140	-.405	.090	.230	-.200	-.145	-.090
Study	Real Life	-.167	-.020	-.097	-.065	.245	-.200	-.140	.155
Setting	Laboratory	-.130	.110	.027	.265	-.140	.115	.085	-.332
	Military	-.490	-.460	-.570	.350	.480	.130	-.460	-.565
Organizational	University/ROTC	-.085	-.385	.030	.205	-.330	-.370	.170	-.395
Setting	Schools	-.020	.250	-.045	-.065	.050	.070	-.080	.170
	Profit-Motive Organizations	-.470	-.100	-.035	-.080	.210	-.240	.200	-.330
	Hospitals	-.315		.080		.460		.170	
Academic	Psychology	-.430	-.285	-.190	.130	.280	.130	-.200	-.365
Discipline	Educational Administration	.055	.150	-.050	-.050	-.165	-.180	-.090	.170
	Commerce	-.315	-.525	-.210	.260	.310	-.070	.185	-.375
Fiedler	Yes	-.500	-.320	-.570	.200	.490	.300	-.360	-.615
Associate	No	-.050	.120	-.047	-.005	.045	-.235	.180	.120
Document	Journal	-.320	-.320	.020	-.330	.280	-.240	.430	.125
Source	Dissertation	-.115	.140	.027	.140	.050	.070	-.045	-.090
	High TV (more valid)	-.320	-.295	-.180	.245	.315	.115	-.150	-.100
Study	Low TV (less valid)	-.040	.150	.027	-.325	.045	-.200	.015	.150
Quality	High MA (more adherence)	-.317	-.320	-.097	.140	.210	.215	.075	-.100
	Low MA (less adherence)	-.120	.140	.000	-.077	.250	-.200	-.042	.115

1. Median correlations predicted by Fiedler's Contingency Model.
2. Fiedler (1967) predicted only direction for Octant VI.

eight sets of reported correlations. For example, the medians in Octant I for "CC:Fiedler" and "CC:Other" are based on the same set of reported correlations as are the medians in Octant I for the other seven variables. Therefore, the median correlations across all variables do not constitute independent data.

The second point concerns the basis for deciding that one octant is supported while another is not. The use of directional congruency, irrespective of magnitude, as the sole criterion by which to assess support leaves open the question of how often the predicted signs are yielded by chance alone. The lack of independence among the medians makes it impossible to provide an answer.

The third and most important point to be made focusses on the misleading impression of support when the magnitude of obtained values is not considered. For instance, the "sign analysis" comparing the bases for determining leadership effectiveness would indicate that group performance yields greater support across octants. Yet, when the medians were graphed (see Figure 7.4, p.164), it was leader performance which showed greater correspondence with Fiedler's curve. Likewise, an inspection of directional congruency for the source of a study suggests that more supportive results emerged from dissertations. Yet, when the medians were graphed, it was the journals which revealed more external congruency (see

Figure 7.12, p.182). These two examples make it clear that support and non-support for any octant of the Model must be assessed on the basis of both magnitude and direction. Only if statistical tests for significant differences between the obtained and antecedent values are used can within-octant support be stated with any confidence. Until this is done, the extent to which any octant is or is not supported must remain in abeyance.

CONCLUSIONS AND DISCUSSION

This chapter has presented and discussed the outcomes yielded by aggregating the reported correlations in terms of median values. The aggregation process has made it clear that both magnitude and direction must be considered together in determining the correspondence between the obtained results and those predicted by Fiedler. In addition to this observation another very important issue — namely partitioned data set combination — emerged from the analysis of the median correlations.

It was noted in Chapter 4 (Figure 4.1, p. 105) that interacting and coacting task groups were to be combined for further analysis only if both groups revealed patterns of LPC-performance correlations which were more similar than dissimilar. Furthermore, if the correlation pattern yielded by nonacting groups ("Leaders") was found to be consistent

with those from the interacting and coacting task groups, all three sets of LPC-performance correlations could be combined. It was also stated that the appearance of more dissimilarity than similarity in any one octant would prevent the combining of the partitioned data sets. The two comparability categories which were subsequently created are subject to the same constraints. The median correlation analysis, particularly the graphic representations, has cast doubt on the legitimacy of combining the partitioned data sets.

The importance of the data set combination issue is related to the credibility of the findings yielded by this meta-analysis. Considerably greater confidence could be placed in the outcomes if the cell sizes were to include all the available LPC-performance correlations in each octant. The need to determine similarities or differences between partitioned sets of data would have been unnecessary had there been a large enough number of octant tests in any one set of data. For instance, had there been a sufficient number of LPC-performance correlations (e.g., 10 or more) in each octant for real-life studies involving only stated interacting task groups within CC:Fiedler for a particular organizational setting, collapsing of the partitioned data sets to increase cell sizes would not have emerged as a problem. That very small n's in fact appeared not only cast considerable doubt on the observation that the Model had

been extensively tested but also made it mandatory that the combining issue be resolved. This issue forms the substance of the next chapter.

Chapter 8

THE PROBLEMATIC NATURE OF COMBINING PARTITIONED DATA SETS

The analysis reported in Chapter 7 raised doubts about the legitimacy of combining the partitioned data sets. This chapter reports an analysis designed to resolve the issue by examining the data in a different way.

The relationship between cell size and combining data sets was noted at the end of the previous chapter. Some extension of that observation is needed to ensure a clear understanding of the problem at hand. A comparison of the two comparability categories (CC:Fiedler and CC:Other) across task group type assumes a similarity among the patterns of the LPC-performance correlations yielded by the four task groups. Likewise, a comparison of any two task group types across comparability categories assumes the similarity of the comparability categories. If the legitimacy of combining the data sets is to be established, the comparisons need to be made both within and across comparability categories and task group types.

These comparisons can be made in one of two different ways. One method involves the use of crosstabulation procedures; the other relies on schematic displays of the data. The second approach was selected for two reasons. First, the usefulness of the potentially more powerful

crosstabulation procedure is limited by the small cell sizes which result from deeper stratification. The effect is shown in Figures 8.1 and 8.2 using Octants II and III as examples.

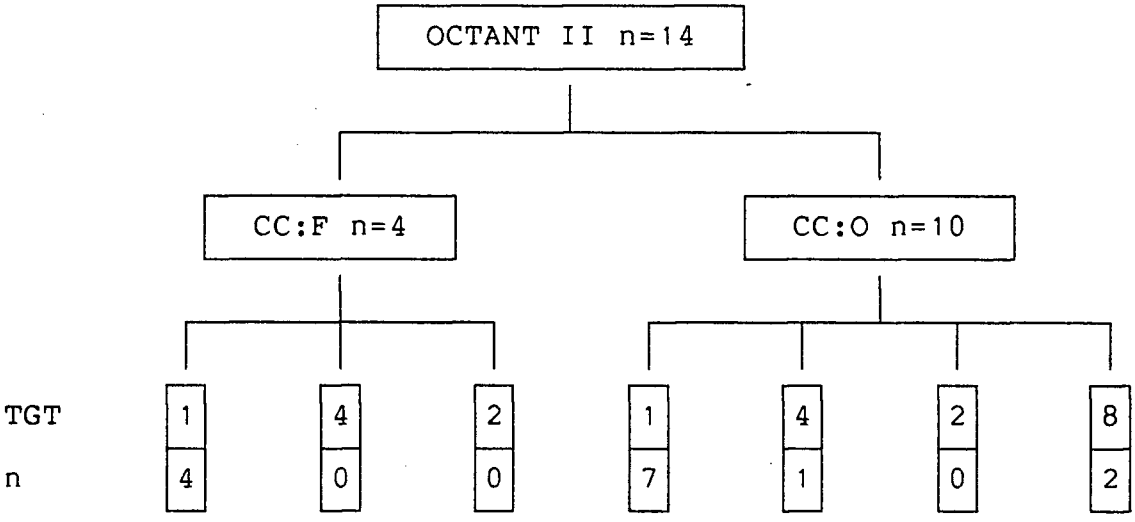


FIGURE 8.1: OCTANT II: CELL SIZE REDUCTION RESULTING FROM STRATIFICATION

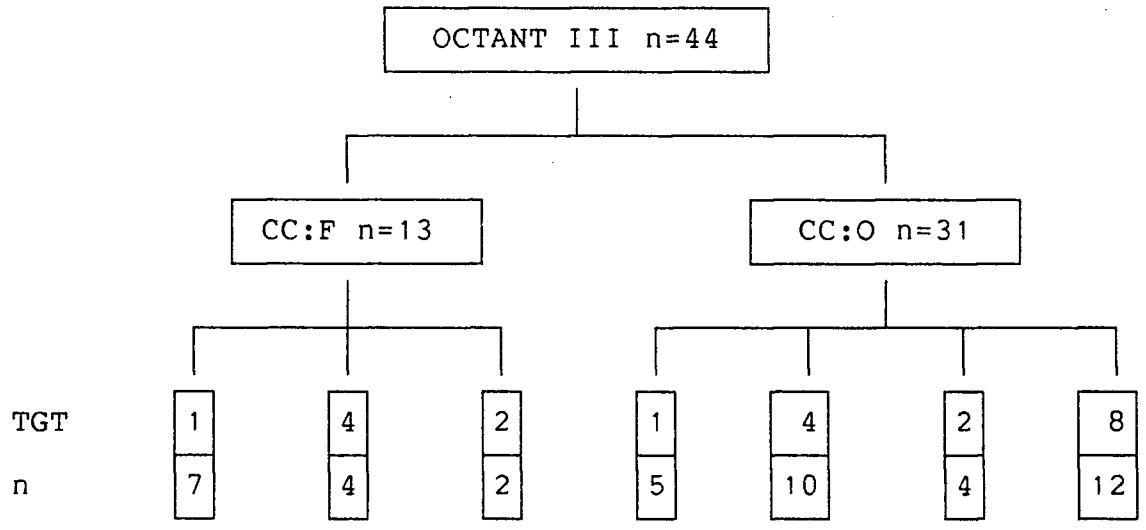


FIGURE 8.2: OCTANT III: CELL SIZE REDUCTION RESULTING FROM STRATIFICATION

Second, the large number of crosstabulations required to ascertain whether there are more similarities than differences in all possible combinations of each comparability category with each task group would create an unmanageable and less than meaningful array of results.

Both these difficulties are minimized with the second method of analysis which relies on schematic displays of data. This method, known as Exploratory Data Analysis, has emerged only fairly recently. In the following section, an explanation of the method precedes a description of its modification for the present study.

EXPLORATORY DATA ANALYSIS: THE METHOD

A detailed explanation of Tukey's Exploratory Data Analysis is beyond the scope of this chapter. However, a brief overview is necessary to provide some understanding of the method. Tukey considers Exploratory Data Analysis as "detective work" which may be numerical or graphical in nature (1977:1). The strength of the evidence provided by the detective work is then evaluated by confirmatory data analysis. In Tukey's opinion, the "techniques for handling and looking" at batches of numbers should be kept as simple as possible (1977:3). The portrayal of the data by visual display can be used to show what Tukey refers to as "separation, asymmetry, irregularities, centering, and width" (1977:21). In other words, the graphics may reveal trends in the normality and shape of the distribution, and in central tendency, dispersion, variability, and range.

To portray the extent to which a set of data exhibits these characteristics, Tukey (1977) suggests two techniques. The first, a box-and-whisker schematic plot, permits some observations about the central tendency and variability of the distribution. By permitting a comparison of two or more distributions, the box-and-whisker graphic indicates the separation between them. The further apart the batches of data, the more different they are; the more the

distributions overlap, the more similar they are (Tukey, 1977). The second, a stem-and-leaf display offers some clues about the shape and width of a distribution of values (Tukey, 1977). For reasons which will become clear as the chapter progresses, only the box-and-whisker technique was used to explore further the feasibility of collapsing of the partitioned data sets in the present study.

Tukey's Box-and-Whisker Technique

In describing the construction of a box-and-whisker schematic plot, Tukey (1977) uses a unique vocabulary. This exposition of the data portrayal technique will rely, where possible, on more familar terms but indicate the Tukey equivalent parenthetically. Figure 8.3 provides a graphic illustration to be read in conjunction with the description which follows. The box portion of the "box-and-whisker" is a rectangle whose upper and lower ends represent, respectively, the 75th and 25th percentiles of a score distribution (the hinges); the median is shown by a bar drawn across the box. Thus, the box contains all the scores within the interquartile range. Scores which lie one-and-one-half times the interquartile range above and below the box are referred to as "inner fences". The scores which lie closest to but inside the inner fences are called "adjacent values". These adjacent values, marked by a dashed crossbar, are joined to the top and bottom of the box

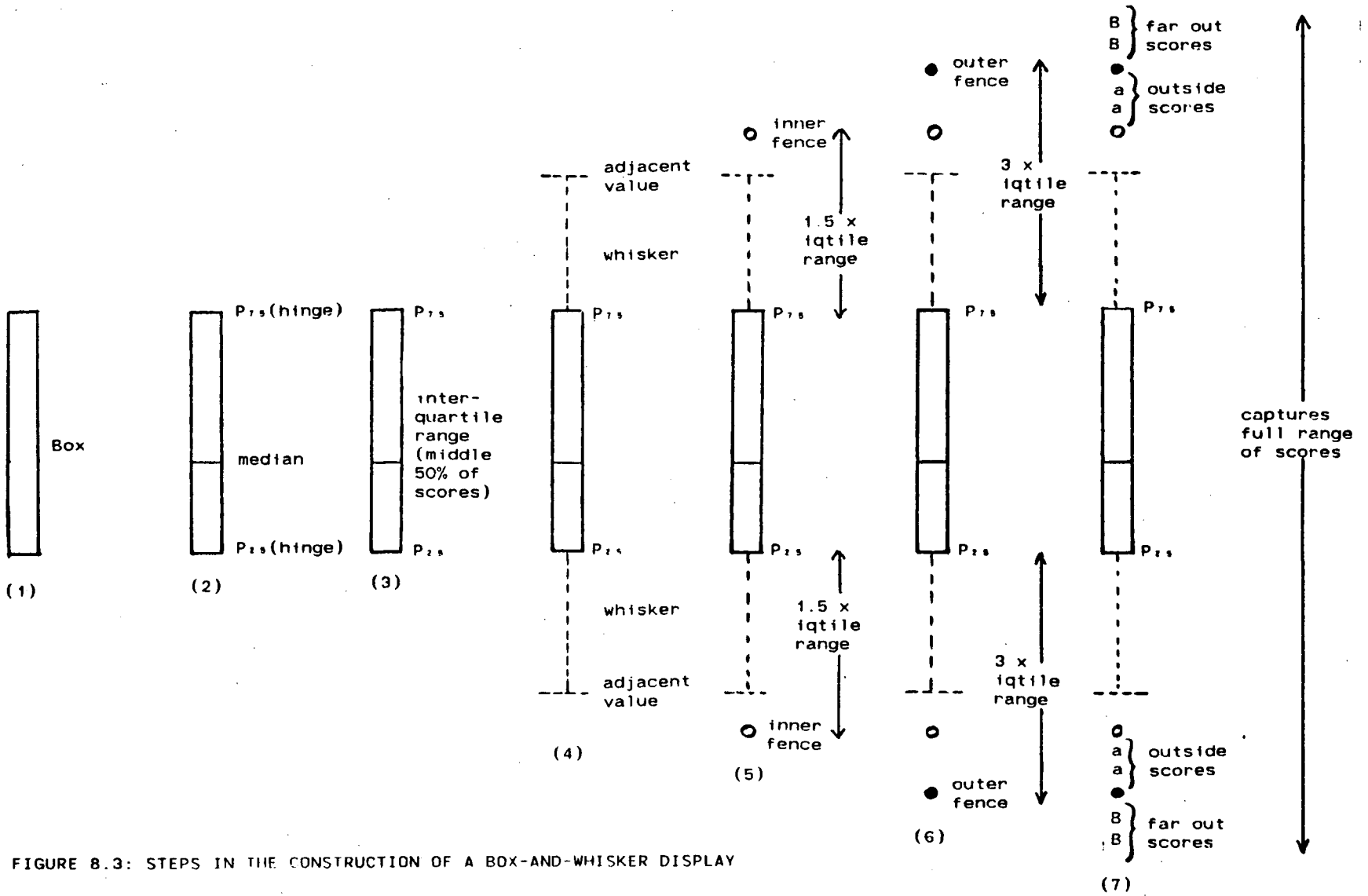


FIGURE 8.3: STEPS IN THE CONSTRUCTION OF A BOX-AND-WHISKER DISPLAY

by a broken line called a "whisker" (Tukey, 1977:44-55).

The schematic also allows for the plotting of two sets of extreme scores (outliers) which fall beyond the inner fences. The scores which lie three times the interquartile range above and below the box are referred to as "outer fences". Any score which falls between the outer and inner fences is referred to as an "outside" score (less extreme outlier); those scores beyond the outer fence are designated as "far out" (more extreme outliers). Collectively, these symbolic designations provide a visual display which permits a comparison of central tendency and variability for two sets of scores.

Modifications to Tukey's Techniques

Several modifications to Tukey's procedures for constructing a box-and-whisker display were needed to accomodate the somewhat unusual nature of the present data[1]. For instance, the limited number of correlations in some octants meant that the extreme value was also both the adjacent value and the inner fence. In yet other cases, there were no correlations lying between the adjacent values

---

[1] The data for the box-and-whisker displays are the within-octant LPC-performance correlations reported by the primary authors rather than the derived medians values used in Chapter 7; the term "scores" thus becomes "correlations".

and the inner fence. The modifications are shown in Figure 8.4.

To be as consistent as possible with Tukey's technique, the whiskers run from the box to the adjacent values but with a solid line to indicate that correlations existed between the hinges (P_{25} and P_{75}) and the adjacent values. The adjacent value itself is marked by a solid crossbar at the end of each whisker (Figure 8.4: Modification A). A broken line extending away from the adjacent value shows that no correlations existed between that value and the inner fence (Figure 8.4: Modification B). The inner fences are designated by either a shaded or an unshaded circle. The unshaded circle denotes the theoretical value of 1.5 times the interquartile range while the shaded circle indicates a "real" value — that is, there was an actual correlation equal to the theoretical value. An unshaded circle joined directly to the box by a broken line indicates that all the correlations were captured within the interquartile range (Figure 8.4: Modification B). The two levels of extreme scores are designated by a small "x" irrespective of distance from the inner fence. Because it was not necessary to distinguish between the two types of outliers, Tukey's outer fences have been omitted (Figure 8.4: Modification C).

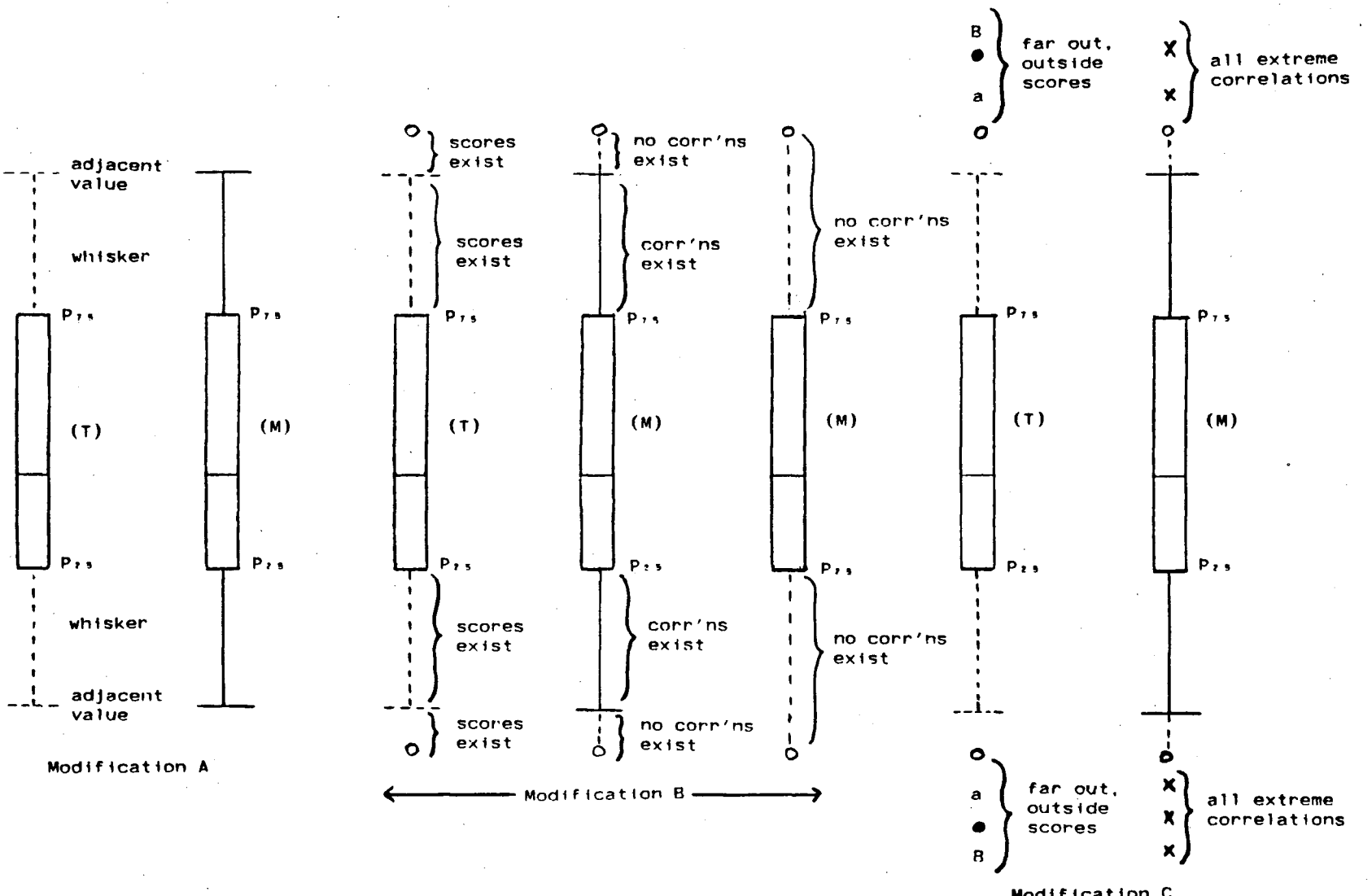


FIGURE 8.4: MODIFICATIONS (M) OF TUKEY'S TECHNIQUES (T) USED IN THE PRESENT STUDY

EXPLORATORY DATA ANALYSIS: THE APPLICATION

The modifications to Tukey's techniques were applied to display the reported LPC-performance correlations ("rho + r" listing) within each of the eight octants. For purposes of this analysis, these data were partitioned first by comparability category and then by task group type. The two stratification variables, rather than the study characteristic covariates, were selected for three reasons. First, it will be recalled from Chapter 5 that two comparability categories were created to reflect the extent to which any given study adhered to Fiedler's methods. This grouping of studies, together with the criticism that the Contingency Model is methodologically-bound, made it necessary to establish that the results yielded by studies classified as "CC:Fiedler" and "CC:Other" were more similar than they were different.

Second, as indicated in Chapter 1, the classification of task groups as interacting and coacting forms an essential part of Fiedler's Model. This distinction, together with the subdivision of interacting groups on the basis of statement or inference and the creation of "nonacting" groups, necessitated that the results generated from each type of task group be examined for similarities and differences. Third, the use of the two stratification

variables made it possible to include all 249 reported correlations in the analysis.

The Exploratory Data Analysis was based on individual octants. If the partitioned data sets are to be combined, all octants should show more similarities than differences across not only comparability categories but also task group types. Should one or more octants be more different than similar for any comparison, combining would not be appropriate. The exploratory analysis of the task group types follows that of the comparability categories.

Analysis of Comparability Categories

The graphic displays of the two comparability categories for each octant are shown in Figure 8.5. It will be recalled that similarity is indicated by extent of overlap between the pairs of boxes and whiskers; differences by the amount of separation. On this basis, the distribution of the reported correlations for CC:Fiedler (shaded box) and CC:Other (unshaded box) appear to be quite similar in Octants I, II, III, VII, and VIII. By contrast, Octants IV, V, and VI show the comparability categories to be quite different one from the other. A comparison of the interquartile ranges for Octants IV and V in particular indicates that the middle 50% of the CC:Fiedler values are grouped above the median in CC:Other. In Octant IV, more than half the CC:Other reported correlations within the

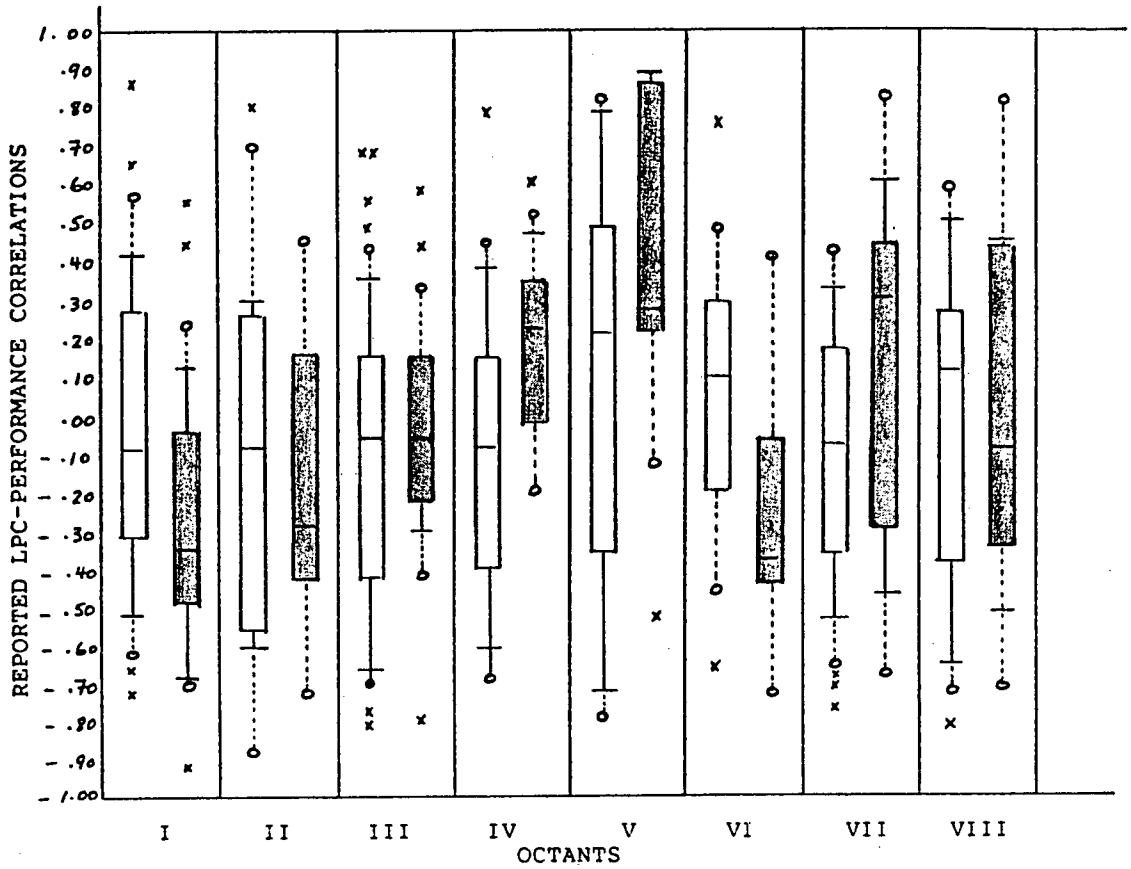


FIGURE 8.5: BOX-AND-WHISKER: COMPARABILITY CATEGORIES

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

CC:OTHER CC:FIEDLER

interquartile range are negative while all the reported correlations in CC:Fiedler are positive. The reverse is the case for Octant VI. Moreover, the median values in Octants IV and VI for both comparability categories are in the opposite direction one from the other. The absence of adjacent values in Octant VI is a clear indication of less variablility in the reported correlations. In CC:Other, for both Octants I and III, there is a greater number of more extreme values (outliers) than in CC:Fiedler. Overall, the within-octant comparisons suggest that there is not sufficient similarity across all octants of the Model to permit combining the two comparability categories.

Analysis of Task Group Types

There are six paired comparisons possible for the four task groups. Of these six, only the stated interacting (TGT 1) and inferred interacting (TGT 4) pairing had a sufficiently large number of reported correlations in most octants to make a meaningful box-and-whisker display of the data. Although the graphics were made for each of other five pairings, the presence of small or empty cells in Octants II, IV, VI, and VIII for coacting groups (TGT 2) and nonacting groups (TGT 8) rendered these comparisons much less useful. However, in an attempt to make some limited observations about the comparability of the findings generated by the task groups, the box-and-whiskers for

stated interacting groups (TGT 1) and nonacting groups (TGT 8) and for stated interacting groups (TGT 1) and coacting groups (TGT 2) are included as examples of the five comparisons. Brief reference will be made to the other three pairings, namely inferred interacting with nonacting, coacting with inferred interacting, and nonacting with coacting.

Comparison of stated and inferred interacting task groups. The reported correlations for task group types stated to be interacting (TGT 1) and inferred to be interacting (TGT 4) are shown schematically in Figure 8.6. The most striking feature of this plot is the separation of the two task group types in Octant VIII. The interquartile ranges not only show no overlap but the lower limit of TGT 1 is 20 correlation values higher than the upper limit of TGT 4. There is also a difference in the variability of the reported correlations for each task group type. With the exception of the outliers, there are no correlations below the 25th percentile for stated interacting groups and none above the 75th percentile for inferred interacting groups.

By contrast, the two task groups are more similar than different in Octants I, II, VI, and VII. In Octants III, IV, and V, the distance between the median values suggests that, in spite of the adjacency, some differences do exist. In Octants III and IV, the medians for the

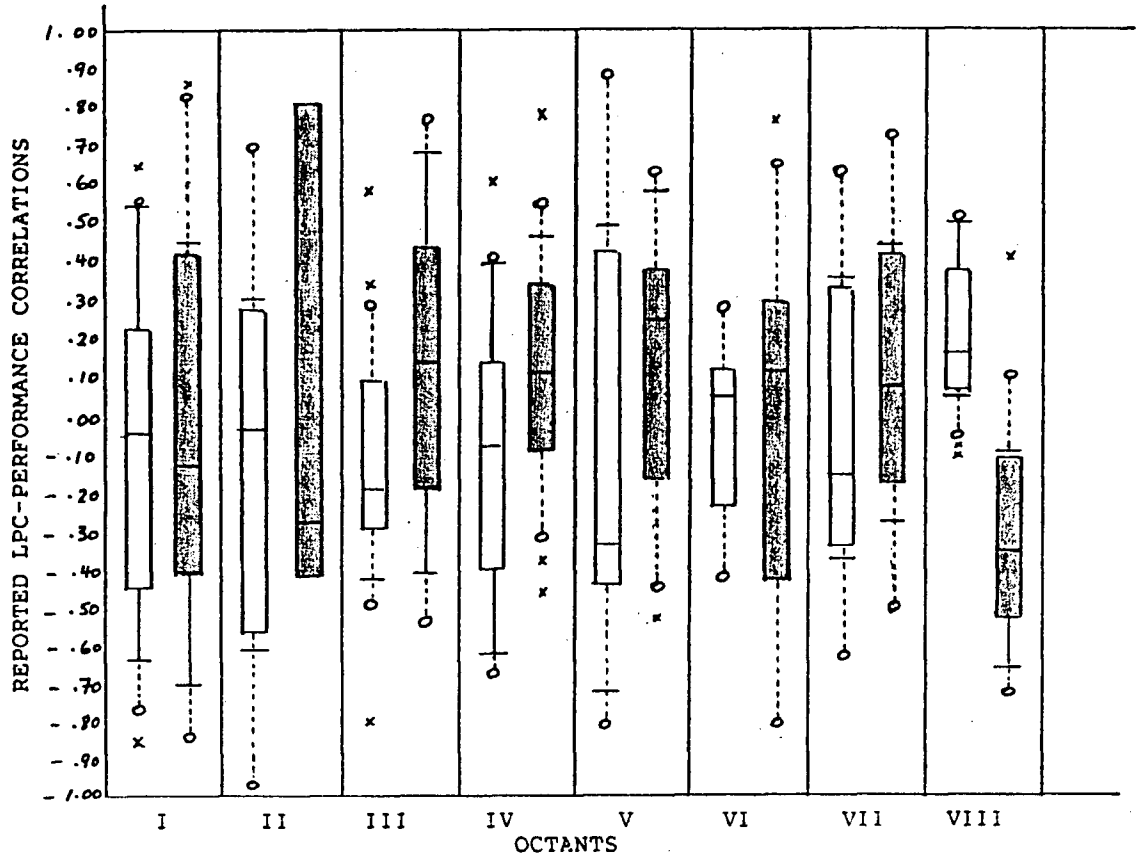


FIGURE 8.6: BOX-AND-WHISKER: TASK GROUP TYPES 1 AND 4

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

TGT 1 TGT 4

inferred interacting groups are either clearly above or just below the upper end of the stated interacting group box. Moreover, the pairs of medians in all three of these octants (III, IV, and V) have opposite signs. The lack of agreement in direction is of particular importance for a Model in which the sign of the correlation is often held to be more important than its magnitude.

The visual display indicates that, within octants, the differences between stated and inferred interacting task groups exceed the similarities. For this reason the two task groups cannot be combined. A supplementary box-and-whisker analysis of these two interacting task groups within each comparability category is reported in Appendix F (Figures F.1 and F.2, pp. 356 and 357). The results are consistent with the above conclusions.

Comparisons of stated interacting groups with coacting and nonacting groups. The graphic displays comparing stated interacting task groups (TGT 1) with each of nonacting (TGT 8) and coacting (TGT 2) task groups (Figures 8.7 and 8.8, respectively) permit meaningful observations only in Octants I, III, V, and VII. In the other four octants there were only two reported correlations for nonacting groups and none for coacting groups.

The "stated interacting-nonacting" display (TGT 1-8: Figure 8.7) shows considerable overlap between the

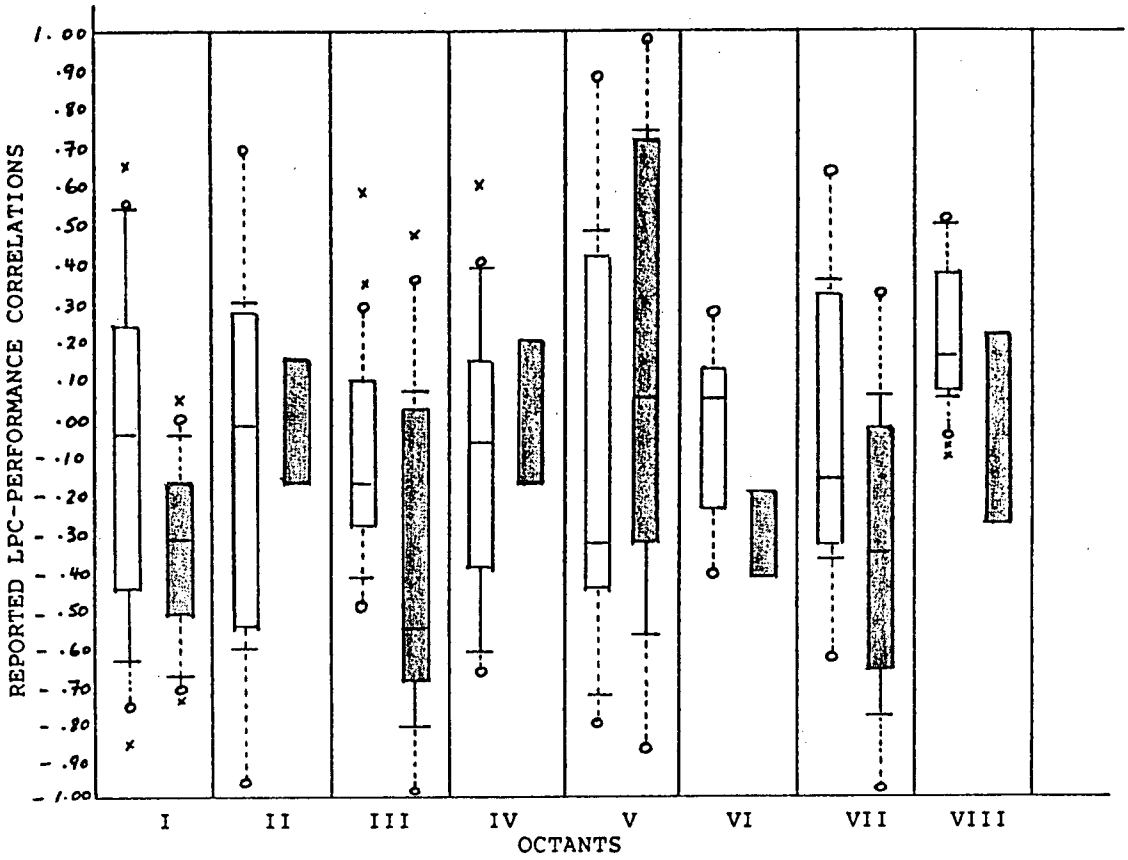


FIGURE 8.7: BOX-AND-WHISKER: TASK GROUP TYPES 1 AND 8

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

TGT 1 TGT 8

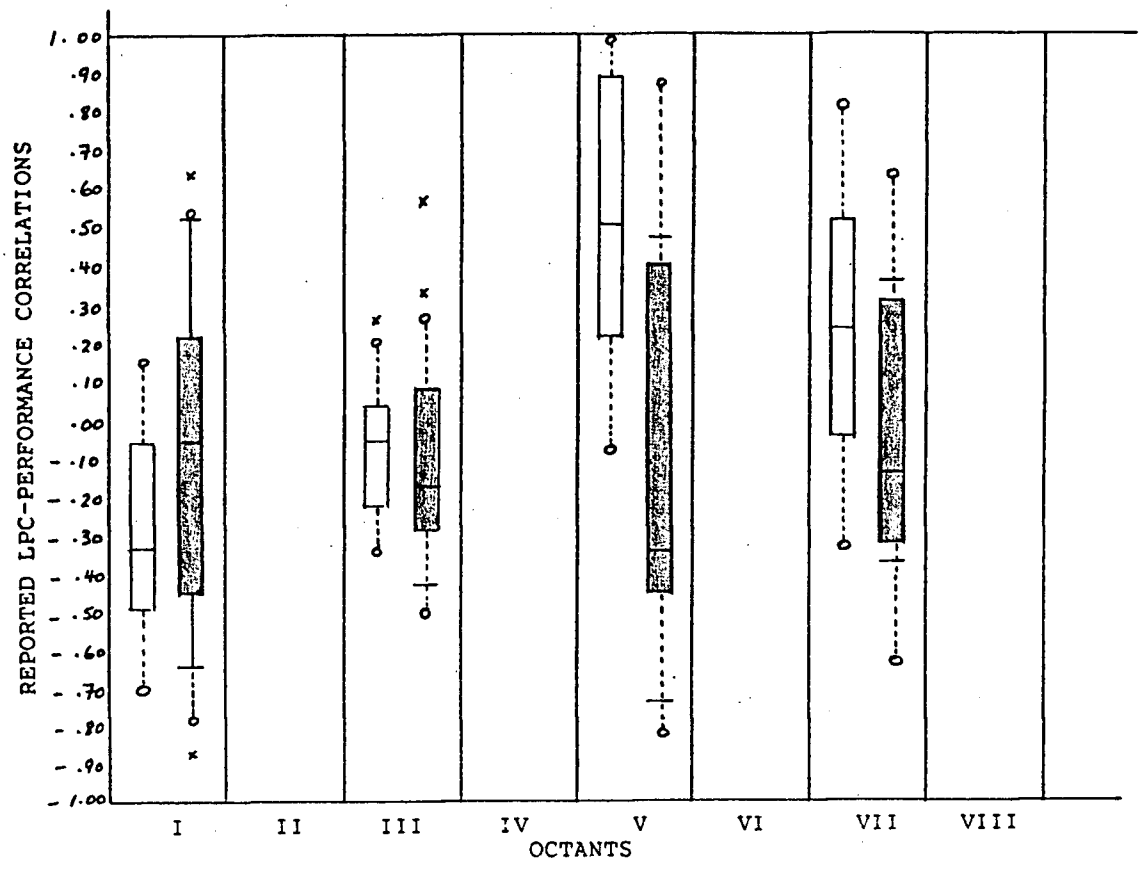


FIGURE 8.8: BOX-AND-WHISKER: TASK GROUP TYPES 1 AND 2

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

TGT 2 TGT 1

box-and-whisker pairs within the odd numbered octants. Despite the lack of directional agreement in Octant V, the within-octant distributions are sufficiently similar to permit combining. Because of the small number of relationships for coacting groups (TGT 1-2: Figure 8.8) in Octants I, III, V, and VII (n=5, 6, 4, and 4, respectively), the comparisons focus only on the interquartile ranges. Viewed in this way, there appears to be a difference between the distributions for stated interacting and coacting groups in Octant V. In Octants I, III, and VII it is likely, given the adjacency of the boxes, that the two task groups are more similar than they are different. However, the lack of consistent similarity across all octants for both comparisons prevents the collapsing of stated interacting task groups with either coacting or nonacting groups.

Comparisons of inferred interacting, coacting, and nonacting task groups. The other three comparisons involving inferred interacting (TGT 4), coacting (TGT 2), and nonacting (TGT 8) groups showed a pattern of similarities and differences not unlike those in the previous pairings. The three displays are shown in Appendix F (pp. 358-360). The only two groups to reveal a clear difference were the coacting and nonacting in Octant VII (TGT 2-8: Figure F.3). The "coacting-inferred interacting" comparison (TGT 2-4: Figure F.4) suggested that these two

groups may be different in Octants I, III, and V. In the remaining octants for both these comparisons, and in those from the inferred interacting and nonacting groups (TGT 4-8: Figure F.5), the overlap of the interquartile ranges suggested that the task groups were at least as similar as they were different. A more detailed discussion of these three pairings is included in Appendix F (pp.358-360).

Summary of Box-and-Whisker Observations

The observations gleaned from the box-and-whisker displays are summarized in Table 8.1. The relationship between each of the seven comparisons and the question of combining the partitioned data sets is reported in terms of "yes", meaning sufficiently similar to combine; "no", meaning differences prevent combining; and "?", meaning the distinction is not clear enough to tell if combining is legitimate.

As noted earlier, the decision to combine the partitioned data sets was contingent upon the presence of more within-octant similarities than differences for both the comparability categories and the task group types. Moreover, it was stated that the similarities should be consistently greater than the differences across all octants for any one comparison. The summary in Table 8.1 clearly indicates that neither of these conditions prevails.

TABLE 8.1: SUMMARY OF BOX-AND-WHISKERS OBSERVATIONS

OBSERVED PAIRS	OCTANT I	II	III	IV	V	VI	VII	VIII
CC:F-CC:O	Yes	Yes	Yes	No	No	No	Yes	Yes
TGT:1-4	Yes	Yes	No	No	No	Yes	Yes	No
TGT:1-8	Yes		Yes		Yes		Yes	
TGT:1-2	Yes		Yes		No		Yes	
TGT:4-8	No		?		Yes		?	
TGT:2-4	Yes		Yes		?		Yes	
TGT:2-8	Yes		?		Yes		No	

Key: Yes = combining legitimate
No = combining not legitimate
? = cannot judge legitimacy of combining

CONCLUSION

The purpose of the box-and-whisker analysis was to resolve the issue of whether or not the partitioned data sets could be combined. The use of Tukey's modified procedures have made it clear that the two comparability categories and the four task group types cannot be collapsed. Regardless of which stratification variable was used to compare them, the reported correlations did not yield consistently similar patterns either within or across octants. This outcome resulted in the decision to

discontinue the meta-analysis for the purpose of establishing which Contingency Model-based findings are consistent, reliable, and valid[2]. The aggregated data do, however, make it possible to ascertain whether there are any empirical reasons which may help to account for the inconsistent pattern of results yielded by the LPC-performance correlations. The search for such reasons constitutes the analysis reported in the next chapter.

---

[2] The original plan of this study provided for Fisher Z transformations of both the reported and antecedent LPC-performance correlations, t tests for significant differences between them, and effect size calculations based on the deviation of the correlations around the mean.

Chapter 9

EMPIRICAL INVESTIGATION OF STUDY VARIABILITY

The preceding three chapters have made it clear that no substantial and unambiguous body of empirical findings has emerged from the studies used in the present meta-analysis. The question of why this should be is important and the data used in the present study permit its investigation. Essentially, the question examined in this chapter is: to what extent is the methodological variability found in the primary studies associated with the variation in their findings? The first major section describes the methodological variability; the second analyzes the relationship between that variability and support for the Contingency Model. The final section presents some conclusions based on both the description and analysis of the methodological variability.

DESCRIPTION OF METHODOLOGICAL VARIABILITY

One of the surprising findings of the present study was the amount of variability in the methods used in the primary studies. The fact that different correlations (rho and r) were used has already been noted, but what is also clear is that no element of Fiedler's Model was treated

uniformly across all studies. The following sections describe the major variations found in the treatment of leadership style (LPC) and the three situational variables.

Variability in Leadership Style (LPC)

Although leadership style was uniformly measured by the LPC instrument, it was not always measured by the same LPC instrument and the scores obtained were not always treated in the same way. Which particular instrument was used (16 items, 18 items) is perhaps less important than is the question of how the obtained scores were treated. This question has two different aspects. The first aspect concerns the method used to dichotomize a group of LPC scores into high or low; the second, the designation of particular scores as high or low LPC.

Methods of score division. Three distinct methods of determining high and low LPC were found. The first is best described as the "relative" method. For instance, subjects were dichotomized relative to the sample mean or median (e.g., Graen et al., 1971a; Grunstad, 1972). In the second method, labelled "extreme", subjects whose LPC scores fell, for example, in the upper and lower thirds of the sample distribution or above and below 0.50 standard units from the mean were designated high or low LPC, respectively (e.g., Rice and Chemers, 1973; Stemler, 1980). The third method of

dividing scores, termed "normative", declared leaders to be high or low LPC on the basis of Posthuma's (1970) norms (e.g., Turner, 1972; Brown and Smolinsky, 1974). In addition to these three methods of determining high and low LPC, ten studies (representing 16 samples) did not divide the scores. Instead, the total LPC scores were treated as values on a continuous variable (e.g., Bell, 1970; Mellor, 1974). A further seven studies (representing ten samples) did not report the method of score division. One study (Sashkin, 1970) (representing four samples) used a six-point scale and was therefore not comparable with the other studies.

Score designation. The second aspect of variability in the treatment of LPC is the designation of which scores are high or low, respectively. The common metric needed to examine this was created by calculating the "item mean" scores (i.e., sum of the item scores divided by the number of items). Only 25 samples (13/38 studies) reported cut-off scores for high and low LPC. Table 9.1 displays the item means for these 25 samples. The range of item means, derived from the scores of 584 leader-subjects across the 25 samples, ran from 2.1 to 5.9 on a scale of one to eight. Values from 4.5 to 5.9 were found exclusively within the scores designating subjects as high LPC. Values from 2.1 to 3.2 were exclusive to the group of scores designating

TABLE 9.1: DESIGNATION OF LPC SCORES AS HIGH OR LOW IN 25 SAMPLES (item mean scores)

High LPC	Low LPC
5.9	4.4
5.4	4.1
4.9	-----
4.5	3.7
4.5	3.5
4.4	3.5
4.1	3.5
4.1	3.2
4.1	3.2
3.8	3.2
-----	2.5
3.6	2.2
3.6	2.1
3.5	
n=294 subjects	n=290 subjects

----- High LPC scores below broken line are below normative mean of 3.71; low LPC scores above broken line are above normative mean of 3.71 (Posthuma, 1970:11).

subjects as low LPC. The scores ranging from 3.5 to 4.4 were high or low LPC depending on the study at hand. These high and low LPC score designations represent a substantial departure from the prescribed cut-off values. The high LPC values below the broken line in Table 9.1 are lower than Posthuma's normative mean of 3.71 (1970:11). Two low LPC values, above the broken line, are higher than the normative mean.

Further, not only are most of the values for low LPC above Fiedler's recommended range of 1.2 to 2.2 but also two of the low LPC values are at or above the suggested minimum of 4.1 for high LPC (1967:44). Moreover, the score designations do not follow Fiedler's later recommendation that values of about 2.0 and 5.0 be used to indicate, respectively, low and high LPC leaders (1971d:129). The item mean scores made it clear that in over half the studies, leaders who were high LPC in one study could have been low LPC in another.

Variability in Leader-Member Relations

It will be recalled from the explanation of the Model in Chapter 1 that leader-member relations is measured by the Group Atmosphere Scale (GAS). According to Fiedler, the level of leader-member relations is assessed by the leader who, therefore, completes the GA scale. While the instrument itself was almost always the means used to

measure leader-member relations, the scale was frequently completed by respondents other than, or in addition to, the leader. The two kinds of variability described above in the treatment of LPC also occurred in the treatment of leader-member relations. Thus, there were three sources of variability to be considered here: (1) methods of score division, (2) score designation, and (3) assessors of leader-member relations.

Methods of score divison. As was the case with LPC, there were three distinct methods of determining "good" and "poor" leader-member relations. In the "relative" method, groups were dichotomized using either the sample median or mean GAS score (e.g., Martin, 1977; Blanchard, 1978, respectively). The "normative" method divided the groups with good and poor leader-member relations on the basis of Posthuma's (1970:12) item mean of 6.60 (e.g., Singe, 1975; Schneier, 1978). The third method, labelled "extreme", used the first and fourth quartiles of the GAS score distribution to designate groups as having, respectively, poor and good leader-member relations (e.g., Martin et al., 1976).

In addition to these three methods, there were a number of studies in which the score division method was unique to the particular study. For instance, only one study used upper and lower thirds of the sample distribution (Hill, 1969a); in another, all the scores were considered

high (Hunt, 1966); yet another used the GAS scores to validate the leader-member relations in sociometrically chosen groups (Saha, 1972); one partialled out all scores between 53 and 57 (Turner, 1972); another used GAS as a post-task validation measure in a laboratory setting (Chemers et al., 1974); of the two studies using scale midpoints, one used a theoretical item mean of 4.5 (Faust, 1972) while the other used the exact middle value of 40.0 (Lanaghan, 1972). There were several studies which either assumed the level of leader-member relations or did not report the score division procedures used to dichotomize the groups.

Score designation. Since all studies used the ten item GAS instrument, the task of comparing scores for this analysis was less difficult than for the LPC scores. Across the studies in which leader-member relations was measured, the total score ranges for good and poor leader-member relations were, respectively, 45 to 69 (n=29 groups) and 44 to 67 (n=26 groups). The two ranges indicate that only one score (44) is exclusive to poor leader-member relations. In other words, in any given study scores between 45 and 67 may result in groups being designated as having good or poor leader-member relations. The overlap in GAS scores is shown in Table F.4 in Appendix F.

Unlike the clearly stated specifications for high and

low LPC, Fiedler offers little direction for declaring good and poor leader-member relations. He does indicate that either a median split (1967:158) or a trichotomy of the scores (1967:115) may be used. More explicit information is provided by Posthuma (1970) in his secondary analysis of 2415 GAS scores. As noted earlier, the item mean was 6.60 on the standard ten item-eight point scale. Applying this benchmark value of 66 to the reported distributions, all of the "poor" GAS scores except one (i.e., 67) fell below the normative mean. By contrast, only three values (67, 68, and 69) in the "good" distribution were above the normative score. If the scores were re-examined by designating a value of 67 as a "good" score (normative mean of 6.6 raised to 67), then only five of the 25 "good LMR" scores would remain as "good" and two of the 22 "poor LMR" scores would become "good LMR".

Assessors of leader-member relations. Fiedler (1967, 1978) has conceptualized this variable in terms of the leader's perception of leader-member relations. The rationale for this leader-centered point of view rests on the assumption that a leader's behaviour in any given situation will depend, in part, on the extent to which he or she perceives the group members to be supportive and loyal irrespective of what the "reality" of the situation may be. It is for this reason that Fiedler's method requires the

leader to complete the GA scale. Not all primary authors, however, agreed that this perspective yielded an accurate appraisal of leader-member relations. Because these authors believed that the variable was better conceptualized as "group atmosphere", the GAS instrument was completed by either the group members only or both the group members and the leader. In the latter case, the GAS scores were combined to yield one composite score by which to designate the group as having good or poor leader-member relations.

Variability in Task Structure Assessment

That someone other than the leader should assess the degree of structure is implicit in the instructions for completion of the task structure instrument (Fiedler, 1967:282). Fiedler and Chemers (1974) indicate that judges should be used (1974:68). Fiedler (1978) states that task structure "should be assessed by the leader's superior [because] self assessments are subject to distortion" (1978:65). Yet, inspite of this testing specification, 11 out of the 27 samples in which task structure was actually measured used either leaders or group members to assess the extent to which the group's primary task was structured or unstructured. In the remaining 22 samples, task structure was not measured, either being held constant or assumed.

Variability in Position Power

The amount of power invested in the position should also be assessed, according to Fiedler, by persons other than the leader. The literature cited in the preceding section regarding the appraisal of structure applies equally to that of position power. It was found, however, that three different groups of assessors — judges (in accordance with Fiedler's method), group members only, and group members plus the leader were used as respondents. The extent to which the variation in the assessors of both position power and task structure may have contributed to the lack of agreement between the antecedent values and the obtained results is reported in the next section of this chapter.

ANALYSIS OF METHODOLOGICAL VARIABILITY

The purpose of this analysis was to determine whether the kinds of methodological variability described above contributed to the disparity between Fiedler's predicted outcomes and the results obtained by aggregating the findings across the primary studies. Operationally, the effects of methodological variability were assessed by reference to the frequency of support for Fiedler's Model. Two indicators of support and non-support were used: (1) the

authors' stated conclusion regarding support for the Model (full, partial, none) and (2) the directional congruency between the predicted medians and the reported correlations. Each of these indicators was crosstabulated with the variations found across the studies. The resulting crosstabulations vary in the total number of samples and octant tests which they display. The reason for this change in n from analysis to analysis lies in the fact that not all primary authors measured all variables. The results are reported first for the methods of score division used with LPC and GAS and second, for the assessment of each situational variable. These five crosstabulations deal with the five kinds of variability described above. A further analysis was performed to examine the effects associated with the use of the summary statistics rho and r, respectively.

Crosstabulation Results for LPC Score Division Methods

It will be recalled from the earlier description that the LPC scores were divided into "high LPC" and "low LPC" by one of four main methods. These methods were relative, normative, extreme, and "totals only". Each method was crosstabulated with three different declarations of support by primary authors and with the amount of support indicated by the frequency of directional congruency.

Authors' stated conclusions. The results of this analysis are shown in Table 9.2. An inspection of the cell

TABLE 9.2: LPC SCORE DIVISION AND SUPPORT FOR MODEL AS INDICATED BY AUTHORS' STATED CONCLUSIONS

Score Division Method	Degree of Support Full	Partial	None	Row total
Relative	1	5	3	9
Normative	1	3	0	4
Extreme	2	4	0	6
Totals Only	0	10	6	16
Column Total	4	22	9	35

frequencies shows that the "totals only" method is more often associated with partial support than is any other method. When partial and full support are combined, there is little to chose among the other three methods. Overall it would appear that the method of dividing LPC scores does not help to explain the disparities among study findings, using the authors' stated conclusions.

Directional congruency. The frequencies of support and non-support for each method, grouped across all octants, are shown in Table 9.3. Regardless of whether the support and non-support frequencies are expressed as a proportion of the overall total (n=181 octant tests) or of the column totals, the relative method shows more non-support than

TABLE 9.3: LPC SCORE DIVISION METHODS AND FREQUENCY OF SUPPORT AS INDICATED BY DIRECTIONAL CONGRUENCY

Score Division Method	Frequency Support	Non-Support	Row Total
Relative	31	39	70
Normative	16	9	25
Extreme	20	4	24
Totals Only	34	28	62
Column Total	101	80	181

support (non-support=22% and support=17% of the overall total; non-support=49% and support=31% of the column totals). The "totals only" method (which does not divide LPC into high and low) shows a different picture. This method, viewed as a percentage of all octant tests, showed support more often than non-support (19% and 15%, respectively). However, as a proportion of either support or non-support, this method produces virtually identical results (34% and 35%). Both the normative and extreme methods yielded more support than non-support regardless of the proportional base (i.e., overall total or column totals). The non-division ("totals only") method accounted for 34% of the directionally congruent results. The other three division methods together accounted for the remaining 66% of the support. It appears, on the basis of both row percentages and cell frequencies, that the extreme method is the most likely to generate findings supportive of the Model, using the criterion of directional congruency.

Comparison of the two indicators. A comparison of the results from the crosstabulations indicates that, regardless of which criterion is used, the "totals only" method is most frequently associated with support for the Model. The only method to be associated with more non-support than support is the relative method when directional congruency is used as the indicator. An inspection of the difference between support and non-support for each method reveals that the use of extreme scores yields proportionally more support irrespective of the indicator. However, the cell frequencies are not sufficiently large to permit any conclusion regarding the method of score division and support for the Contingency Model. Neither of these crosstabulations has yielded results which help to account for the disparity between the reported and antecedent correlations.

Crosstabulation Results for GAS Score Division Methods

The same four score division methods used to classify the procedures for dichotomizing the LPC scores were applied to the GAS scores. It will be recalled that this scale operationalizes the leader-member relations variable. These four categories of methods were crosstabulated with the two indicators of support for the Model in order to determine if

there was a systematic relationship between support for the Model and the method of dividing scores.

Authors' stated conclusions. The results of this analysis are shown in Table 9.4. An inspection of the cell

TABLE 9.4: GAS SCORE DIVISION AND SUPPORT FOR MODEL AS INDICATED BY AUTHORS' STATED CONCLUSIONS

Score Division Method	Degree of Support Full	Partial	None	Row total
Relative	0	14	5	19
Normative	5	8	4	17
Extreme	0	3	0	3
Totals Only	0	1	0	1
Column Total	5	26	9	40

frequencies revealed that there was virtually no difference between the relative and normative methods when full and partial support were combined. The fact that only four studies used the extreme and totals only methods makes it impossible to assess their effects. Table 9.4 suggests that there may be a relationship between the way in which groups are deemed to have good or poor leader-member relations and the support conclusion reached by the author(s) of a study. It is therefore possible that some of the disparity between the obtained and antecedent values may be explained by methodological variability in the measurement of leader-member relations.

Directional congruency. The frequency of support and non-support for each method, grouped across all octants, is shown in Table 9.5. Those studies which used the extreme

TABLE 9.5: GAS SCORE DIVISION METHODS AND FREQUENCY OF SUPPORT AS INDICATED BY DIRECTIONAL CONGRUENCY

Score Division Method	Frequency Support	Non-Support	Row Total
Relative	43	45	88
Normative	57	54	111
Extreme	8	20	28
Totals Only	6	0	6
Column Total	114	119	233

method (e.g., upper and lower thirds of the sample distribution) yielded more than twice as many non-supportive as supportive results. The other observation to be made about the division of GAS scores concerns the frequency not only with which the relative and normative methods were used but also the amounts of support and non-support with which they were associated. Although the relative method was not used as often, it yielded roughly the same frequencies of support and non-support as did the normative method. However, when viewed as a percentage of column totals, the normative methods produced both more support (50%) and

non-support (45%) than did the relative procedure in which the percentages were the same (38% support, 38% non-support). With the possible exception of the extreme method, methodological variability with respect to the division of GAS scores seems to have very little connection with directional congruency.

Comparison of the two indicators. The frequency of support (full and partial) associated with the relative method is considerably greater than non-support (14 to 5, respectively) when the authors' stated conclusion is used as the indicator. When the indicator is directional congruency, the frequency of support and non-support is almost identical (43 and 45, respectively). Regardless of which indicator is used, the normative method produces a pattern of results very similar to that of the relative method. The extreme and totals only methods are more frequently linked with either limited support or non-support than they are with support. Taken together, the crosstabulation results suggest that support for the Model is more often associated with the normative and relative methods of dividing GAS scores. However, there is no clear indication that the use of the extreme and totals only methods accounts for the disparity between the obtained results and those predicted by Fiedler.

Crosstabulation Results for Assessment of Leader-Member Relations

The earlier description of methodological variability identified three different groups of respondents who assessed the level of leader-member relations on the GAS instrument. The procedure prescribed by the Model indicates that the leaders should complete the scale. It is the purpose of this analysis, using both authors' stated conclusions and directional congruency, to ascertain whether adherence to Fiedler's recommendations yields more support for the Model than does non-adherence.

Authors' stated conclusions. As an indicator of support, the authors' stated conclusions were crosstabulated with the groups of respondents who completed the GAS instrument. The results are shown in Table 9.6. The cell frequencies indicate that Fiedler's procedure was followed in nearly three quarters of the samples (31/42). Moreover, of those 31 samples, 25 found some measure of support for the Model. It would appear from the obtained frequencies that methodological adherence with respect to the operationalization of leader-member relations is linked with the authors' conclusions regarding support.

TABLE 9.6: GAS INSTRUMENT COMPLETION AND SUPPORT FOR MODEL AS INDICATED BY AUTHORS' STATED CONCLUSIONS

LMR Assessors	Degree of Support Full	Partial	None	Row total
Leaders	6	19	6	31
Group Members	0	3	0	3
Leadrs+Grp Membrs	2	4	2	8
Column Total	8	26	8	42

Directional congruency. The frequency of support and non-support for each group of assessors is shown in Table 9.7. It is interesting to note that nearly all the octant

TABLE 9.7: GAS INSTRUMENT COMPLETION AND FREQUENCY OF SUPPORT AS INDICATED BY DIRECTIONAL CONGRUENCY

LMR Assessors	Frequency Support	Non-Support	Row Total
Leaders	106	65	171
Group Members	5	11	16
Leaders+Group Members	18	34	52
Column Total	129	110	239

tests (n=249) are represented in the total[1]. In light of the fact that this instrument was used in 96% of the tests, it would appear that there is much greater adherence to

[1] In the remaining ten octant tests, the level of leader-member relations was assumed to be good or poor; no validation of the assumption was reported.

Fiedler's method in the measurement of leader-member relations than in that of either task structure or position power.

In addition to the very consistent use of the prescribed instrument in the operationalization of leader-member relations, over 71% (171/239) of the octant tests show that the assessment was done in accordance with Fiedler's method. Moreover, when subjects other than the leaders completed the instrument, non-supportive results were more frequent than supportive ones. Of the 171 tests in which the leaders themselves filled out the GA scale, 71% (106/171) yielded directional congruency. However, it should also be noted that 59% (65/110) of the non-support came from this same set of octant tests. By contrast, when either of the other two groups were asked to assess the level of leader-member relations, non-support always exceeded support — by a margin of almost two to one. These outcomes provide some evidence that deviation from Fiedler's prescribed method is related to the lack of correspondence between the obtained and antecedent values.

Comparison of the two indicators. Both the authors' stated conclusions and the directional congruency indicators show that methodological adherence is more frequently associated with support and that methodological deviation is more frequently associated with non-support. The cell

frequencies for both indicators suggest that methodological variability, at least in the assessment of leader-member relations, helps to account for the disparity between Fiedler's predictions and the obtained results.

Crosstabulation Results for Assessment of Task Structure

The earlier description of methodological variability indicated that respondents other than the leader's superiors or judges assessed the degree to which the group's primary task was structured or unstructured. It was the purpose of this analysis, using both indicators of support, to ascertain whether the use of assessors different from those prescribed by Fiedler was connected with frequency of support for the Model.

Authors' stated conclusions. The results of this analysis are shown in Table 9.8. Although the cell sizes are too small for any definitive statement, two trends are apparent. When either superiors or external judges are used to assess task structure, there is a greater tendency to conclude at least some support for the Model. By contrast, less support and more non-support is found when either leaders or group members assess the degree of structure. In light of these frequencies, it would seem that less adherence to Fiedler's method may yield less supportive findings across studies.

TABLE 9.8: TASK STRUCTURE ASSESSMENT AND SUPPORT FOR MODEL AS INDICATED BY AUTHORS' STATED CONCLUSIONS

TS Assessors	Degree of Support Full	Partial	None	Row total
Superiors	1	8	1	10
Judges	0	5	1	6
Leaders	1	2	3	6
Group Members	0	3	1	4
Column Total	2	18	6	26

Directional congruency. This crosstabulation was based on the same four groups of assessors used in the analysis of the authors' stated conclusions. The frequencies displayed in Table 9.9 suggest the absence of any systematic relationship between support and procedure.

TABLE 9.9: TASK STRUCTURE ASSESSMENT AND FREQUENCY OF SUPPORT AS INDICATED BY DIRECTIONAL CONGRUENCY

TS Assessors	Frequency Support	Non-Support	Row Total
Superiors	24	17	41
Judges	19	16	35
Leaders	19	16	35
Group Members	7	10	17
Column Total	69	59	128

Neither the cell sizes nor the differences between the support and non-support frequencies are large enough to form the basis for any definitive statement. It can only be observed that there is considerable departure from the prescribed method.

When the four groups are combined with respect to procedural adherence and non-adherence, Fiedler's method is followed in about 59% (76/128) of the octant tests. Together, "superiors" and "judges" accounted for 62% (43/69) of the directional congruency. However, these same two groups also revealed 56% (33/59) of the non-support. As assessors of task structure, "leaders" and "group members" together were associated with a greater proportion of non-support than support (44% and 38%, respectively). The results of this crosstabulation are in agreement with those from the authors' stated conclusions analysis in that methodological variation in the treatment of the task structure variable appears to make little difference in the frequency of support yielded by the primary studies.

Crosstabulations Results for Assessment of Position Power

The earlier description of methodological variability indicated that the level of position power (strong or weak) should be assessed by either the leader's superior or by judges. In some studies, however, position power was assessed by the leaders themselves or by the group members

or by both the leaders and the group members. It was the purpose of this analysis to ascertain whether these three deviant procedures may have contributed to the lack of correspondence between the antecedent and obtained results.

Authors' stated conclusions. The frequencies yielded by the crosstabulation of this indicator with the five groups of assessors are displayed in Table 9.10. Clearly,

TABLE 9.10: POSITION POWER ASSESSMENT AND SUPPORT FOR MODEL AS INDICATED BY AUTHORS' STATED CONCLUSIONS

PP Assessors	Degree of Support Full	Partial	None	Row total
Superiors	1	3	1	5
Judges	0	4	2	6
Leaders	1	2	2	5
Group Members	0	2	1	3
Leaders+Group Members	0	2	0	2
Column Total	2	13	6	21

the cell sizes are too small to permit any meaningful statement regarding the relationship between frequency of support and methodological variability.

Directional congruency. This crosstabulation was based on the same five groups of assessors used for the analysis of the authors' stated conclusions. The results

are shown in Table 9.11. The cell frequencies indicate that

TABLE 9.11: POSITION POWER ASSESSMENT AND FREQUENCY OF SUPPORT AS INDICATED BY DIRECTIONAL CONGRUENCY

PP Assessors	Frequency Support	Non-Support	Row Total
Superiors	22	17	39
Judges	11	8	19
Leaders	16	17	33
Group Members	9	3	12
Leaders+Group Members	7	5	12
Column Total	65	50	115

Fiedler's method (superiors or judges) is used as often as non-Fiedler methods (n=58 and n=57, respectively). When viewed as a proportion of the total (n=115), there is essentially no difference in the percentage of support and non-support yielded by any one assessment procedure. A comparison of the support and non-support columns shows little difference in the frequencies for each group of assessors. With the possible exception of "group members" (more support than non-support), it does not seem to matter how position power is assessed. As was the case with the assessment of task structure, there appears to be no systematic association between methodological variability and non-support for the Contingency Model.

Crosstabulation Results for the Summary Statistic

When the obtained values (i.e., median correlations) were compared with the antecedent values, the Pearson r and Spearman rho statistics yielded rather different results (see Chapter 6 and Appendix F). The findings, aggregated on the basis of the Pearson r, corresponded more closely with the predicted curve in Octants I, II, III, and VIII; the Spearman rho did so in the other four octants. In other words, for the Model as a whole, the aggregated results provided no clear-cut basis for selecting one summary statistic over the other — despite Fiedler's method which indicates use of the rank-order correlation. It is, therefore, of some interest to examine the relationship between the summary statistic used and the two indicators of support.

Authors' stated conclusions. The crosstabulated frequencies are shown in Table 9.12. A visual inspection of the cell frequencies would suggest that the summary statistic has little, if any, association with the authors' stated conclusions regarding support for the Contingency Model. It would appear that the correlation statistic used to report study findings does not help to account for the disparity in results when the authors' stated conclusions are used as the criterion.

TABLE 9.12: SUMMARY STATISTIC AND FREQUENCY OF SUPPORT FOR MODEL AS INDICATED BY AUTHORS' STATED CONCLUSIONS

Summary Statistic	Degree of Support Full	Partial	None	Row total
rho	2	19	5	26
r	6	16	5	27
Column Total	8	35	10	53

Directional congruency. The frequencies resulting from the crosstabulation of this indicator of support with the summary statistic used to report the findings for the octant tests are displayed in Table 9.13. There are three

TABLE 9.13: SUMMARY STATISTIC AND FREQUENCY OF SUPPORT AS INDICATED BY DIRECTIONAL CONGRUENCY

Summary Statistic	Frequency Support	Non-Support	Row Total
rho	79	53	132
r	56	61	117
Column Total	135	114	249

points to be made about these data. First, the Spearman rho has been used more often than the Pearson r to report the LPC-performance correlation. Second, as a proportion of the row totals, the Spearman rho is associated with 20% more supportive than non-supportive tests. The Pearson r is associated with roughly equal percentages of support and non-support (48% and 52%, respectively). Third, as a

proportion of the column totals, the Spearman rho is again associated with a higher percentage of support than is the Pearson r (59% and 41%, respectively). By contrast, the Pearson r is associated with a higher percentage of non-support than is the Spearman rho (54% to 46%, respectively). These results suggest that the recommended Spearman rho is associated with support more often than is the Pearson r. However, the pattern of cell frequencies suggests that the choice of summary statistic accounts for little, if any, of the lack of correspondence between the obtained and predicted results.

Summary of Results

The purpose of the analysis of methodological variability was to ascertain whether departures from Fiedler's prescribed procedures were associated with support for the Model. The relationship was assessed by using the authors' stated conclusions and directional congruency as indicators of support. The results for both indicators are summarized below.

1. LPC Score Division (Tables 9.2 and 9.3)

    1.1 there was considerable variation in methods used to divide scores;
    1.2 there was extensive overlap between high and low LPC scores (Table 9.1);
    1.3 approximately 50% of the samples did not report LPC scores;

1.4 the relative method was the most frequently used way to divide the LPC scores;
1.5 the extreme method showed highest percentage of support relative to non-support;
1.6 the "totals only" (i.e., non-division) was associated with support more often than any other method.

2. GAS Score Division (Tables 9.4 and 9.5)

2.1 there was less variability than for LPC but still considerable;
2.2 there was almost complete overlap between good and poor GAS scores;
2.3 the normative method was the most frequently used way to divide the GAS scores;
2.4 the "totals only" was the least used method;
2.5 the normative method was associated with greatest proportion of support, followed by relative method.

3. Leader-Member Relations Assessment (Tables 9.6 and 9.7)

3.1 the majority of studies used this scale to measure leader-member relations;
3.2 the cell frequencies show a trend toward a relationship between support and methodological adherence;
3.3 leaders were the most frequently used assessors of leader-member relations;
3.4 the highest support frequency was associated with leader-completion of GAS;
3.5 the lowest support frequency was linked with "leader+group member" completion of GAS.

4. Task Structure Assessment (Tables 9.8 and 9.9)

4.1 there was considerable departure from Fiedler's method;
4.2 the authors' stated conclusions and support were more frequently associated when task structure was assessed by superiors and judges;

4.3 task structure was assessed most frequently by leaders' superiors; least frequently by group members;
4.4 superiors and judges together were associated with the highest percentage of both support and non-support;
4.5 judges and leaders were used to assess task structure equally often with identical results for support and non-support.

5. Position Power Assessment (Tables 9.10 and 9.11)

5.1 there was more procedural variability than with task structure;
5.2 the frequencies indicated that methodological adherence was unrelated to the authors' stated conclusions;
5.3 position power was assessed most frequently by the leaders' superiors;
5.4 group members are involved least frequently in the assessment of position power;
5.5 support and non-support were associated about equally with each group of assessors.

6. Summary Statistic (Tables 9.12 and 9.13)

6.1 rho was used more frequently than r across all octants tests;
6.2 rho was associated with both greater frequency and proportion of support than was r;
6.3 r was associated with more non-support than support.

CONCLUSIONS AND DISCUSSION

The analysis of methodological variability has not provided a satisfactory explanation of the disparity between Fiedler's predictions and the results of the aggregated findings from the primary studies. The analysis has,

however, raised four potentially fruitful points, the discussion of which forms the substance of the final section of this chapter.

First, the crosstabulation results have indicated that support is more often associated with adherence to Fiedler's prescribed procedures for testing or applying the Model. Conversely, non-support is more frequently connected with departures from those procedures. This finding reinforces the criticism that the Contingency Model is method-bound (e.g., Graen et al., 1970; Barrow, 1977). The question raised by this analysis is not so much concerned with the extent of methodological adherence as with the fact that many of the studies purporting to test the Contingency Model have, in fact, tested a variation of the Model. This, of course, leaves unanswered the question of the appropriateness of the Model for all the situations in which it has been applied.

The second point raised by the consideration of methodological variability concerns the method used to dichotomize a group of LPC scores into high or low and the designation of particular scores as high or low LPC. The crosstabulation results point toward two seemingly antithetical conclusions. On the one hand, dividing subjects into high and low LPC is necessary if any octant of the Model is to be tested with experimental rigour. Moreover, such division is needed because of Fiedler's

(1967) distinction between leader behaviour and leadership style. On the other hand, the non-division of LPC scores — that is, simply using all the scores of all the subjects within a given octant for the calculation of either rho or r — simplifies the problem of retaining a sufficient number of participants to make the octant(s) testing a viable enterprise. Neither procedure is without reproach yet both are necessary if the Contingency Model is to be recognized as a valid tool for leadership research.

The third point raised by the analysis of methodological variability concerns the conceptualization of position power within the methodologically-deviant studies. Fiedler (1967, 1971d,1978) states that the amount of power a leader has is the same as the amount the organization vests in the position. Other researchers (e.g., Blanchard, 1978; Csoka, 1975; Mellor, 1974) consider this viewpoint to be both inadequate and inappropriate. They have suggested that it matters little what the organization believes to be the power ascribed to the role. Rather they focus on their attention on the role incumbent — on what the leader perceives his or her position power to be irrespective of the formal organizational structure.

The fourth point raised by the analysis of methodological variability concerns alternative explanations for the disparity between the antecedent and obtained results. The crosstabulations revealed, on the one hand,

very little evidence to account for the lack of association between frequency of support and the assessment procedures for task structure and position power. On the other hand, the analyses showed that there was a systematic relationship between the assessment of leader-member relations and the frequency of support across studies. In the absence of empirical evidence to account for the difference in the crosstabulation results, a speculative explanation is offered.

It is at least plausible that the difference lies in the nature of what is being perceived (i.e., assessed). When leader-member relations is evaluated by either the group members or the group members and the leader jointly, the perceptions of group members are likely to be quite different from those of the leader. That these differences in perception may exist is not totally without empirical support. It is noted by Mitchell et al. that:

> The leader's rating of group atmosphere seems particularly suspect. Its use depends upon the assumption that the leader can accurately judge the climate of the group, but it is his position which seems most likely to create the possibility of inaccurate perception (1970:13).

The authors provide some evidence of low inter-correlations between GAS scores when the instrument is completed by leaders and by group members, respectively. From a somewhat different perspective, Argyris observed that, although leaders require sensitivity to evaluate leader-member relations, Fiedler does not research the "predictive

validity of sensitivity" in high and low LPC persons (1976:651). It is also possible that the GA scale may be subject to response bias in that the leader may wish to have it believed that his or her group members are loyal and supportive — a view which may not be shared by the group members.

By contrast, both the structure of the task and the power vested in the position by the organization are much more stable in character and perhaps less subject to response bias. Both dimensions are likely to be perceived more similarly by all groups of assessors. This suggestion would help to explain why the assessor of leader-member relations is related to the frequency of support while the appraiser of task structure and position power apparently has very little connection with the frequency of support for the Contingency Model.

Whatever the explanation may be, it appears that methods of dividing both LPC and GAS scores, together with the assessors of leader-member relations, are associated with frequency of support for the Contingency Model. It also seems that methodological variation in the assessment of task structure and position power has little relationship with frequency of support. However, it is recognized that the presence of more than one procedural deviation in any single study could, across all studies, mask the overall effect of even minor departures from Fiedler's method. It

is possible that further research, specific to different combinations of methodological variations, may help to provide a more definitive answer.

Chapter 10

SUMMARY, DISCUSSION, AND IMPLICATIONS FOR FUTURE RESEARCH

SUMMARY OF FINDINGS

This study originated with the observation that Fiedler's Contingency Model of Leadership Effectiveness was among the most widely cited and yet the most controversial in the leadership literature. The study was designed to undertake a meta-analysis of the findings in Fiedler-based studies in order to try to establish which of them were consistent, reliable, and valid. If such findings were identified, then the concern would become the formulation of a theory which could account for them and from which hypotheses could be generated to guide future research. If consistent, reliable, and valid findings did not emerge, then the aim would be to determine the extent to which methodological, as opposed to theoretical, inadequacies might explain their absence.

The key elements of Fiedler's Model are leadership effectiveness, leadership style (LPC), and the situation (leader-member relations, task structure, and position power). The Model itself uses eight different configurations of the values of the situational variables to define eight octants. The Model predicts that a particular

direction and magnitude of correlation between performance and leadership style will be found for each octant. Thus, the studies to be subjected to meta-analysis would have to be those which included at least three elements: a measure of style, the identification of octants (or their configurations of the values of the variables) and the LPC-performance correlation.

From the original pool of 402 documents, 221 were found to be relevant to the Model. These 221 documents were of three kinds: (1) primary studies reported in sufficient detail to permit an appraisal of their quality, (2) primary studies in which the detail was not sufficient to allow a critical appraisal, and (3) non-primary studies (e.g., critiques or reviews of research). The primary studies constituted the potential data base for the meta-analysis.

In order to ensure the comparability of all the primary studies, each study was examined for the way in which the five variables of Fiedler's Model had been treated. This initial assessment of comparability revealed that 74 studies did not test individual octants of the Model (or did not use a correlation statistic to report the results). These studies, designated "Non-Octant", were deleted from the data base. The remaining 38 studies, labelled "Octant" studies, all tested one or more individual octants of the Model and reported the findings in correlational terms. In addition, each of the 38 studies

used some form of Fiedler's LPC scale, assessed the level of each situational variable in an acceptable way, and used a direct measure of effectiveness based on either group or leader performance. These studies comprised the data base for the meta-analysis.

Although all 38 of the octant studies displayed these five characteristics, they did not all do so to the same extent. A second assessment of comparability revealed that some studies adhered closely to Fiedler's method with respect to each element while others deviated in the treatment of one or two of the elements. These deviations were sufficiently substantial to warrant the creation of two categories of comparability. Those studies which adhered most closely to Fiedler's methods were designated "Comparability Category: Fiedler" (CC:Fiedler); those adhering less closely were designated "Comparability Category: Other" (CC:Other).

Since all octant studies reported an LPC-performance correlation, the first analysis examined the frequency of support for Fiedler's Model using two criteria of support: (1) the primary study authors' stated conclusions and (2) the directional congruency between the reported correlations and those predicted by Fiedler for each octant. The analysis of authors' conclusions showed that 15% (6/40) of studies declared full support for the Model while 22% (9/40)

concluded non-support[1]. The remaining 63% (25/40) declared partial support. An analysis by comparability category showed that of the 12 studies classified as "CC:Fiedler", two declared full support while ten concluded partial support. Of the 28 studies within the "CC:Other" group, four studies concluded full support, 15 declared partial support, and nine concluded no support for the Model.

The application of the directional congruency criterion of support to the 249 reported LPC-performance correlations (octant tests) showed that 54% (135/249) of the correlations were in the direction predicted by Fiedler while 46% (114/249) were opposite to Fiedler's predicted direction. Those studies classified as "CC:Fiedler" yielded a total of 83 octants tests of which 69% (57/83) were in the direction predicted by Fiedler while 31% (31/83) were opposite to Fiedler's predicted direction. The "CC:Other" studies yielded a total of 166 octant tests of which 53% (88/166) were in the direction predicted by Fiedler while 47% (78/166) were opposite to Fiedler's predicted direction.

A second phase of the analysis examined not only the direction but also the magnitude of the within-octant median correlations computed from the reported correlations for each octant. This analysis permitted a direct comparison

---

[1] The total number of studies is 40 rather than 38 because two CC:Fiedler studies used two different task group types and were therefore counted twice.

not only with Fiedler's predicted values but also among the obtained median values grouped in a variety of ways according to several different variables. The obtained medians for the two comparability categories corresponded one with another only in Octants I, III, and V. The curve yielded by the "CC:Fiedler" median correlations corresponded more closely with Fiedler's predicted curve than did the curve yielded by the "CC:Other" median correlations. A lack of correspondence was also found between or among the obtained median correlations when the octant tests were grouped, for example, by task group type, study setting (real-life or laboratory), organizational setting (e.g., military or schools), or the basis for assessing leadership effectiveness (leader or group performance). A comparison of obtained and predicted values revealed almost complete correspondence for the military studies and almost no correspondence for the school-based studies. There was greater correspondence between the obtained and predicted curves when leadership effectiveness was based on leader rather than group performance.

The results of this median correlation analysis raised serious doubts regarding the comparability of the findings themselves. More specifically, the analysis suggested that neither the comparability categories nor the task group types were sufficiently similar to permit combining all the octant tests. This observation led to a

third assessment of comparability which was based on Exploratory Data Analysis (Tukey, 1977). This analysis showed that the findings yielded by the studies both in the comparability categories and in the task group types were not sufficiently similar to permit combining the partitioned data sets and thus confirmed the picture suggested by the previous analysis.

In light of this outcome an attempt was made to determine whether the methodological variability found in the primary studies was associated with the lack of consistent findings. The effects of methodological variability were assessed by reference to the frequency of support as indicated by the authors' stated conclusions and the directional congruency between the predicted medians and the reported correlations. Because the results generated by crosstabulating each of these criteria with the variations in method did not reveal a clear pattern of relationships, no satisfactory explanation was provided for the observed lack of correspondence. The analysis did, however, reveal some important findings. First, for example, there was extensive overlap in the scores used to designate subjects as high or low LPC and in the scores used to designate groups as having good or poor leader-member relations. Second, Fiedler's procedures for assessing leader-member relations were followed more closely than were his procedures for assessing either task structure or position

power. Overall, there was a tendency for supportive results to be associated with adherence to Fiedler's method and, conversely, for non-supportive results to be associated with deviation from it.

DISCUSSION

Three aspects of these findings merit some extended discussion, namely: (1) the assessment of a leader's effectiveness, (2) the organizational settings in which the studies were conducted, (3) the extent to which the Model has been tested and deemed to be supported. Each of these aspects is discussed in a separate section in the following pages.

The Assessment of a Leader's Effectiveness

It was noted in Chapter 7 (p. 159) that about one third of the octant tests were based on an evaluation of the leader's performance. The Model, however, specifies that leadership effectiveness is synonymous with the effectiveness of the group's performance. That is to say, a leader is effective to the extent that his or her group is effective, by definition. Moreover, and consistent with Rice (1975), the median correlations based upon LPC-leader performance corresponded more closely with Fiedler's predicted values than did those based on LPC-group

performance (see Figure 7.4, p. 162).

The studies reviewed for this meta-analysis revealed considerable divergence with respect to whose task was being evaluated as the measure of effectiveness and whose task was being assessed for structure. For example, in none of the laboratory studies was it clear whether the assessment of effectiveness was based solely on the performance of the group members or on the performance of both the leader and the group members. The real-life studies revealed four different combinations of task structure and leadership effectiveness. Those studies which were consistent with the Model assessed the structure of the group's task and measured effectiveness on the basis of the group's performance on that task. Other studies assessed the structure of the leader's task and evaluated the leader's performance. Some studies were not only inconsistent with the Model, they were also inconsistent within themselves. For example, some primary authors assessed the structure of the group's task but evaluated the leader's performance while others assessed the structure of the leader's task but evaluated the group's performance. It should be noted that studies using these last two combinations were not included in the data base.

Since the specifications of the Model are clear with respect to whose task is to be assessed for structure and whose performance is to be evaluated, the reasons for these

inconsistencies cannot lie within the Model itself. They may lie, however, with the possibility that, for some particular purpose, a researcher deems it more appropriate to examine the structure of the leader's task and, therefore, evaluate the leader's performance. Whatever the rationale may be for this particular methodological departure, it is reasonable to conclude that in spite of the lack of consistency with the Model, studies basing effectiveness on leader rather than group performance yield results which correspond more closely to those predicted by the Model.

The question which remains unanswered, however, is why the leader performance studies are more supportive of the Model. If, for example, these primary authors have based the assessment of both leader-member relations and position power on leader perception, it might be plausible to suggest that the relationship between leadership style and leader performance would be contingent upon the leader's perception of the situation. If this were to be the case, it could then be argued defensibly that one of the present problems with the Contingency Model arises from the combinations of leader and non-leader assessment of the variables, namely leader assessment of leader-member relations, non-leader assessment of task structure and position power, and non-leader performance.

The Organizational Setting

The second aspect to be discussed also arises from the median correlation analysis reported in Chapter 7. The graphic displays (Figure 7.6, p. 169) which compared the obtained values with those predicted by Fiedler showed that (1) studies conducted in military settings corresponded most closely with the antecedent curve, (2) those conducted in business or industrial settings showed some correspondence, and (3) those conducted in school settings bore little resemblance to Fiedler's curve. The discussion focusses on the nature of schools as task groups and the measurement of the task structure variable in an attempt to account for the disparity between the obtained results and those predicted by the Model.

The nature of schools as task groups. It will be recalled from the description of the Contingency Model in Chapter 1 (pp. 15-16) that Fiedler has identified two kinds of task groups: (1) those in which the task requires interdependent action by the group members (interacting groups) and (2) those in which the group members work relatively independently (coacting groups). Fiedler (1971d) has suggested further that there may be a distinction between coacting task groups and coacting training groups.

According to Fiedler (1971d), only coacting task groups yield results similar to the predictions for interacting groups. If both types of task groups are predicted to yield results consistent with the Model, then it should be expected that schools, classified by Fiedler (1971d) as coacting task groups, should also yield results consistent with the predictions of the Model. Clearly this was not the case for the studies included in this meta-analysis. If schools are considered as coacting training groups, it might be possible to account for the obtained disparity. This explanation of disparity seems unlikely, however, since only the pupils in a school could reasonably be considered as "training" groups (groups which exist for the benefit of the individual as distinct from the organization). A more plausible explanation arises from the measurement of the task structure variable.

The treatment of task structure. Studies which assessed the task of teachers for structure typically designated that task as unstructured (e.g., Garland, 1973). There is, however, some evidence in the teaching effectiveness literature (e.g., Bennett, 1976; Rutter et al., 1979) to indicate that teaching may sometimes be a structured task. For a structured task, only Octants I, II, V, and VI are relevant. For an unstructured task, only Octants III, IV, VII, and VIII are relevant. In other

words, it is possible that, because of the designation as an unstructured task, the groups of teachers were assigned to the "wrong" octants. If this were the case, it would help to account for the disparity in the results.

Alternatively, it may be that teaching is both a structured and an unstructured task depending upon the nature not only of what is being taught but also the pupils to whom it is being taught. If this is the case, then there is no way of designating the task of teachers as either structured or unstructured. Given this viewpoint, the task structure variable is inappropriate to coacting groups — at least in schools. It is therefore possible that the way in which the task structure variable has been measured in school-based studies may help to account for the disparity between Fiedler's predictions and the obtained results in school settings.

The Extent of Model Testing

The third aspect to be discussed concerns the statement, frequently encountered in the Fiedler-based literature, that the Contingency Model has been extensively tested and generally supported. For example, Schriesheim and Kerr note that the Model has been widely tested and "these tests have taken place in a wide variety of organizational settings" (1973:13). Fiedler states that the

> Contingency Model has been extensively tested in a wide variety of laboratory and field conditions. Most of these studies have supported the Model (1978:67).

The results yielded by the present study suggest that neither of the above quoted statements is wholly accurate. While the Model has indeed been tested in a large number of different kinds of organizations, it has by no means been intensively tested in any of them except the military. Moreover, and again with exception of the military, the methods used in the testing have not been consistent either within a particular type of organization or across various different types of organizations. Further, Fiedler's statement that most of the studies have supported the Model is without sound grounds. In the absence of established testing criteria beyond correlational direction, the term "support" lacks clear meaning.

Nor is the clarity improved by the existence of the conflicting results yielded by the two indicators of support used in the present study. When the primary authors' stated conclusions were used as the criterion, 78% (31/40) found some measure of support but only 15% (6/40) declared full support. When the directional congruency of 249 correlations was used as the criterion, only 54% (135/249) were in the direction predicted by Fiedler. When magnitude was added to direction for the median correlation analysis, there was even less support for the predictions of the

Model. However, before this conflicting evidence is interpreted as meaning that the Model should — in keeping with the opinion of Schriesheim and Kerr (1977) — be retired from the leadership arena, it should be recognized that the conflicting evidence may not be a function of the Model itself but rather of the substantial amount of inconsistency in its testing. Whether or not the Contingency Model is a viable tool for leadership research cannot be stated unequivocally until such time as testing procedures are both established and, perhaps more importantly, adhered to. It is the purpose of the final section of this chapter to suggest some ways in which the testing procedures should be standardized before the Model can be declared valid or invalid.

FUTURE DIRECTIONS FOR CONTINGENCY MODEL RESEARCH

One of the most important findings to emerge from the present study was the lack of comparability created by the wide variety of methods used not only across the primary studies but also between the primary studies and the Model. In light of this outcome, and of the specific findings which underlie it, several recommendations to guide future research with the Contingency Model are in order.

The first recommendation concerns the score values used to designate subjects as "high LPC" or "low LPC" and

groups as having "good" or "poor" leader-member relations. As long as each variable of the Model continues to be thought of in dichotomous terms, there is a clear need to update the normative values established by Posthuma (1970). Once these values have been revised, they can be used as the cut-off points for both LPC and GAS scores, thus eliminating the variability in score division methods. The use of normative values to dichotomize both LPC and leader-member relations was shown by the present study to be associated with results corresponding more closely to those predicted by the Model[2].

It should be noted that the data coded for this meta-analysis included the LPC and GAS scores not only from the 38 studies used in the data base but also from the 74 primary studies which were deleted from the analysis. The availability of this information should make the task of updating the normative values less difficult. The main problem in doing so is to ensure that all scores included in the update represent leader assessment of leader-member relations. It is therefore recommended that:

---

[2] Normative scores for LPC (high, middle, and low) are provided in the training manual, Improving Leadership Effectiveness (Fiedler et al., 1976). The values do not correspond exactly with the present normative mean of 3.71 provided by Posthuma (1970).

1. Normative scores, based on an update of Posthuma's (1970) values for LPC and GAS, be consistently used as the cut-off points to dichotomize the LPC scores as high or low and leader-member relations as good or poor.

The second recommendation concerns the task structure and position power variables. The Model specifically requires that both be assessed by either the leader's supervisors or knowledgeable judges. The findings of the present study indicated that this procedure was not consistently followed. It is speculated that the appraisal of either variable may yield different opinions when assessed by either the leader or the group members or both the leader and the group members. If this indeed were the case, it would have important implications for the assignment of groups to octants. Clearly, such inconsistency would yield quite disparate results for any given octant being tested. Therefore, if the research is to be regarded as a test of the Model, the assessment must be done in accordance with its specifications. It is therefore recommended that:

2. Task structure and position power be consistently assessed in accordance with Fiedler's method of using either the leader's superiors or knowledgeable judges.

The third recommendation concerns the consistency between whose task is assessed for structure and whose task performance is evaluated as the measure of effectiveness.

In all too many studies, the group's task was assessed for structure but the performance evaluated was that of the leader. This procedure clearly violates the Model which stipulates that the leader is effective to the extent the group performs its primary task. It is therefore recommended that:

3. Group performance, clearly distinguished from that of the leader in both real-life and laboratory settings, be consistently used as the basis for assessing leadership effectiveness.

Researchers who disagree either with this method of determining the effectiveness of the leadership or with the method of assessing the situational variables should not report their findings as a test of the Contingency Model but rather as a modification of the Model.

The fourth recommendation concerns the criteria for determining that a test of one or more octants of the Model does, in fact, support Fiedler's predictions. It has been observed by a number of writers (e.g., Graen et al., 1970; Mitchell et al., 1970) that acceptance of those results which are directionally congruent with Fiedler's predictions constitutes too lenient a criterion by which to assess the validity of the Model. In any one study, it is seldom the case that sample sizes are large enough to yield statistical significance. For example, in a study by Radtke (1978), one of the reported LPC-group performance correlations was -0.92 for Octant I but still nonsignificant (n=3 groups, $p<0.12$). The problems of sample size and statistical significance can

be resolved by aggregating the correlations (as was the original intent of the present study), converting them to Fisher Z scores and then testing for mean differences using the t statistic. However, in the absence of comparability, the results of such testing would be meaningless. Until such time as a sufficient number of comparable tests exist in any one octant, an alternate solution is proposed. It seems reasonable to suggest that some benchmark value above and below the predicted medians be used as a criterion to assess the correspondence in magnitude of the directionally congruent correlations. While it is recognized that this procedure is also lenient and cannot be tested statistically, it does at least take into account both magnitude and direction as criteria for the acceptance or non-acceptance of any given octant test. It is therefore recommended that:

4. A temporary testing criterion, based on a benchmark value above and below Fiedler's predicted medians, be established as a means of assessing whether a given test has supported the octant or a given study has supported the Model.

The final recommendation concerns the reasons for the disparate results yielded by the military, business-industrial, and school studies relative to each other and to the Model. The data coded for the present study make it possible to examine, through further crosstabulation procedures, the combinations of methods used to assess the three situational variables, leadership

effectiveness, and the score division methods used within each of the above organizational settings. The purpose of the comparisons is to search for methodological distinctions which may help to account for the close correspondence of the military and business-industrial results and the lack of correspondence of the school studies. Should it be found that there are few substantial differences in the methods within each subsample, consideration would have to be given to the possibility that organizational differences exist which make the Model more appropriate for some populations than for others. It is therefore recommended that:

5. Further analysis of the available data be conducted to determine the combinations of methods used in the assessment of both the situational variables and leadership effectiveness in an attempt to account for the correspondence yielded by some subsamples of the studies and the lack of correspondence yielded by other subsamples of the studies.

Neither this source of variability nor any of those discussed previously can be stated with any certainty to account for the lack of correspondence between the obtained results and those predicted by Fiedler's Contingency Model. It is only when standardized procedures such as those recommended above have been established and consistently applied that any attempt to develop an explanatory framework should be undertaken.

BIBLIOGRAPHY

Adair, J.G.
1978 The combined probabilities of 345 studies: only half the story? The Behavioral and Brain Sciences, 3:386-87.

Allen, D.B.
1972 Heterogeneity of Research Interests and Effectiveness of University Departments. Technical Report 73-38 (USOE, OEG 0-70-3347), Organizational Research, Department of Psychology, University of Washington, Seattle.

Alpander, G.G.
1974 Planning management training programs for organizational development. Personnel Journal, 53(1):15-25.

Anderson, L.R.
1964 Some Effects of Leadership Training on Intercultural Discussion Groups. Technical Report 18 (ONR 1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Anderson, L.S.
1980 "Billy Budd", "The Caine Mutiny" and Fiedler's "Contingency Model": An Inferential Analysis of Leadership. Unpublished Doctoral dissertation, University of North Carolina.

Arbuthnot, J.
1968 Relationships Among Psychological Differentiations and Leadership Style. Unpublished Master's thesis, Cornell University.

Argyris, C.
1976 Theories of action that inhibit individual learning. American Psychologist, 31(9):638-654.

Arnett, M.D.
1978a Attitudinal correlates of least preferred co-worker score. Psychological Reports, 43(3, Pt. 1):962.

Arnett, M.D.
1978b Behavioral correlates of least preferred co-worker scale. Psychological Reports, 43(3, Pt. 2): 1102.

Arnett, M.D.
1978c Least preferred co-worker score as a measure of cognitive complexity. Perceptual and Motor Skills, 47(2):567-574.

Ashour, A.S.
1973a The Contingency Model of Leadership Effectiveness: an evaluation. Organizational Behavior and Human Performance, 9(3):339-355.

Ashour, A.S.
1973b Further discussion of Fiedler's Contingency Model of Leadership Effectiveness. Organizational Behavior and Human Performance, 9(3):369-376.

Ayer, J.G.
1968 Effects of Success and Failure of Interpersonal and Task Performance upon Leader Perception and Behavior. Technical Report 26 (MD 2060), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Bagley, M.C.
1972 Situational Leadership in Graduate Departments of Physical Education. Unpublished Doctoral dissertation, University of Illinois.

Barrow, J.C.
1977 The variables of leadership: a review and conceptual framework. Academy of Management Review, 2(2):231-246.

Bates, P.A.
1967 Leadership Performance at Two Managerial Levels in the Telephone Company. Unpublished Bachelor's thesis, University of Illinois.

Beach, L.R. and B. Beach
1978 A note on judgments of situational favorableness and probability of success. Organizational Behavior and Human Performance, 22(1):69-74.

Beach, B.H., T.R. Mitchell, and L.R. Beach
1975 Components of Situational Favorableness and Probability of Success. Technical Report 75-66, Organizational Research, Department of Psychology, University of Washington, Seattle.

Beebe, R.J.
1974 The Least Preferred Coworker Score of the Leader and the Productivity of Small Interacting Task Groups in Octants II and IV of the Fiedler Contingency Model. Unpublished Doctoral dissertation, The College of William and Mary, Virginia.

Bell, N.H.
1970 Leadership Effectiveness Within the North Carolina Department of Community Colleges: An Application and Extension of F.E. Fiedler's Contingency Model to Co-Acting Groups. Unpublished Doctoral dissertation, North Carolina State University at Raleigh.

Bennett, M.
1977 Testing management theories cross-culturally. Journal of Applied Psychology, 62(5):578-581.

Bennett, N.
1976 Teaching Style and Pupil Progress. London: Open Books.

Bird, A.M.
1977 Team structure and success as related to cohesiveness and leadership. Journal of Social Psychology, 103(2):217-223.

Blades, J.W. and F.E. Fiedler
1973 Participative Management, Member Intelligence, and Group Performance. Technical Report 73-40, Organizational Research, Department of Psychology, University of Washington, Seattle.

Blades, J.W. and F.E. Fiedler
1976 The Influence of Intelligence, Task Ability, and Motivation on Group Performance. Technical Report 76-78 (ONR 170-761, ARI-G0005), Organizational Research, Department of Psychology, University of Washington, Seattle.

Blanchard, L.
1978 The Leadership Effectiveness of Wisconsin Elementary School Principals. Unpublished Doctoral dissertation, University of Wisconsin.

Bons, P.M.
1974 The Effect of Changes in Leadership Environment on the Behavior of Relationship and Task-Motivated Leaders. Unpublished Doctoral dissertation, University of Washington.

Bons, P.M., A.R. Bass, and S.S. Komorita
1970 Changes in leadership style as a function of military experience and type of command. Personnel Psychology, 23(4):551-568.

Bons, P.M. and F.E. Fiedler
1976 Changes in organizational leadership and the behavior of relationship- and task-motivated leaders. Administrative Science Quarterly, 21(3):433-472.

Borden, D.F.
1976 Leadership Experience and Organizational Performance: A Review of the Literature. Technical Report 76-85 (ARI: DAHC-19-73-G-0005), Organizational Research, Department of Psychology, University of Washington, Seattle.

Borg, W.R. and M.D. Gall
1983 Educational Research: An Introduction (4th ed.). New York: Longmans.

Braxton, R.B. and W.L. Crosby
1974 A Comparative Analysis of Results Using Three Leadership Style Measurement Instruments. Unpublished Master's thesis, Air Force Institute of Technology, Air University (AD 787200).

Brodbeck, M. (ed.)
1968 Readings in the Philosophy of the Social Sciences. London: Collier-MacMillan.

Brown, D.I. and R.A. Smolinski
1974 The Contingency Model of Leadership Effectiveness and its Implications for Military Managers. Unpublished Master's thesis, Air Force Institute of Technology, Air University (AD 787 184).

Browne, T.J.
1974 Empirical Verification of Fiedler's LPC. Unpublished Master's thesis, University of Victoria, B.C., Canada.

Bugental, R.W.G.
1969 A Reinterpretation of Fiedler's ASo Score for Assumed Similarity of Opposites. Unpublished Doctoral dissertation, University of California, Los Angeles.

Butterfield, D.A.
1968 An Integrative Approach to the Study of Leadership Effectiveness in Organizations. Unpublished Doctoral dissertation, University of Michigan.

Carlson, J.
1977 The Effect of Leadership Style and Leader Position Power on Group Performance. Unpublished Doctoral dissertation, University of Maryland.

Chapman, J.B.
1974 A Comparative Analysis of Male and Female Leadership Styles in Similar Work Environments. Unpublished Doctoral dissertation, University of Nebraska, Lincoln.

Chemers, M.M.
1969 Cross-cultural training as a means for improving situational favourableness. Human Relations, 22(6):531-546.

Chemers, M.M.
1970 The relationship between birth order and leadership style. Journal of Social Psychology, 80:243-244.

Chemers, M.M., R.W. Rice, E. Sundstrom, and W.M. Butler
1974 Leader LPC, Training, and Effectiveness: An Experimental Examination. Technical Report No. 73-42 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Chemers, M.M., and G.J. Skrzypek
1972 The Contingency Model of Leadership Effectiveness. Journal of Personality and Social Psychology, 24(2):172-177.

Chemers, M.M. and D.A. Summers
1968 Group Atmoshpere and the Perception of Group Favorableness. Technical Reports 71 (ONR-1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Cohen, P.A.
1981 Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. Review of Educational Research, 51(3):281-309.

Cohen, P.A, J.A. Kulik, and C.C. Kulik
1982 Educational outcomes of tutoring: a meta-analysis of findings. American Educational Research Journal, 19(2):237-248.

Cook, T.D., and L.C. Leviton
1980 Reviewing the literature: a comparison of traditional methods with meta-analysis. Journal of Personality, 48(4):449-472.

Cooper, H.M.
1982 Scientific quidelines for conducting integrative research reviews. Review of Educational Research, 52(2):291-302.

Csoka, L.S.
1972a Intelligence: A Critical Variable in Leadership Experience. Technical Report 72-34 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Csoka, L.S.
1972b A Validation of the Contingency Model Approach to Leadership Experience and Training. Technical Report 72-32 (ONR-177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Csoka, L.S.
1973 Organic and Mechanistic Organizational Climates and the Contingency Model. Technical Report No. 73-41 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Csoka, L.S.
1975 Relationship between organizational climate and the situational favorableness dimension of Fiedler's Contingency Model. Journal of Applied Psychology, 60:273-277.

Csoka, L.S., and F.E. Fiedler
1971 The Effect of Leadership Experience and Training in Structured Military Tasks — A Test of the Contingency Model. Technical Report No. 71-21 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Csoka, L.S. and F.E. Fiedler
1972 The effect of military leadership training: a test of the Contingency Model. Organizational Behavior and Human Performance, 8(3):395-407.

Danielson, R.R.
1974 Leadership in Coaching: Description and Evaluation. Unpublished Doctoral dissertation, University of Alberta.

Deveau, R.J.
1976 The Relationship Between the Leadership Effectiveness of First-line Supervisors and Measures of Authoritarianism, Creativity, General Intelligence, and Leadership Style. Unpublished Doctoral dissertation, Boston University School of Education.

DeZonia, R.H.
1958 The Relationship Between Psychological Distance and Effective Task Performance. Unpublished Doctoral dissertation, University of Illinois.

Doyle, W.J. and W.P. Ahlbrand
1973 Hierarchical group performance and leader orientation. Administrator's Notebook, 22(3).

Dubin, R.
1965 Supervision and productivity: empirical findings and theoretical considerations, in L. Broom (ed.), Leadership and Productivity: Some Facts of Industrial Life. San Francisco: Chandler.

Duffy, J.F.
1973 A Within-Subjects Validity Generalization of Fiedler's Contingency Model of Leadership. Unpublished Doctoral dissertation, Iowa State University.

Eagly, A.H.
1970 Leadership style and role differentiation as determinants of group effectiveness. Journal of Personality, 38(4):509-524.

Ebersole, P.D.
1969 Investigation of Fiedler's Leadership Effectiveness Model in a Quasi-Therapeutic Situation with Hospitalized Patients. Unpublished Doctoral dissertation, Univeristy of California, Los Angeles.

Evans, M.G.
1973 A leader's ability to differentiate the subordinate's perception of the leader, and the subordinate's performance. Personnel Psychology, 26:385-395.

Evans, M.G. and J. Dermer
1974 What does the least preferred co-worker scale really measure? A cognitive interpretation. Journal of Applied Psychology, 59:202-206.

Fahy, F.A.
1972 The Contingency Model and the Schools: A Study of the Relationships Between Teaching Effectiveness, Leadership Style, and the Favourableness of the Situation. Unpublished Doctoral dissertation, Rutgers University.

Faulkner, R.T.B.
1980 Leadership Style of Principal Related to Degree of Bureaucratization of School. Unpublished Doctoral dissertation, University of Ottawa.

Faust, R.M.
1972 Leader Styles of Union Stewards: A Test of Fiedler's Model. Unpublished Doctoral dissertation, University of Iowa.

Fearon, D.S.
1973 An Investigation of Fiedler's Contingency Theory of Leadership Effectiveness as It Applies to Community College Citizens Advisory Councils: An Exploratory Study. Unpublished Doctoral dissertation, University of Connecticut.

Fiedler, F.E.
1958 Leader Attitudes and Group Effectiveness. Final Report of ONR Project NR 170-106, N6-ori-07135, University of Illinois Press, Urbana.

Fiedler, F.E.
1964 A Contingency Model of Leadership Effectiveness, in L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol. 1:149-190. New York: Academic Press.

Fiedler, F.E.
1967 A Theory of Leadership Effectiveness. New York: McGraw-Hill.

Fiedler, F.E.
1968 Leadership Experience and Leader Performance — Another Hypothesis Shot to Hell. Technical Report No. 70 (ONR 1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Fiedler, F.E.
1969 Style or circumstance: the leadership enigma. Psychology Today 2(10):38-43.

Fiedler, F.E.
1970a On the Death and Transfiguration of Leadership Training. Technical Report No. 70-16(ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1970b Personality, Motivational Systems, and Behavior of High and Low LPC Persons. Technical Report No. 70-12 (ONR 1834-36), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1971a Leadership. New York: General Learning Press.

Fiedler, F.E.
1971b Note on the methodology of the Graen, Orris, and Alvares studies testing the Contingency Model. Journal of Applied Psychology, 55(3):202-204.

Fiedler, F.E.
1971c Personality and Situational Determinants of Leader Behavior. Technical Report No. 71-18 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1971d Validation and extension of the Contingency Model of Leadership Effectiveness: a review of empirical findings. Psychological Bulletin, 76(2):128-148.

Fiedler, F.E.
1972a How do you make leaders more effective? New answers to an old puzzle. Organizational Dynamics, 1(2):3-18.

Fiedler, F.E.
1972b Leadership Experience and Leadership Training — Some New Answers to an Old Problem. Technical Report No. 72-36 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1972c Personality, motivational systems, and behavior of high and low LPC persons. Human Relations, 25(5):391-412.

Fiedler, F.E.
1973a The Contingency Model — a reply to Ashour. Organizational Behavior and Human Performance, 9(3):356-368.

Fiedler, F.E.
1973b Predicting the effects of leadership training and experience from the Contingency Model: a clarification. Journal of Applied Psychology, 57(2):110-113.

Fiedler, F.E.
1973c Stimulus/response: the trouble with leadership training is that it doesn't train leaders. Psychology Today, 6(9):23-30, 92.

Fiedler, F.E.
1973d Toward a Comprehensive System of Leadership Utilization. Technical Report No. 73-50 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1974 The Contingency Model — new directions for leadership utilization. Contemporary Business, Autumn, 65-79.

Fiedler, F.E.
1976a Intelligence and Group Performance: A Multiple Screen Model. Technical Report No. 76-79 (ARI-G-0005) Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1976b The leadership game: matching the man to the situation. Organizational Dynamics, 4(3):6-16.

Fiedler, F.E.
1977 A rejoinder to Schriesheim and Kerr's premature obituary of the Contingency Model, in J.G. Hunt and L.L. Larson (eds.), Leadership: The Cutting Edge. Carbondale: Southern Illinois University Press.

Fiedler, F.E.
1978 The Contingency Model and the dynamics of the leadership process, in L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol. 11. New York: Academic Press.

Fiedler, F.E.
1979 Responses to Sergiovanni. Educational Leadership, 36(9): 394-396.

Fiedler, F.E. and M.M. Chemers
1974 Leadership and Effective Management. Glenview: Scott, Foresman.

Fiedler, F.E., M.M. Chemers, and L. Mahar
1976 Improving Leadership Effectiveness: The Leader Match Concept (rev. ed.). New York: John Wiley.

Fiedler, F.E. and A.F. Leister
1977 Leader intelligence and task performance: a test of a multiple screen model. Organizational Behavior and Human Performance, 20(1):1-14.

Fiedler, F.E., G. O'Brien, and D. Ilgen
1969 The effect of leadership style upon the performance and adjustment of volunteer teams operating in a stressful foreign environment. Human Relations, 22(6):503-514.

Fiedler, F.E., E.H. Potter, M.M. Zais, and W.A. Knowlton
1979 Organizational stress, and the use and misuse of managerial intelligence and experience. Journal of Applied Psychology, 64(6):635-647.

Fishbein, M.
1966 Unpublished manuscript (cited in Fiedler, 1967:29).

Fishbein, M., E. Landy, and G. Hatch
1969a A consideration of two assumptions underlying Fiedler's Contingency Model for Leadership Effectiveness. American Journal of Psychology, 82(4):457-473.

Fishbein, M., E. Landy, and G. Hatch
1969b Some determinants of an individual's esteem for the least preferred co-worker: an attitudinal analysis. Human Relations, 22(2):173-188.

Flynn, M.L.
1972 Relationships Between Special Education Administrators' Responses to Fiedler's Least Preferred Coworker Measure and Their Responses to Selected Hypothetical Group Task Situations. Unpublished Doctoral dissertation, University of Georgia.

Foa, U.G., T.R. Mitchell and F.E. Fiedler
1971 Differentiation matching. Behavioral Science, 16(2):130-142.

Fox, W.M.
1976 Reliabilities, means, and standard deviations for LPC scales: instrument refinement. Academy of Management Journal, 19(3):450-461.

Fox, W.M.
1982a A test of Octant I of Fiedler's Contingency Model with training dependent coacting task groups. Replications in Social Psychology, 2(1):47-49.

Fox W.M.
1982b Behavioral correlates of leader LPC score as moderated by situational favorability. Replications in Social Psychology, in press.

Fox, W.M., W.A. Hill and W.H. Guertin
1971 Factor Analysis of Least Preferred Coworker Scales. Technical Report 70-1, College of Business, University of Florida, Gainesville.

French, J.R. and B.H. Raven
1958 Legitimate power, coercive power, and observability in social influence. Sociometry, 21:83-97. (cited in Fiedler, 1967:22 as "1956")

Gale, G.
1979 Theory of Science: An Introduction to the History, Logic, and Philosophy of Science. New York: McGraw-Hill.

Garland, P.
1973 The Effect of Principal-Teacher Interaction in Secondary School Environments: An Empirical Study. Unpublished Doctoral dissertation, University of Ottawa.

Geyer, L.J. and J.W. Julian
1973 Manipulation of situational favorability in tests of the Contingency Model. Journal of Psychology, 84(1):13-21.

Gilbert, M.A.
1977 An Organizational Approach to the Study of Productivity, Efficiency, and Satisfaction of AAA High-School Basketball Teams Based on Fiedler's Contingency Model and Taylor and Bowers' Survey of Organizational Conditions. Unpublished Doctoral dissertation, University of Oregon.

Glass, G.V.
1976 Primary, Secondary, and Meta-analysis of Research. Paper presented to the American Educational Research Association, San Francisco.

Glass, G.V.
1978 Integrating findings: the meta-analysis of research, in L.S. Shulman (ed.), Review of Research in Education, Vol. 5. Itasca: Peacock.

Glass, G.V., B. McGaw, and M.L. Smith
1981 Meta-Analysis in Social Research. Beverly Hills: Sage.

Glass, G.V. and M.L. Smith
1979 Meta-analysis of research on class size and achievement. Educational Evaluation and Policy Analysis, 1(1):2-16.

Gochman, I. and F.E. Fiedler
1975 The Effect of Situational Favorableness on Leader and Member Perceptions of Leader Behavior. Technical Report No. 75-64 (ONR N00014-67-A-0103-0013), Organizational Research, Department of Psychology, University of Washington, Seattle.

Goldstein, G.S. and C.W. Cole
1969 Leadership variables. The Journal of the Scientific Laboratories, Denison University, 50:59-67.

Gottheil, E. and C.G. Lauterbach
1969 Leader and squad attributes contributing to mutual esteem among squad members. Journal of Social Psychology, 77:69-78.

Graen, G.B., D. Alvares, J.B. Orris, and J.A. Martella
1970 The Contingency Model of Leadership Effectiveness: antecedent and evidential results. Psychological Bulletin, 74(4):285-296.

Graen, G.B., J.B. Orris, and K.M. Alvares
1971a Contingency Model of Leadership Effectiveness: some experimental results. Journal of Applied Psychology, 55(3):196-201.

Graen, G.B., J.B. Orris, and K.M. Alvares
1971b Contingency Model of Leadership Effectiveness: some methodological issues. Journal of Applied Psychology, 55(3):205-210.

Graham, W.K.
1968 Description of leader behavior and evaluation of leaders as a function of LPC. Personnel Psychology, 21:457-464.

Graham, W.K.
1973 Leader behavior, esteem for the least preferred coworker, and group performance. Journal of Social Psychology, 90:59-66.

Green, S.G. and D.M. Nebeker
1974 Leader Behavior: Autonomous or Interactive? Technical Report 74-62 (ONR 177-473), Organizational Research, Department of Psychology, University of Washington, Seattle.

Green, S.G. and D.M. Nebeker
1977 The effects of situational factors and leadership style on leader behavior. Organizational Behavior and Human Performance, 19(2):368-377.

Green, S., D. Nebeker and A. Boni
1976 Personality and situational effects on leader behavior. Academy of Management Journal, 19(2):184-194.

Gruenfeld, L. and J. Arbuthnot
1968 Field independence, achievement values, and the evaluation of a competency related dimension on the least preferred co-worker (LPC) measure. Perceptual and Motor Skills, 27:991-1002.

Gruenfeld, L.W., D.E. Rance and P. Weissenburg
1969 The behavior of task-oriented (low LPC) and socially-oriented (high LPC) leaders under several conditions of social support. Journal of Social Psychology, 79:99-107.

Grunstad, N.L.
1972 The Determination of Leader Behavior From the Interaction of Fiedler's LPC and Situational Favorableness. Unpublished Doctoral dissertation, Ohio State University.

Hansson, R.O. and F.E. Fiedler
1972 Perceived Similarity, Personality, and Attraction to Large Organizations. Technical Report 72-39 (ONR-177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Hardy, R.C.
1969 The Effect of Leadership Style on the Performance of Small Classroom Groups: A Test of the Contingency Model. Unpublished Doctoral dissertation, Indiana University.

Hardy, R.C.
1972 A developmental study of relationships between birth order and leadership style for two distinctly different American groups. Journal of Social Psychology, 87:147-148.

Hardy, R.C.
1975 A test of poor leader-member relations cells of the Contingency Model on elementary school children. Child Development, 46(4):958-964.

Hardy, R.C. and J.F. Bohren
1975 The effect of experience on teacher effectiveness: A test of the Contingency Model. Journal of Psychology, 89(1):159-163.

Hardy, R.C., J. Carey, B. Eberwein, and J. Eliot
1976 Comparison of task- and relation-oriented individuals' ability to perceive viewpoints of others. Perceptual and Motor Skills, 42(3, Pt. 2):1028-1030.

Hardy, R.C., S. Sack, and F. Harpine
1973 An experimental test of the Contingency Model on small classroom groups. Journal of Psychology, 85(1):3-16.

Harré, R.
1972 The Philosophies of Science. London: Oxford University Press.

Hartley, S.S.
1979 Results of Meta-Analysis of the Effects on Mathematics Achievement of Different Instructional Modes. Paper presented to the American Educational Research Association, San Francisco.

Hawkins, C.A.
1962 A Study of Factors Mediating a Relationship Between Leader Rating Behavior and Group Productivity. Unpublished Doctoral dissertation, University of Minnesota, Minneapolis.

Hawley, D.E.
1969 A Study of the Relationship Between the Leader Behavior and Attitudes of Elementary School Principals. Unpublished Master's thesis, University of Saskatchewan.

Hearold, S.L.
1979 Meta-Analysis of the Effects of Television on Social Behavior. Paper presented to the American Educational Research Association, San Francisco.

Hedges, L.V. and I. Olkin
1982 Analyses, reanalyses, and meta-analysis. Contemporary Education Review, 1(3):157-165.

Hedges, L.V. and W. Stock
1983 The effects of class size: an examination of rival hypotheses. American Educational Research Journal, 20(1):63-85.

Helmich, D.L.
1975a Corporate succession: an examination. Academy of Management Journal, 18(3):429-441.

Helmich, D.L.
1975b The executive interface and president's leadership behavior. Journal of Business Research, 3(1):43-52.

Hemphill, J.K.
1961 Why people attempt to lead, in L. Petrullo and B.M. Bass (eds.), Leadership and Interpersonal Behavior. New York: Holt.

Hendrix, W.H.
1976 Contingency Approaches to Leadership: A Review and Synthesis. Technical Report 76-17 (Air Force Human Resources Laboratory), Lackland Air Force Base, Texas (ED 130013).

Hersey, P. and K.H. Blanchard
1972 Management of Organizational Behavior: Utilizing Human Resources (2nd ed.). Englewood Cliffs: Prentice Hall.

Hewett, T.T. and G.E. O'Brien
1971 The Effects of Work Organization, Leadership Style and Member Compatability upon Small Group Productivity. Technical Report No. 71-22 (ONR 177-473), Organizational Research, Department of Psychology, University of Washington, Seattle.

Hill, W.A.
1969a A situational approach to leadership effectiveness. Journal of Applied Psychology, 53(6):513-517.

Hill, W.A.
1969b The LPC leader: a cognitive twist, in W.G. Scott and P.P. Le Breton (eds.), Academy of Management Proceedings (29th Annual Meeting), August:125-130.

Hills, R.J.
1975 Some notes on the methodology of science for researchers and administrators in education, in W.G. Monahan (ed.), Theoretical Dimensions of Educational Administration. New York: MacMillan.

Hopfe, M.E.
1968 A Case Study of Leadership in a College Using Fiedler's Contingency Model. Unpublished Doctoral dissertation, Ohio University.

Hosking, D.M.
1978 A Critical Evaluation of Fiedler's Predictor Measures of Leadership Effectiveness. Unpublished Doctoral dissertation, University of Warwick.

Hosking, D.M. and C.A. Schriesheim
1978 Improving leadership effectiveness: the Leader Match concept. Administrative Science Quarterly, 23(3):496-505.

House, R.J.
1971 A path goal theory of leader effectiveness. Administrative Science Quarterly, 16(3):321-338.

Hovey, D.E.
1974 The low powered leader confronts a messy problem: a test of Fiedler's theory. Academy of Management Journal, 17(2):358-362.

Hunt, J.G.
1966 A Test of Fiedler's Leadership Contingency Model in Three Organizational Settings. Unpublished Doctoral dissertation, University of Illinois, Urbana.

Hunt, J.G.
1971 Leadership style effects at two managerial levels in a simulated organization. Administrative Science Quarterly, 16(4):476-485.

Hunt, J.G., R.N. Osborn and L.L. Larson
1975 Upper level technical orientation and first-level leadership within a noncontingency and contingency framework. Academy of Management Journal, 18(3):476-488.

Hutchins, E.B.
1958 Task-Oriented and Quasi-Therapeutic Role Functions of the Small Group Leader. Unpublished Doctoral dissertation, University of Illinois, Urbana.

Ilgen, D.R. and G. O'Brien
1968 The Effects of Task Organization and Member Compatability on Leader-Member Relations in Small Groups. Technical Report No. 58 (ONR-1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Inciong, P.A.
1974 Leadership Styles and Team Success. Unpublished Doctoral dissertation, University of Utah.

Jackson, G.B.
1980 Methods for integrative reviews. Review of Educational Research, 50(3):438-460.

Jackson, J.J.
1977 Leadership styles and recreation programme delivery systems. Canadian Journal of Applied Sports Sciences, 2(4):207-211.

Jacobs, H.L.
1975 A Critical Evaluation of Fiedler's Contingency Model of Leadership Effectiveness in its Application to Inter-Disciplinary Task-Groups in Public School Settings. Unpublished Doctoral dissertation, Columbia University Teachers College.

Jacoby, J.
1968 Creative ability of task-oriented versus person-oriented leaders. The Journal of Creative Behavior, 2(4):249-253.

Johnson, D.W., R.T. Johnson, and G. Maruyama
1983 Interdependence and interpersonal attraction among heterogeneous and homogeneous individuals: a theoretical formulation and a meta-analysis of the research. Review of Educational Research, 53(1):5-54.

Johnson, R. and B. Ryan
1976 A test of the Contingency Model of Leadership Effectiveness. Journal of Applied Social Psychology, 6(2):177-185.

Jones, J.R. and M. Johnson
1972 LPC as a modifier of leader-follower relationships. Academy of Management Journal, 15(2):185-196.

Justis, R.T., B.L. Kedia, and D.B. Stephens
1978 The effect of position power and perceived task competence on trainer effectiveness: a parital utilization of Fiedler's Contingency Model of Leadership. Personnel Psychology, 31(1): 83-93.

Kaplan, A.
1964 The Conduct of Inquiry: A Methodology for Behavioral Science. San Francisco: Chandler.

Katz, D. and R.L. Kahn
1966 The Social Psychology of Organizations. New York: Wiley.

Keane, F.J.
1976 The Relationship of Sex, Teacher-Leadership Style and Teacher-Leader Behavior in Teacher-Student Interaction. Unpublished Doctoral dissertation, Boston University.

Kennedy, J.K.
1982 Middle LPC leaders and the Contingency Model of Leadership Effectiveness. Organizational Behavior and Human Performance, 30(1):1-14.

Kerlinger, F.N.
1973 Foundations of Behavioral Research (2nd ed.). New York: Holt, Rinehart and Winston.

Kerr, S. and A. Harlan
1973 Predicting the effectiveness of leadership training and experience from the Contingency Model: some remaining problems. Journal of Applied Psychology, 57(2):114-117.

Konar-Goldband, E., R.W. Rice, and W. Monkarsh
1979 Time-phased interrelationships of group atmosphere, group performance, and leader style. Journal of Applied Psychology, 64(4):401-409.

Korman, A.K.
1973 On the development of contingency theories of leadership: some methodological considerations and a possible alternative. Journal of Applied Psychology, 58(3):384-387.

Kuehl, C.R.
1971 Small Group Productivity as Related to Leadership Style, Role Differentiation, Status Differentiation and Organizational Backgrounds of Members. Unpublished Doctoral dissertation, University of Iowa.

Kulik, J.A., R.L. Bangert, and G.W. Williams
1983 Effects of computer-based teaching on secondary school students. Journal of Educational Psychology, in press.

Kulik, C.C. and J.A. Kulik
1982 Effects of ability grouping on secondary school students: a meta-analysis of evaluation findings. American Educational Research Journal, 19(3):415-428.

Kulik, J.A., C.C. Kulik, and P.A. Cohen
1980 Effectiveness of computer-based college teaching: a meta-analysis of findings. Review of Educational Research, 50(4):525-544.

Lanaghan, R.C.
1972 Leadership Effectiveness in Selected Elementary Schools. Unpublished Doctoral dissertation, University of Illinois, Urbana-Champaign.

Larson, L.L. and K. Rowland
1973 Leadership style, stress, and behavior in task performance. Organizational Behavior and Human Performance, 9(3):407-420.

Larson, L.L. and K. Rowland
1974 Leadership style and cognitive complexity. Academy of Management Journal, 17(1):37-45.

Lord, R.G., J.S. Phillips, and M.C. Rush
1980 Effects of sex and personality on perceptions of emergent leadership, influence, and social power. Journal of Applied Psychology, 65(2):176-182.

Luiten W., W. Ames, and G. Ackerson
1980 A meta-analysis of the effects of advance organizers on learning and retention. American Educational Research Journal, 17(2):211-218.

Mai-Dalton, R.
1976 The Influence of Training and Position Power on the Behavior of Male and Female Leaders. Technical Report 76-86, Organizational Research, Department of Psychology, University of Washington, Seattle.

Marascuilo, L.A. and J.R. Levin
1970 Appropriate post hoc comparisons for interaction and nested hypotheses in analysis of variance designs: the elimination of Type IV errors. American Educational Research Journal, 7(3):397-421.

Margeneau, H.
1950 The Nature of Physical Reality. Toronto: McGraw-Hill.

Martin, W.J.
1977 Fiedler's Contingency Model in Classroom Teaching Situations. Unpublished Doctoral dissertation, University of Utah.

Martin, Y.M., G.B. Isherwood, and R.E. Lavery
1976 Leadership effectiveness in teacher probation committees. Educational Administrative Quarterly, 12(2):87-99.

McGaw, B. and K.R. White
1981 Meta-Analysis of Empirical Research (Research Training Presession Manual). Presented to the American Educational Research Association, Los Angeles.

McGrath, J.E. and I. Altman
1966 Small Group Research. New York: Holt, Rinehart, and Winston.

McKague, T.R.
1968 A Study of the Relationship Between School Organizational Behavior and the Variables of Bureaucratization and Leader Attitudes. Unpublished Doctoral dissertation, University of Alberta.

McKague, T.R.
1970 LPC — a new perspective on leadership. Educational Administration Quarterly, 6(3):1-14.

McMahon, J.T.
1972 The contingency theory: logic and method revisited. Personnel Psychology, 25:697-710.

McNamara, V.D.
1967 A Descriptive Analytic Study of Directive-Permissive Variation in the Leader Behavior of Elementary School Principals. Unpublished Master's thesis, University of Alberta.

McNamara, V.D.
1968 Leadership, Staff, and School Effectiveness. Unpublished Doctoral dissertation, University of Alberta.

Mellor, K.P.
1974 An Investigation of Fiedler's Contingency Model of Leadership Effectiveness as It Applies to Elementary Principals in Rhode Island. Unpublished Doctoral dissertation, University of Connecticut.

Meuwese, W.A.T.
1964 The Effect of the Leader's Ability and Interpersonal Attitudes on Group Creativity Under Varying Conditions of Stress. Unpublished Doctoral dissertation, University of Amsterdam, Amsterdam.

Michaelson, L.K.
1973 Leader orientation, leader behavior, group effectiveness and situational favorability: an empirical extension of the Contingency Model. Organizational Behavior and Human Performance, 9(2):226-245.

Millard, R.J.
1981 A Comparative Analysis of Male and Female Management Style and Perceived Behavior Patterns. Unpublished Doctoral dissertation, Brigham Young University.

Miller, T.I.
1979 Comparative Effects of Drug Therapy, Psychotherapy and Drug Plus Psychotherapy: A Meta-Analysis. Paper presented to the American Educational Research Association, San Francisco.

Miskel, C.G.
1974 Public School Principals' Leader Style, Organizational Situation, and Effectiveness. Final Report (NIE, DHEW-NE-G-00-3-0141), University of Kansas, Lawrence.

Miskel, C.G.
1977 Principals' attitudes toward work and co-workers, situational factors, perceived effectiveness, and innovation effort. Educational Administration Quarterly, 13(2):51-70.

Mitchell, T.R.
1969 Leader Complexity, Leadership Style, and Group Performance. Unpublished Doctoral dissertation, University of Illinois.

Mitchell, T.R.
1970 The construct validity of three dimensions of leadership research. Journal of Social Psychology, 80:89-94.

Mitchell, T.R., A. Biglan, G.R. Oncken, and F.E. Fiedler
1970 The Contingency Model: Criticisms and Suggestions. Technical Report No. 70-11 (ONR N00014-67-A-0103-0012), Organizational Research, Department of Psychology, University of Washington, Seattle.

Mitchell, T.R., J.R. Larson and S.G. Green
1977 Leader behavior, situational moderators, and group performance: an attributional analysis. Organizational Behavior and Human Performance, 18(1):254-268.

Monahan, W.G. (ed.)
1975 Theoretical Dimensions of Educational Administration. New York: MacMillan.

Muller, H.P.
1970 Relationship between time-span of discretion, leadership behavior, and Fiedler's LPC scores. Journal of Applied Psychology, 54(2):140-144.

Nayar, E.S.K., H. Touzard and D.A. Summers
1968 Training tasks, and mediator orientation in heterocultural negotiations. Human Relations, 21(3):283-294.

Nealey, S.M. and M.R. Blood
1968 Leadership performance of nursing supervisors at two organizational levels. Journal of Applied Psychology, 52(5):414-422.

Nealey, S.M. and F.E. Fiedler
1968 Leadership functions of middle managers. Psychological Bulletin, 70(5):313-329.

Nealey, S.M. and T.W. Owen
1970 A multitrait-multimethod analysis of predictors and criteria of nursing performance. Organizational Behavior and Human Performance, 5:348-365.

Nebeker, D.M.
1972 A Model for Supervisor Behavior: An Integration of Fiedler's Contingency Model and Expectancy Theory. Unpublished Doctoral dissertation, University of Washington.

Nebeker, D.M.
1973 Situational and Perceived Environmental Uncertainty: An Integrative Approach. Technical Report No. 73-48 (ARPA, ONR-177-473), Organizational Research, Department of Psychology, University of Washington, Seattle.

Nebeker, D.M.
1975 Situational favorability and environmental uncertainty: an integrative study. Administrative Science Quarterly, 20(2):281-294.

Nebeker, D.M., L.R. Beach and S.G. Green
1974 Situational Favorability and the Perception of Uncertainty: An Experimental Demonstration. Technical Report 74-61 (ONR 177-472), Organization Research, Department of Psychology, University of Washington, Seattle.

Nebeker, D.M. and R.O. Hansson
1972 Confidence in Human Nature and Leader Style. Technical Report 72-37 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Ninane, P. and F.E. Fiedler
1970 Member reactions to success and failure of task groups. Human Relations, 23(1):3-13.

O'Brien, G.E., A. Biglan and J. Penna
1972 Measurement of the distribution of potential influence and participation in groups and organizations. Journal of Applied Psychology, 56(1):11-18.

O'Brien, G.E. and F.E. Fiedler
1977 Measurement of interactive effects of leadership style and group structure upon group performance. Australian Journal of Psychology, 29(1):59-71.

O'Brien, G.E., F.E. Fiedler and T. Hewett
1971 The effects of programmed culture training upon the performance of volunteer medical teams in Central America. Human Relations, 24(3):209-231.

Pflaum, S.W., H.J. Walberg, M.L. Karegians, and S.P. Rasher
1980 Reading instruction: a quantitative analysis. Educational Researcher, 9(7):12-18.

Posthuma, A.B.
1970 Normative Data on the Least-Preferred Co-worker Scale (LPC) and the Group Atmosphere Questionnaire (GA). Technical Report 70-8, (ARPA Order 454, N00014-67-A-0103-0013, Office of Naval Research), Organizational Research, Department of Psychology, University of Washington, Seattle.

Potter, E.H. and F.E. Fiedler
1981 Stress and the utilization of staff member intelligence and experience. Academy of Management Journal, 24(2):361-376.

Prothero, J. and F.E. Fiedler
1974 The Effect of Situational Change on Individual Behavior and Performance: An Extension of the Contingency Model. Technical Report 74-59 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Radtke, E.M.T.
1978 Leadership Effectiveness of Hospital-Diploma, Community College, and University Prepared Head Nurses: Interrelationships of Leadership Style and Situational Favorableness on Group Performance in Patient-Units of General Hospitals. Unpublished Doctoral dissertation, Wayne State University.

Ranhosky, F.W.
1978 Personality and Behavioral Correlates of Leadership Effectiveness of Assistant Principals. Unpublished Doctoral dissertation, Fordham University.

Reavis, C.A. and V.J. Derlaga
1976 Test of a Contingency Model of teacher effectiveness. Journal of Educational Research, 69(6):221-225.

Redfield, D.L. and E.W. Rousseau
1981 A meta-analysis of experimental research on teacher questioning behavior. Review of Educational Research, 51(2):237-245.

Reilly, A.J.
1968 The Effects of Different Leadership Styles on Group Performance: A Field Experiment. Iowa Industrial Relations Center, Iowa State University, Ames.

Rice, R.W.
1975 The Esteem For Least Preferred Coworker (LPC) Score: What Does It Measure? Unpublished Doctoral dissertation, University of Utah.

Rice, R.W.
1978 Psychometric properties of the esteem for least preferred co-worker (LPC) scale. Academy of Management Review, 3(1):106-118.

Rice, R.W.
1979 Reliability and validity of the LPC scale: a reply. Academy of Management Review, 4(2):291-294.

Rice, R.W.
1981 Leader LPC and follower satisfaction: a review. Organizational Behavior and Human Performance, 28(1):1-25.

Rice, R.W., L.R. Bender, and A.G. Vitters
1980 Leader sex, follower attitudes toward women, and leadership effectiveness: a laboratory experiment. Organizational Behavior and Human Performance, 25(1):46-78.

Rice, R.W. and M.M. Chemers
1973 Predicting the emergence of leaders using Fiedler's Contingency Model of Leadership Effectiveness. Journal of Applied Psychology, 57:281-287.

Rice, R.W. and M.M. Chemers
1975 Personality and situational determinants of leader behavior. Journal of Applied Psychology, 60(1):20-27.

Rice, R.W. and F.J. Seaman
1981 Interval analysis of the least preferred co-worker (LPC) scale. Educational and Psychological Measurement, 41(1):109-120.

Rice, R.W., F.J. Seaman, and D.J. Garvin
1978 An empirical examination of the esteem for Least Preferred Co-worker (LPC) construct. Journal of Psychology, 98(2):195-205.

Rosenthal, R.
1978 Combining results of independent studies. Psychological Bulletin, 85:185-193.

Rosenthal, R.
1979 The "file-drawer" problem and tolerance for null results. Psychological Bulletin, 86:638-641.

Rosenthal, R. and D.B. Rubin
1978 Interpersonal expectancy effects: the first 345 studies. The Behavioral and Brain Sciences, 3:377-415.

Rubin, G.J.
1971 A Modified Contingency Model for Leadership Effectiveness. Unpublished Doctoral dissertation, University of Tennessee.

Rudner, R.S.
1966 Philosophy of Science. Englewood-Cliffs: Prentice-Hall.

Rutter, M., B. Maughan, P. Mortimore, and J. Ouston
1979 Fifteen Thousand Hours. Cambridge: Harvard University Press.

Saha, S.K.
1972 An Experimental Validation and Extension of Fiedler's Contingency Model of Leadership Effectiveness. Unpublished Doctoral dissertation, University of British Columbia.

Sashkin, M.
1970 Supervisory Leadership in Problem Solving Groups: Experimental Tests of Fred Fiedler's "Theory of Leadership Effectiveness" in the Laboratory Using Role Play Methods. Unpublished Doctoral dissertation, University of Michigan.

Sashkin, M., F.C. Taylor and R.C. Tripathi
1974 An analysis of situational moderating effects on the relationship between LPC and other psychological measures. Journal of Applied Psychology, 59:731-740.

Schmidt, D.E.
1976 Cognitive Homogeneity and the Contingency Model of Leadership Effectiveness: A Preliminary Study. Technical Report No. 76-83 (ARI-G0005), Organizational Research, Department of Psychology, University of Washington, Seattle.,

Schmidt, D.E. and F.E. Fiedler
1976 The Contingency Model and Leader Behavior: An Instrumentality Approach. Technical Report No. 78-62 (ARI-G-0005), Organizational Research, Department of Psychology, University of Washington, Seattle.

Schneier, C.E.
1978 The Contingency Model of Leadership: an extension to emergent leadership and leader's sex. Organizational Behavior and Human Performance, 21:220-239.

Schneier, C.E.
1979 Measuring cognitive complexity: developing reliability, validity, and norm tables for a personality instrument. Educational and Psychological Measurement, 39(3):599-612.

Schriesheim, C.A.
1979 Social desirability and leader effectiveness. Journal of Social Psychology, 108(1):89-94.

Schriesheim, C.A., B.D. Bannister, and W.H. Money
1979 Psychometric properties of the LPC scale: an extension of Rice's review. Academy of Management Review, 4(2):287-290.

Schriesheim, C.A. and S. Kerr
1977 Theories and measures of leadership: a critical appraisal of current and future directions, in J.G. Hunt and L.L. Larson (eds.), Leadership: The Cutting Edge. Carbondale: Southern Illinois University Press.

Sepic, F.T.
1979 The Effects of Leadership Style, Functional Authority and Stress on the Relationships of Supervisory Experience and Managerial Performance in the Hotel and Restaurant Industry. Unpublished Doctoral dissertation, University of Washington, Seattle.

Shaw, M.E.
1963 Scaling Group Tasks: A Method for Dimensional Analysis. Technical Report No. 1 (Office of Naval Research Contract NR 170-266, Nonr-58011), University of Florida.

Shiflett, S.C.
1973 The Contingency Model of Leadership Effectiveness: some implications of its statistical and methodological properties. Behavioral Science, 18:429-440.

Shiflett, S.C.
1974 Stereotyping and esteem for one's least preferred co-worker. Journal of Social Psychology, 93:55-65.

Shiflett, S.C., R.G. Downey, and P.J. Duffy
1979 The Effects of Multidimensionality on the Predictive and Construct Validity of the LPC Scale. U.S. Army Research Institute for the Behavioral and Social Sciences. (AD-A072 313)

Shiflett, S.C. and S.M. Nealey
1972 The effects of changing leader power: a test of situational engineering. Organizational Behavior and Human Performance, 7(2):371-382.

Shima, H.
1968 The relationship between the leader's mode of interpersonal cognition and the performance of the group. Japanese Psychological Research, 10:13-30.

Simpson, S.N.
1980 Comment on meta-analysis of research on class size and achievement. Educational Evaluation and Policy Analysis, 2(3):81-83.

Singe, A.L.
1975 A Study of the Relationship Between Work Group Performance and Leader Motivation, Leader Behavior, and Situational Favorableness: An Application of the Contingency Theory of Leadership Effectiveness to Group Supervision in Multi Unit Elementary Schools. Unpublished Doctoral dissertation, University of Connecticut.

Skrzypek, G.J.
1969 The Relationship of Leadership Style to Task Structure, Position Power, and Leader Member Relations. Technical Report 34, U.S. Military Academy, West Point.

Smith, M.L.
1980 Publication bias and meta-analysis. Evaluation in Education: An International Review Series, 4(1):22-24.

Smith, M.L. and G.V. Glass
1980 Meta-analysis of research on class size and its relationship to attitudes and instruction. American Educational Research Journal, 17(4):419-433.

Smith, R.G.
1973 The Effects of Leadership Style, Leader Position Power, and Problem-Solving Method on Group Performance. Unpublished Doctoral dissertation, University of Maryland.

Stemler, J.G.
1980 Fiedler's LPC Scale: Behavioral and Attitudinal Correlates. Unpublished Doctoral dissertation, University of Cincinnati.

Stinson, J.E.
1972 Least preferred coworker as a measure of leadership style. Psychological Reports, 30:930.

Stinson, J.E. and L. Tracy
1974 Some disturbing characteristics of the LPC score. Personnel Psychology, 27:447-485.

Stock, W.A., M.A. Okun, M.J. Haring, W. Miller, C. Kinney, and R.W. Ceurvorst
1982 Rigor in data synthesis: a case study of reliability in meta-analysis. Educational Researcher, 11(6):10-14.

Stogdill, R.M.
1959 Individual Behavior and Group Achievement. New York: Oxford University Press.

Stogdill, R.M.
1971 Organizational leadership, in R.E. Stockhouse, V.F. Phillips and E. Owens (eds.), Frontiers of Leadership: The United States Air Force Academy Program (1970). Air Force Office of Scientific Research, Air Force Systems Command, USAF AFQSR-TR-71-1857, August.

Stone, V.
1979 Principal's Leadership Style, Situational Control and School Effectiveness. Unpublished Doctoral dissertation, University of California, Berkeley.

Strube, M.J. and J.E. Garcia
1981 A meta-analytic investigation of Fiedler's Model of Leadership Effectiveness. Psychological Bulletin, 90(2):307-321.

Suppe, F. (ed.)
1979 The Structure of Scientific Theories (2nd ed.). Urbana: University of Illinois Press.

Taylor, D.
1975 An Investigation of the Effects of the Relative Stabilities of Some Prominent Measures of Leadership. Unpublished Master's thesis, University of Washington.

Thomson, J.G.
1972 An Empirical Study of the Relationship Between Ontario Secondary Principals' Leadership Effectiveness and the Helping Relationship in Ontario Secondary Teachers. Unpublished Doctoral dissertation, University of Ottawa.

Tucker, J.H.
1977 Leadership in Autonomous Group Environments, Task/Person Orientation and Interpersonal Competance. Unpublished Doctoral dissertation, Georgia State (University School of Arts and Sciences).

Tukey, J.W.
1977 Exploratory Data Analysis. Reading: Addison-Wesley.

Turner, J.H.
1972 The Contingency Model of Leadership Effectiveness: An Empirical Investigation and Evaluation. Unpublished Doctoral dissertation, The City University of New York. (incorrectly cited as "Tumes" in Fiedler, 1978)

Utecht, R.E.
1970 An Empirical Study of Military Leadership Success Using Fiedler's Contingency Theory. Unpublished Doctoral dissertation, Arizona State University.

Utz, R.D.
1980 Fiedler's Situational Leadership Theory as a Measurement of Leadership Variables in Public Middle and Junior High Schools. Unpublished Doctoral dissertation, University of Michigan.

Van Gundy, A.B.
1975 Leadership Styles of College and University Presidents: An Application of Fiedler's Contingency Model. Unpublished Doctoral dissertation, Ohio State University.

Van Gundy, A.B. and L.L. Haynes
1978 A comparison of college presidents using Fiedler's Contingency Model. Journal of Negro Education, 47(3):215-229.

Vecchio, R.P.
1977 An empirical examination of the validity of Fiedler's Model of Leadership Effectiveness. Organizational Behavior and Human Performance, 19(1):180-206.

Vecchio, R.P.
1979 A test of the cognitive complexity interpretation of the least preferred coworker scale. Educational and Psychological Measurement, 39(3):523-526.

Wagner, C. and L. Wheeler
1969 Model, need, and cost effects in helping behavior. Journal of Personality and Social Psychology, 12(2):111-116.

Watkins, J.F.
1969 An inquiry into the principal-staff relationship. Journal of Educational Research, 63(1):11-15.

Wearing, A. and D.W. Bishop
1970 The Contingency Model and the Functioning of Military Squads. Technical Report No. 70-15 (ONR 1834-36), Organizational Research, Department of Psychology, University of Washington, Seattle.

White, K.R.
1979 The Relationship Between Socioeconomic Status and Academic Achievement. Paper presented to the American Educational Research Association, San Francisco.

Williams, L.B. and W.K. Hoy
1971 Principal-staff relations: situational mediator of effectiveness. Journal of Educational Administration, 9(1):66-73.

Wood, M.T. and R.S. Sobel
1970 Effects of similarity of leadership style at two levels of management on the job satisfaction of the first level manager. Personnel Psychology, 23:577-590.

Woodward, J.
1965 Industrial Organization: Theory and Practice. London: Oxford University Press.

Wyrick, F.I.
1972 The Effect of Departmental Leadership on Faculty Satisfaction and Departmental Effectiveness in a Big Ten University. Unpublished Doctoral dissertation, University of Illinois, Urbana.

Yukl, G.
1970 Leader LPC scores: attitude dimensions and behavioral correlates. Journal of Social Psychology, 80:207-212.

Yunker, G.W.
1973 A Field Test of Fiedler's Contingency Model of Leadership and a Proposed Revised Model. Unpublished Doctoral dissertation, Southern Illinois University, Carbondale.

APPENDIX A

OCTANT STUDIES
(100 Series)

OCTANT STUDIES[1]

Studies classified as "CC:Fiedler" are indicated by an asterisk (*) in this list.

Bagley, M.C.
1972 Situational Leadership in Graduate Departments of Physical Education. Unpublished Doctoral dissertation, University of Illinois.

Beebe, R.J.
1974 The Least Preferred Coworker Score of the Leader and the Productivity of Small Interacting Task Groups in Octants II and IV of the Fiedler Contingency Model. Unpublished Doctoral dissertation, The College of William and Mary, Virginia.

Bell, N.H.
1970 Leadership Effectiveness Within the North Carolina Department of Community Colleges: An Application and Extension of F.E. Fiedler's Contingency Model to Co-Acting Groups. Unpublished Doctoral dissertation, North Carolina State University at Raleigh.

Blanchard, L.
1978 The Leadership Effectiveness of Wisconsin Elementary School Principals. Unpublished Doctoral dissertation, University of Wisconsin.

Brown, D.I. and R.A. Smolinski
1974 The Contingency Model of Leadership Effectiveness and its Implications for Military Managers. Unpublished Master's thesis, Air Force Institute of Technology, Air University (AD 787 184).

Butterfield, D.A.
1968 An Integrative Approach to the Study of Leadership Effectiveness in Organizations. Unpublished Doctoral dissertation, University of Michigan.

---

[1] There are 40 rather than 38 references on this list because Csoka and Fiedler (1971), Csoka (1972b), and Csoka and Fiedler (1972) were regarded as one "multiple sample" study.

Chemers, M.M., R.W. Rice, E. Sundstrom, and W.M. Butler
1974 Leader LPC, Training, and Effectiveness: An Experimental Examination. Technical Report No. 73-42 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

* Chemers, M.M., and G.J. Skrzypek
1972 The Contingency Model of Leadership Effectiveness. Journal of Personality and Social Psychology, 24(2):172-177.

Csoka, L.S.
1972b A Validation of the Contingency Model Approach to Leadership Experience and Training. Technical Report 72-32 (ONR-177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Csoka, L.S.
1975 Relationship between organizational climate and the situational favorableness dimension of Fiedler's Contingency Model. Journal of Applied Psychology, 60:273-277.

Csoka, L.S., and F.E. Fiedler
1971 The Effect of Leadership Experience and Training in Structured Military Tasks — A Test of the Contingency Model. Technical Report No. 71-21 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Csoka, L.S. and F.E. Fiedler
1972 The effect of military leadership training: a test of the Contingency Model. Organizational Behavior and Human Performance, 8(3):395-407.

Faust, R.M.
1972 Leader Styles of Union Stewards: A Test of Fiedler's Model. Unpublished Doctoral dissertation, University of Iowa.

Gilbert, M.A.
1977 An Organizational Approach to the Study of Productivity, Efficiency, and Satisfaction of AAA High-School Basketball Teams Based on Fiedler's Contingency Model and Taylor and Bowers' Survey of Organizational Conditions. Unpublished Doctoral dissertation, University of Oregon.

* Graen, G.B., J.B. Orris, and K.M. Alvares
1971a Contingency Model of Leadership Effectiveness: some experimental results. Journal of Applied Psychology, 55(3):196-201.

Grunstad, N.L.
1972 The Determination of Leader Behavior From the Interaction of Fiedler's LPC and Situational Favorableness. Unpublished Doctoral dissertation, Ohio State University.

* Hill, W.A.
1969 A situational approach to leadership effectiveness. Journal of Applied Psychology, 53(6):513-517.

Hopfe, M.E.
1968 A Case Study of Leadership in a College Using Fiedler's Contingency Model. · Unpublished Doctoral dissertation, Ohio University.

Hovey, D.E.
1974 The low powered leader confronts a messy problem: a test of Fiedler's theory. Academy of Management Journal, 17(2):358-362.

* Hunt, J.G.
1966 A Test of Fiedler's Leadership Contingency Model in Three Organizational Settings. Unpublished Doctoral dissertation, University of Illinois, Urbana.

Inciong, P.A.
1974 Leadership Styles and Team Success. Unpublished Doctoral dissertation, University of Utah.

Jacobs, H.L.
1975 A Critical Evaluation of Fiedler's Contingency Model of Leadership Effectiveness in its Application to Inter-Disciplinary Task-Groups in Public School Settings. Unpublished Doctoral dissertation, Columbia University Teachers College.

Kuehl, C.R.
1971 Small Group Productivity as Related to Leadership Style, Role Differentiation, Status Differentiation and Organizational Backgrounds of Members. Unpublished Doctoral dissertation, University of Iowa.

Lanaghan, R.C.
1972 Leadership Effectiveness in Selected Elementary Schools. Unpublished Doctoral dissertation, University of Illinois, Urbana-Champaign.

Martin, W.J.
1977 Fiedler's ContingencyModel in Classroom Teaching Situations. Unpublished Doctoral dissertation, University of Utah.

Martin, Y.M., G.B. Isherwood, and R.E. Lavery
1976 Leadership effectiveness in teacher probation committees. Educational Administrative Quarterly, 12(2):87-99.

* Mellor, K.P.
1974 An Investigation of Fiedler's Contingency Model of Leadership Effectiveness as It Applies to Elementary Principals in Rhode Island. Unpublished Doctoral dissertation, University of Connecticut.

* Radtke, E.M.T.
1978 Leadership Effectiveness of Hospital-Diploma, Community College, and University Prepared Head Nurses: Interrelationships of Leadership Style and Situational Favorableness on Group Performance in Patient-Units of General Hospitals. Unpublished Doctoral dissertation, Wayne State University.

Rice, R.W. and M.M. Chemers
1973 Predicting the emergence of leaders using Fiedler's Contingency Model of Leadership Effectiveness. Journal of Applied Psychology, 57:281-287.

Rubin, G.J.
1971 A Modified Contingency Model for Leadership Effectiveness. Unpublished Doctoral dissertation, University of Tennessee.

* Saha, S.K.
1972 An Experimental Validation and Extension of Fiedler's Contingency Model of Leadership Effectiveness. Unpublished Doctoral dissertation, University of British Columbia.

Sashkin, M.
1970 Supervisory Leadership in Problem Solving Groups: Experimental Tests of Fred Fiedler's "Theory of Leadership Effectiveness" in the Laboratory Using Role Play Methods. Unpublished Doctoral dissertation, University of Michigan.

Sashkin, M., F.C. Taylor and R.C. Tripathi
1974 An analysis of situational moderating effects on the relationship between LPC and other psychological measures. Journal of Applied Psychology, 59:731-740.

Schneier, C.E.
1978 The Contingency Model of Leadership: an extension to emergent leadership and leader's sex. Organizational Behavior and Human Performance, 21:220-239.

* Shiflett, S.C. and S.M. Nealey
1972 The effects of changing leader power: a test of situational engineering. Organizational Behavior and Human Performance, 7(2):371-382.

* Singe, A.L.
1975 A Study of the Relationship Between Work Group Performance and Leader Motivation, Leader Behavior, and Situational Favorableness: An Application of the Contingency Theory of Leadership Effectiveness to Group Supervision in Multi Unit Elementary Schools. Unpublished Doctoral dissertation, University of Connecticut.

Stemler, J.G.
1980 Fiedler's LPC Scale: Behavioral and Attitudinal Correlates. Unpublished Doctoral dissertation, University of Cincinnati.

* Turner, J.H.
1972 The Contingency Model of Leadership Effectiveness: An Empirical Investigation and Evaluation. Unpublished Doctoral dissertation, The City University of New York. (incorrectly cited as "Tumes" in Fiedler, 1978)

Van Gundy, A.B.
1975 Leadership Styles of College and University Presidents: An Application of Fiedler's Contingency Model. Unpublished Doctoral dissertation, Ohio State University.

Yunker, G.W.
1973 A Field Test of Fiedler's Contingency Model of Leadership and a Proposed Revised Model. Unpublished Doctoral dissertation, Southern Illinois University, Carbondale.

APPENDIX B

NON-OCTANT STUDIES
(100 Series)

NON-OCTANT STUDIES

Anderson, L.R.
1964 Some Effects of Leadership Training on Intercultural Discussion Groups. Technical Report 18 (ONR 1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Arbuthnot, J.
1968 Relationships Among Psychological Differentiations and Leadership Style. Unpublished Master's thesis, Cornell University.

Arnett, M.D.
1978c Least preferred co-worker score as a measure of cognitive complexity. Perceptual and Motor Skills, 47(2):567-574.

Bons, P.M.
1974 The Effect of Changes in Leadership Environment on the Behavior of Relationship and Task-Motivated Leaders. Unpublished Doctoral dissertation, University of Washington.

Bons, P.M. and F.E. Fiedler
1976 Changes in organizational leadership and the behavior of relationship- and task-motivated leaders. Administrative Science Quarterly, 21(3):433-472.

Bons, P.M., A.R. Bass, and S.S. Komorita
1970 Changes in leadership style as a function of military experience and type of command. Personnel Psychology, 23(4):551-568.

Browne, T.J.
1974 Empirical Verification of Fiedler's LPC. Unpublished Master's thesis, University of Victoria, B.C., Canada.

Bugental, R.W.G.
1969 A Reinterpretation of Fiedler's ASo Score for Assumed Similarity of Opposites. Unpublished Doctoral dissertation, University of California, Los Angeles.

Carlson, J.
1977 The Effect of Leadership Style and Leader Position Power on Group Performance. Unpublished Doctoral dissertation, University of Maryland.

Chapman, J.B.
1974 A Comparative Analysis of Male and Female Leadership Styles in Similar Work Environments. Unpublished Doctoral dissertation, University of Nebraska, Lincoln.

Chemers, M.M.
1969 Cross-cultural training as a means for improving situational favourableness. Human Relations, 22(6):531-546.

DeZonia, R.H.
1958 The Relationship Between Psychological Distance and Effective Task Performance. Unpublished Doctoral dissertation, University of Illinois.

Eagly, A.H.
1970 Leadership style and role differentiation as determinants of group effectiveness. Journal of Personality, 38(4):509-524.

Evans, M.G. and J. Dermer
1974 What does the least preferred co-worker scale really measure? A cognitive interpretation. Journal of Applied Psychology, 59:202-206.

Fiedler, F.E., G. O'Brien, and D. Ilgen
1969 The effect of leadership style upon the performance and adjustment of volunteer teams operating in a stressful foreign environment. Human Relations, 22(6):503-514.

Fishbein, M., E. Landy, and G. Hatch
1969a A consideration of two assumptions underlying Fiedler's Contingency Model for Leadership Effectiveness. American Journal of Psychology, 82(4):457-473.

Flynn, M.L.
1972 Relationships Between Special Education Administrators' Responses to Fiedler's Least Preferred Coworker Measure and Their Responses to Selected Hypothetical Group Task Situations. Unpublished Doctoral dissertation, University of Georgia.

Garland, P.
1973 The Effect of Principal-Teacher Interaction in Secondary School Environments: An Empirical Study. Unpublished Doctoral dissertation, University of Ottawa.

Gottheil, E. and C.G. Lauterbach
1969 Leader and squad attributes contributing to mutual esteem among squad members. Journal of Social Psychology, 77:69-78.

Graham, W.K.
1968 Description of leader behavior and evaluation of leaders as a function of LPC. Personnel Psychology, 21:457-464.

Graham, W.K.
1973 Leader behavior, esteem for the least preferred coworker, and group performance. Journal of Social Psychology, 90:59-66.

Green, S.G. and D.M. Nebeker
1974 Leader Behavior: Autonomous or Interactive? Technical Report 74-62 (ONR 177-473), Organizational Research, Department of Psychology, University of Washington, Seattle.

Green, S.G. and D.M. Nebeker
1977 The effects of situational factors and leadership style on leader behavior. Organizational Behavior and Human Performance, 19(2):368-377.

Green, S., D. Nebeker and A. Boni
1976 Personality and situational effects on leader behavior. Academy of Management Journal, 19(2):184-194.

Gruenfeld, L. and J. Arbuthnot
1968 Field independence, achievement values, and the evaluation of a competency related dimension on the least preferred co-worker (LPC) measure. Perceptual and Motor Skills, 27:991-1002.

Gruenfeld, L.W., D.E. Rance and P. Weissenburg
1969 The behavior of task-oriented (low LPC) and socially-oriented (high LPC) leaders under several conditions of social support. Journal of Social Psychology, 79:99-107.

Hardy, R.C.
1969 The Effect of Leadership Style on the Performance of Small Classroom Groups: A Test of the Contingency Model. Unpublished Doctoral dissertation, Indiana University.

Hardy, R.C.
1972 A developmental study of relationships between birth order and leadership style for two distinctly different American groups. Journal of Social Psychology, 87:147-148.

Hardy, R.C.
1975 A test of poor leader-member relations cells of the Contingency Model on elementary school children. Child Development, 46(4):958-964.

Hardy, R.C. and J.F. Bohren
1975 The effect of experience on teacher effectiveness: A test of the Contingency Model. Journal of Psychology, 89(1):159-163.

Hardy, R.C., J. Carey, B. Eberwein, and J. Eliot
1976 Comparison of task- and relation-oriented individuals' ability to perceive viewpoints of others. Perceptual and Motor Skills, 42(3, Pt. 2):1028-1030.

Hardy, R.C., S. Sack, and F. Harpine
1973 An experimental test of the Contingency Model on small classroom groups. Journal of Psychology, 85(1):3-16.

Hawley, D.E.
1969 A Study of the Relationship Between the Leader Behavior and Attitudes of Elementary School Principals. Unpublished Master's thesis, University of Saskatchewan.

Hosking, D.M.
1978 A Critical Evaluation of Fiedler's Predictor Measures of Leadership Effectiveness. Unpublished Doctoral dissertation, University of Warwick.

Hunt, J.G.
1971 Leadership style effects at two managerial levels in a simulated organization. Administrative Science Quarterly, 16(4):476-485.

Hutchins, E.B.
1958 Task-Oriented and Quasi-Therapeutic Role Functions of the Small Group Leader. Unpublished Doctoral dissertation, University of Illinois, Urbana.

Jones, J.R. and M. Johnson
1972 LPC as a modifier of leader-follower relationships. Academy of Management Journal, 15(2):185-196.

Keane, F.J.
1976 The Relationship of Sex, Teacher-Leadership Style and Teacher-Leader Behavior in Teacher-Student Interaction. Unpublished Doctoral dissertation, Boston University.

Konar-Goldband, E., R.W. Rice, and W. Monkarsh
1979 Time-phased interrelationships of group atmosphere, group performance, and leader style. Journal of Applied Psychology, 64(4):401-409.

Larson, L.L. and K. Rowland
1973 Leadership style, stress, and behavior in task performance. Organizational Behavior and Human Performance, 9(3):407-420.

Larson, L.L. and K. Rowland
1974 Leadership style and cognitive complexity. Academy of Management Journal, 17(1):37-45.

Mai-Dalton, R.
1976 The Influence of Training and Position Power on the Behavior of Male and Female Leaders. Technical Report 76-86, Organizational Research, Department of Psychology, University of Washington, Seattle.

McKague, T.R.
1968 A Study of the Relationship Between School Organizational Behavior and the Variables of Bureaucratization and Leader Attitudes. Unpublished Doctoral dissertation, University of Alberta.

McNamara, V.D.
1968 Leadership, Staff, and School Effectiveness. Unpublished Doctoral dissertation, University of Alberta.

Meuwese, W.A.T.
1964 The Effect of the Leader's Ability and Interpersonal Attitudes on Group Creativity Under Varying Conditions of Stress. Unpublished Doctoral dissertation, University of Amsterdam, Amsterdam.

Michaelson, L.K.
1973 Leader orientation, leader behavior, group effectiveness and situational favorability: an empirical extension of the Contingency Model. Organizational Behavior and Human Performance, 9(2):226-245.

Millard, R.J.
1981 A Comparative Analysis of Male and Female Management Style and Perceived Behavior Patterns. Unpublished Doctoral dissertation, Brigham Young University.

Mitchell, T.R.
1969 Leader Complexity, Leadership Style, and Group Performance. Unpublished Doctoral dissertation, University of Illinois.

Mitchell, T.R.
1970 The construct validity of three dimensions of leadership research. Journal of Social Psychology, 80:89-94.

Muller, H.P.
1970 Relationship between time-span of discretion, leadership behavior, and Fiedler's LPC scores. Journal of Applied Psychology, 54(2):140-144.

Nayar, E.S.K., H. Touzard and D.A. Summers
1968 Training tasks, and mediator orientation in heterocultural negotiations. Human Relations, 21(3):283-294.

Nealey, S.M. and T.W. Owen
1970 A multitrait-multimethod analysis of predictors and criteria of nursing performance. Organizational Behavior and Human Performance, 5:348-365.

Nebeker, D.M.
1972 A Model for Supervisor Behavior: An Integration of Fiedler's Contingency Model and Expectancy Theory. Unpublished Doctoral dissertation, University of Washington.

Ninane, P. and F.E. Fiedler
1970 Member reactions to success and failure of task groups. Human Relations, 23(1):3-13.

Ranhosky, F.W.
1978 Personality and Behavioral Correlates of Leadership Effectiveness of Assistant Principals. Unpublished Doctoral dissertation, Fordham University.

Reilly, A.J.
1968 The Effects of Different Leadership Styles on Group Performance: A Field Experiment. Iowa Industrial Relations Center, Iowa State University, Ames.

Rice, R.W. and M.M. Chemers
1975 Personality and situational determinants of leader behavior. Journal of Applied Psychology, 60(1):20-27.

Rice, R.W., F.J. Seaman, and D.J. Garvin
1978 An empirical examination of the esteem for Least Preferred Co-worker (LPC) construct. Journal of Psychology, 98(2):195-205.

Schriesheim, C.A.
1979 Social desirability and leader effectiveness. Journal of Social Psychology, 108(1):89-94.

Sepic, F.T.
1979 The Effects of Leadership Style, Functional Authority and Stress on the Relationships of Supervisory Experience and Managerial Performance in the Hotel and Restaurant Industry. Unpublished Doctoral dissertation, University of Washington, Seattle.

Shiflett, S.C.
1974 Stereotyping and esteem for one's least preferred co-worker. Journal of Social Psychology, 93:55-65.

Shima, H.
1968 The relationship between the leader's mode of interpersonal cognition and the performance of the group. Japanese Psychological Research, 10:13-30.

Smith, R.G.
1973 The Effects of Leadership Style, Leader Position Power, and Problem-Solving Method on Group Performance. Unpublished Doctoral dissertation, University of Maryland.

Stinson, J.E.
1972 Least preferred coworker as a measure of leadership style. Psychological Reports, 30:930.

Stinson, J.E. and L. Tracy
1974 Some disturbing characteristics of the LPC score. Personnel Psychology, 27:447-485.

Stone, V.
1979 Principal's Leadership Style, Situational Control and School Effectiveness. Unpublished Doctoral dissertation, University of California, Berkeley.

Thomson, J.G.
1972 An Empirical Study of the Relationship Between Ontario Secondary Principals' Leadership Effectiveness and the Helping Relationship in Ontario Secondary Teachers. Unpublished Doctoral dissertation, University of Ottawa.

Utecht, R.E.
1970 An Empirical Study of Military Leadership Success Using Fiedler's Contingency Theory. Unpublished Doctoral dissertation, Arizona State University.

Van Gundy, A.B. and L.L. Haynes
1978 A comparison of college presidents using Fiedler's Contingency Model. Journal of Negro Education, 47(3):215-229.

Watkins, J.F.
1969 An inquiry into the principal-staff relationship. Journal of Educational Research, 63(1):11-15.

Williams, L.B. and W.K. Hoy
1971 Principal-staff relations: situational mediator of effectiveness. Journal of Educational Administration, 9(1):66-73.

Wood, M.T. and R.S. Sobel
1970 Effects of similarity of leadership style at two levels of management on the job satisfaction of the first level manager. Personnel Psychology, 23:577-590.

Wyrick, F.I.
1972 The Effect of Departmental Leadership on Faculty Satisfaction and Departmental Effectiveness in a Big Ten University. Unpublished Doctoral dissertation, University of Illinois, Urbana.

APPENDIX C

"300 SERIES" DOCUMENTS

"300 SERIES" DOCUMENTS

Anderson, L.S.
1980 "Billy Budd", "The Caine Mutiny" and Fiedler's "Contingency Model": An Inferential Analysis of Leadership. Unpublished Doctoral dissertation, University of North Carolina.

Argyris, C.
1976 Theories of action that inhibit individual learning. American Psychologist, 31(9):638-654.

Arnett, M.D.
1978b Behavioral correlates of least preferred co-worker scale. Psychological Reports, 43(3, Pt. 2): 1102.

Ashour, A.S.
1973a The Contingency Model of Leadership Effectiveness: An evaluation. Organizational Behaviour and Human Performance, 9(3):339-355.

Ashour, A.S.
1973b Further discussion of Fiedler's Contingency Model of Leadership Effectiveness. Organizational Behavior and Human Performance, 9(3):369-376.

Barrow, J.C.
1977 The variables of leadership: a review and conceptual framework. Academy of Management Review, 2(2):231-246.

Beach, L.R. and B. Beach
1978 A note on judgments of situational favorableness and probability of success. Organizational Behavior and Human Performance, 22(1):69-74.

Blades, J.W. and F.E. Fiedler
1973 Participative Management, Member Intelligence, and Group Performance. Technical Report 73-40, Organizational Research, Department of Psychology, University of Washington, Seattle.

Blades, J.W. and F.E. Fiedler
1976 The Influence of Intelligence, Task Ability, and Motivation on Group Performance. Technical Report 76-78 (ONR 170-761, ARI-G0005), Organizational Research, Department of Psychology, University of Washington, Seattle.

Borden, D.F.
1976 Leadership Experience and Organizational Performance: A Review of the Literature. Technical Report 76-85 (ARI: DAHC-19-73-G-0005), Organizational Research, Department of Psychology, University of Washington, Seattle.

Chemers, M.M.
1970 The relationship between birth order and leadership style. Journal of Social Psychology, 80:243-244.

Csoka, L.S.
1972a Intelligence: A Critical Variable in Leadership Experience. Technical Report 72-34 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Csoka, L.S.
1973 Organic and Mechanistic Organizational Climates and the Contingency Model. Technical Report No. 73-41 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Doyle, W.J. and W.P. Ahlbrand
1973 Hierarchical group performance and leader orientation. Administrator's Notebook, 22(3).

Ebersole, P.D.
1969 Investigation of Fiedler's Leadership Effectiveness Model in a Quasi-Therapeutic Situation with Hospitalized Patients. Unpublished Doctoral dissertation, Univeristy of California, Los Angeles.

Fiedler, F.E.
1968 Leadership Experience and Leader Performance — Another Hypothesis Shot to Hell. Technical Report No. 70 (ONR 1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Fiedler, F.E.
1969 Style or circumstance: the leadership enigma. Psychology Today 2(10):38-43.

Fiedler, F.E.
1970a On the Death and Transfiguration of Leadership Training. Technical Report No. 70-16(ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1970b Personality, Motivational Systems, and Behavior of High and Low LPC Persons. Technical Report No. 70-12 (ONR 1834-36), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1971b Note on the methodology of the Graen, Orris, and Alvares studies testing the Contingency Model. Journal of Applied Psychology, 55(3):202-204.

Fiedler, F.E.
1971c Personality and Situational Determinants of Leader Behavior. Technical Report No. 71-18 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1971d Validation and extension of the Contingency Model of Leadership Effectiveness: a review of empirical findings. Psychological Bulletin, 76(2):128-148.

Fiedler, F.E.
1972a How do you make leaders more effective? New answers to an old puzzle. Organizational Dynamics, 1(2):3-18.

Fiedler, F.E.
1972b Leadership Experience and Leadership Training — Some New Answers to an Old Problem. Technical Report No. 72-36 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1973a The Contingency Model — a reply to Ashour. Organizational Behavior and Human Performance, 9(3):356-368.

Fiedler, F.E.
1973b Predicting the effects of leadership training and experience from the Contingency Model: a clarification. Journal of Applied Psychology, 57(2):110-113.

Fiedler, F.E.
1973c Stimulus/response: the trouble with leadership training is that it doesn't train leaders. Psychology Today, 6(9):23-30, 92.

Fiedler, F.E.
1973d Toward a Comprehensive System of Leadership Utilization. Technical Report No. 73-50 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1974 The Contingency Model — new directions for leadership utilization. Contemporary Business, Autumn, 65-79.

Fiedler, F.E.
1976a Intelligence and Group Performance: A Multiple Screen Model. Technical Report No. 76-79 (ARI-G-0005) Organizational Research, Department of Psychology, University of Washington, Seattle.

Fiedler, F.E.
1976b The leadership game: matching the man to the situation. Organizational Dynamics, 4(3):6-16.

Fiedler, F.E.
1979 Responses to Sergiovanni. Educational Leadership, 36(9): 394-396.

Fiedler, F.E. and A.F. Leister
1977 Leader intelligence and task performance: a test of a multiple screen model. Organizational Behavior and Human Performance, 20(1):1-14.

Fiedler, F.E., E.H. Potter, M.M. Zais, and W.A. Knowlton
1979 Organizational stress, and the use and misuse of managerial intelligence and experience. Journal of Applied Psychology, 64(6):635-647.

Foa, U.G., T.R. Mitchell and F.E. Fiedler
1971 Differentiation matching. Behavioral Science, 16(2):130-142.

Fox, W.M.
1976 Reliabilities, means, and standard deviations for LPC scales: instrument refinement. Academy of Management Journal, 19(3):450-461.

Fox, W.M.
1982a A test of Octant I of Fiedler's Contingency Model with training dependent coacting task groups. Replications in Social Psychology, 2(1):47-49.

Fox, W.M., W.A. Hill and W.H. Guertin
1971 Factor Analysis of Least Preferred Coworker Scales. Technical Report 70-1, College of Business, University of Florida, Gainesville.

Geyer, L.J. and J.W. Julian
1973 Manipulation of situational favorability in tests of the Contingency Model. Journal of Psychology, 84(1):13-21.

Graen, G.B., D. Alvares, J.B. Orris, and J.A. Martella
1970 The Contingency Model of Leadership Effectiveness: antecedent and evidential results. Psychological Bulletin, 74(4):285-296.

Graen, G.B., J.B. Orris, and K.M. Alvares
1971b Contingency Model of Leadership Effectiveness: some methodological issues. Journal of Applied Psychology, 55(3):205-210.

Hansson, R.O. and F.E. Fiedler
1972 Perceived Similarity, Personality, and Attraction to Large Organizations. Technical Report 72-39 (ONR-177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Hawkins, C.A.
1962 A Study of Factors Mediating a Relationship Between Leader Rating Behavior and Group Productivity. Unpublished Doctoral dissertation, University of Minnesota, Minneapolis.

Helmich, D.L.
1975a Corporate succession: an examination. Academy of Management Journal, 18(3):429-441.

Helmich, D.L.
1975b The executive interface and president's leadership behavior. Journal of Business Research, 3(1):43-52.

Hendrix, W.H.
1976 Contingency Approaches to Leadership: A Review and Synthesis. Technical Report 76-17 (Air Force Human Resources Laboratory), Lackland Air Force Base, Texas (ED 130013).

Hill, W.A.
1969b The LPC leader: a cognitive twist, in W.G. Scott and P.P. Le Breton (eds.), Academy of Management Proceedings (29th Annual Meeting), August:125-130.

Hosking, D.M. and C.A. Schriesheim
1978 Improving leadership effectiveness: the Leader Match concept. Administrative Science Quarterly, 23(3):496-505.

Hunt, J.G., R.N. Osborn and L.L. Larson
1975 Upper level technical orientation and first-level leadership within a noncontingency and contingency framework. Academy of Management Journal, 18(3):476-488.

Jackson, J.J.
1977 Leadership styles and recreation programme delivery systems. Canadian Journal of Applied Sports Sciences, 2(4):207-211.

Jacoby, J.
1968 Creative ability of task-oriented versus person-oriented leaders. The Journal of Creative Behavior, 2(4):249-253.

Justis, R.T., B.L. Kedia, and D.B. Stephens
1978 The effect of position power and perceived task competence on trainer effectiveness: a parital utilization of Fiedler's Contingency Model of Leadership. Personnel Psychology, 31(1): 83-93.

Kerr, S. and A. Harlan
1973 Predicting the effectiveness of leadership training and experience from the Contingency Model: some remaining problems. Journal of Applied Psychology, 57(2):114-117.

Korman, A.K.
1973 On the development of contingency theories of leadership: some methodological considerations and a possible alternative. Journal of Applied Psychology, 58(3):384-387.

Lord, R.G., J.S. Phillips, and M.C. Rush
1980 Effects of sex and personality on perceptions of emergent leadership, influence, and social power. Journal of Applied Psychology, 65(2):176-182.

McKague, T.R.
1970 LPC — a new perspective on leadership. Educational Administration Quarterly, 6(3):1-14.

McMahon, J.T.
1972 The contingency theory: logic and method revisited. Personnel Psychology, 25:697-710.

Mitchell, T.R., A. Biglan, G.R. Oncken, and F.E. Fiedler
1970 The Contingency Model: Criticisms and Suggestions. Technical Report No. 70-11 (ONR N00014-67-A-0103-0012), Organizational Research, Department of Psychology, University of Washington, Seattle.

Mitchell, T.R., J.R. Larson and S.G. Green
1977 Leader behavior, situational moderators, and group performance: an attributional analysis. Organizational Behavior and Human Performance, 18(1):254-268.

Nealey, S.M. and F.E. Fiedler
1968 Leadership functions of middle managers. Psychological Bulletin, 70(5):313-329.

O'Brien, G.E., A. Biglan and J. Penna
1972 Measurement of the distribution of potential influence and participation in groups and organizations. Journal of Applied Psychology, 56(1):11-18.

O'Brien, G.E. and F.E. Fiedler
1977 Measurement of interactive effects of leadership style and group structure upon group performance. Australian Journal of Psychology, 29(1):59-71.

O'Brien, G.E., F.E. Fiedler and T. Hewett
1971 The effects of programmed culture training upon the performance of volunteer medical teams in Central America. Human Relations, 24(3):209-231.

Posthuma, A.B.
1970 Normative Data on the Least-Preferred Co-worker Scale (LPC) and the Group Atmosphere Questionnaire (GA). Technical Report 70-8, (ARPA Order 454, N00014-67-A-0103-0013, Office of Naval Research), Organizational Research, Department of Psychology, University of Washington, Seattle.

Potter, E.H. and F.E. Fiedler
1981 Stress and the utilization of staff member intelligence and experience. Academy of Management Journal, 24(2):361-376.

Rice, R.W.
1975 The Esteem For Least Preferred Coworker (LPC) Score: What Does It Measure? Unpublished Doctoral dissertation, University of Utah.

Rice, R.W.
1979 Reliability and validity of the LPC scale: a reply. Academy of Management Review, 4(2):291-294.

Rice, R.W.
1981 Leader LPC and follower satisfaction: a review. Organizational Behavior and Human Performance, 28(1):1-25.

Schneier, C.E.
1979 Measuring cognitive complexity: developing reliability, validity, and norm tables for a personality instrument. Educational and Psychological Measurement, 39(3):599-612.

Schriesheim, C.A., B.D. Bannister, and W.H. Money
1979 Psychometric properties of the LPC scale: an extension of Rice's review. Academy of Management Review, 4(2):287-290.

Shiflett, S.C.
1973 The Contingency Model of Leadership Effectiveness: some implications of its statistical and methodological properties. Behavioral Science, 18:429-440.

Strube, M.J. and J.E. Garcia
1981 A meta-analytic investigation of Fiedler's Model of Leadership Effectiveness. Psychological Bulletin, 90(2):307-321.

Vecchio, R.P.
1979 A test of the cognitive complexity interpretation of the least preferred coworker scale. Educational and Psychological Measurement, 39(3):523-526.

Wagner, C. and L. Wheeler
1969 Model, need, and cost effects in helping behavior. Journal of Personality and Social Psychology, 12(2):111-116.

Wearing, A. and D.W. Bishop
1970 The Contingency Model and the Functioning of Military Squads. Technical Report No. 70-15 (ONR 1834-36), Organizational Research, Department of Psychology, University of Washington, Seattle.

Yukl, G.
1970 Leader LPC scores: attitude dimensions and behavioral correlates. Journal of Social Psychology, 80:207-212.

APPENDIX D

"400 SERIES" DOCUMENTS

"400 SERIES" DOCUMENTS

Allen, D.B.
1972 Heterogeneity of Research Interests and Effectiveness of University Departments. Technical Report 73-38 (USOE, OEG 0-70-3347), Organizational Research, Department of Psychology, University of Washington, Seattle.

Alpander, G.G.
1974 Planning management training programs for organizational development. Personnel Journal, 53(1):15-25.

Arnett, M.D.
1978a Attitudinal correlates of least preferred co-worker score. Psychological Reports, 43(3, Pt. 1):962.

Ayer, J.G.
1968 Effects of Success and Failure of Interpersonal and Task Performance upon Leader Perception and Behavior. Technical Report 26 (MD 2060), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Beach, B.H., T.R. Mitchell, and L.R. Beach
1975 Components of Situational Favorableness and Probability of Success. Technical Report 75-66, Organizational Research, Department of Psychology, University of Washington, Seattle.

Braxton, R.B. and W.L. Crosby
1974 A Comparative Analysis of Results Using Three Leadership Style Measurement Instruments. Unpublished Master's thesis, Air Force Institute of Technology, Air University (AD 787200).

Chemers, M.M. and D.A. Summers
1968 Group Atmoshpere and the Perception of Group Favorableness. Technical Reports 71 (ONR-1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Danielson, R.R.
1974 Leadership in Coaching: Description and Evaluation. Unpublished Doctoral dissertation, University of Alberta.

Deveau, R.J.
1976 The Relationship Between the Leadership Effectiveness of First-line Supervisors and Measures of Authoritarianism, Creativity, General Intelligence, and Leadership Style. Unpublished Doctoral dissertation, Boston University School of Education.

Fishbein, M., E. Landy, and G. Hatch
1969b Some determinants of an individual's esteem for the least preferred co-worker: an attitudinal analysis. Human Relations, 22(2):173-188.

Fox W.M.
1982b Behavioral correlates of leader LPC score as moderated by situational favorability. Replications in Social Psychology, in press.

Gochman, I. and F.E. Fiedler
1975 The Effect of Situational Favorableness on Leader and Member Perceptions of Leader Behavior. Technical Report No. 75-64 (ONR N00014-67-A-0103-0013), Organizational Research, Department of Psychology, University of Washington, Seattle.

Hewett, T.T. and G.E. O'Brien
1971 The Effects of Work Organization, Leadership Style and Member Compatability upon Small Group Productivity. Technical Report No. 71-22 (ONR 177-473), Organizational Research, Department of Psychology, University of Washington, Seattle.

Ilgen, D.R. and G. O'Brien
1968 The Effects of Task Organization and Member Compatability on Leader-Member Relations in Small Groups. Technical Report No. 58 (ONR-1834-36), Group Effectiveness Research Laboratory, University of Illinois, Urbana.

Kennedy, J.K.
1982 Middle LPC leaders and the Contingency Model of Leadership Effectiveness. Organizational Behavior and Human Performance, 30(1):1-14.

Miskel, C.G.
1974 Public School Principals' Leader Style, Organizational Situation, and Effectiveness. Final Report (NIE, DHEW-NE-G-00-3-0141), University of Kansas, Lawrence.

Miskel, C.G.
1977 Principals' attitudes toward work and co-workers, situational factors, perceived effectiveness, and innovation effort. Educational Administration Quarterly, 13(2):51-70.

Nealey, S.M. and M.R. Blood
1968 Leadership performance of nursing supervisors at two organizational levels. Journal of Applied Psychology, 52(5):414-422.

Nebeker, D.M., L.R. Beach and S.G. Green
1974 Situational Favorability and the Perception of Uncertainty: An Experimental Demonstration. Technical Report 74-61 (ONR 177-472), Organization Research, Department of Psychology, University of Washington, Seattle.

Nebeker, D.M. and R.O. Hansson
1972 Confidence in Human Nature and Leader Style. Technical Report 72-37 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Prothero, J. and F.E. Fiedler
1974 The Effect of Situational Change on Individual Behavior and Performance: An Extension of the Contingency Model. Technical Report 74-59 (ONR 177-472), Organizational Research, Department of Psychology, University of Washington, Seattle.

Rice, R.W., L.R. Bender, and A.G. Vitters
1980 Leader sex, follower attitudes toward women, and leadership effectiveness: a laboratory experiment. Organizational Behavior and Human Performance, 25(1):46-78.

Rice, R.W. and F.J. Seaman
1981 Interval analysis of the least preferred co-worker (LPC) scale. Educational and Psychological Measurement, 41(1):109-120.

Schmidt, D.E.
1976 Cognitive Homogeneity and the Contingency Model of Leadership Effectiveness: A Preliminary Study. Technical Report No. 76-83 (ARI-G0005), Organizational Research, Department of Psychology, University of Washington, Seattle.

Schmidt, D.E. and F.E. Fiedler
1976 The Contingency Model and Leader Behavior: An Instrumentality Approach. Technical Report No. 78-62 (ARI-G-0005), Organizational Research, Department of Psychology, University of Washington, Seattle.

Shiflett, S.C., R.G. Downey, and P.J. Duffy
1979 The Effects of Multidimensionality on the Predictive and Construct Validity of the LPC Scale. U.S. Army Research Institute for the Behavioral and Social Sciences. (AD-A072 313)

Taylor, D.
1975 An Investigation of the Effects of the Relative Stabilities of Some Prominent Measures of Leadership. Unpublished Master's thesis, University of Washington.

Vecchio, R.P.
1977 An empirical examination of the validity of Fiedler's Model of Leadership Effectiveness. Organizational Behavior and Human Performance, 19(1):180-206.

APPENDIX E

ANNOTATED BIBLIOGRAPHY OF REJECTED STUDIES

ANNOTATED BIBLIOGRAPHY OF REJECTED STUDIES

The 13 primary studies listed below were deleted from the original pool of 402 documents. An indication of the specific reasons for rejecting each study is given below the citation. In all cases, departures from Fiedler's Model were numerous or substantial enough to render the study not comparable with other Fiedler-based studies.

Bennett, M.
1977 Testing management theories cross-culturally. Journal of Applied Psychology, 62(5):578-581.

The LPC instrument was a five-point rather than an eight-point scale. The scores were therefore not comparable with those from other studies. In addition, the situational variables were not used to test any octant of the Model.

Bird, A.M.
1977 Team structure and success as related to cohesiveness and leadership. Journal of Social Psychology, 103(2):217-223.

Leader-member relations, conceptualized as team cohesion, was used as a dependent variable. Task structure assessed as "structured" based on the teams' win-loss record. Performance was also evaluated on the same win-loss record. The leaders' (coaches) LPC scores were assessed by both the group members (players) and the leaders.

Duffy, J.F.
1973 A Within-Subjects Validity Generalization of Fiedler's Contingency Model of Leadership. Unpublished Doctoral dissertation, Iowa State University.

Leadership style was not measured by the LPC

instrument. The assessment of leader-member relations was based on the subordinates' (group members) perceptions using a non-Fiedlerian instrument. Task structure and position power were assessed on the basis of job descriptions and no description was given of the validation of either measure.

Evans, M.G.
1973 A leader's ability to differentiate the subordinate's perception of the leader, and the subordinate's performance. Personnel Psychology, 26:385-395.

The 21-item "LPC" instrument contained only six pairs of adjectives found in any of Fiedler's 1958, 1964, 1967, 1974, 1976, and 1978 versions. The item variance of the "LPC" scores, rather than the scores, were used as the predictor of performance.

Fahy, F.A.
1972 The Contingency Model and the Schools: A Study of the Relationships Between Teaching Effectiveness, Leadership Style, and the Favorableness of the Situation. Unpublished Doctoral dissertation, Rutgers University.

Task structure was assessed on the basis of scores on the Pupil Control Ideology Instrument. Position power of the cooperating (sponsor) teacher was assessed by student teachers using the LPC instrument. The performance assessed was that of the student teachers.

Faulkner, R.T.B.
1980 Leadership Style of Principal Related to Degree of Bureaucratization of School. Unpublished Doctoral dissertation, University of Ottawa.

The amount of situational favourableness was based only on the GAS scores. Both task structure and position power were assumed on the basis of whether a school was "traditional" or "open". The relationship between the LPC scores of principals and the bureaucratic structure of schools is unclear.

Fearon, D.S.
1973 An Investigation of Fiedler's Contingency Theory of Leadership Effectiveness as It Applies to Community College Citizens Advisory Councils: An Exploratory Study. Unpublished Doctoral dissertation, University of Connecticut.

No inter-rater reliability measure was reported for the four judges' scores on the task structure and position power assessments to permit interpretation of the disparity among these reported scores. Of five scores reported for two measures, three appear to have been computed inaccurately. One of the rho statistics was also inaccurately calculated. These shortcomings make the assignment of groups to octants uncertain. They also cast doubt on the reliability of the reported correlations.

Goldstein, G.S. and C.W. Cole
1969 Leadership variables. The Journal of the Scientific Laboratories, Denison University, 50:59-67.

The supervisors' (leaders) task was assessed for structure while the department heads' (group members) performance was evaluated as the measure of effectiveness. The designation of groups as structured and unstructured is unclear. It appears as if the assessment was reversed during the conduct of the study.

Johnson, R. and B. Ryan
1976 A test of the Contingency Model of Leadership Effectiveness. Journal of Applied Social Psychology, 6(2):177-185.

The LPC instrument was not used to measure leadership style. By using instructionally induced directive and non-directive behaviour, the authors treated leadership as an independent, manipulated variable. This procedure clearly violates Fiedler's assumption that LPC is a relatively stable personality trait. In addition, there was a 25% (9/40) sample attrition yet all 40

leaders' scores were used in the performance correlations. The assessment of position power was validated using only one question. This study was also rejected in the Strube and Garcia (1981) investigation (see Chapter 3, p.68).

Reavis, C.A. and V.J. Derlaga
1976 Test of a Contingency Model of teacher effectiveness. Journal of Educational Research, 69(6):221-225.

The LPC instrument was not used to measure leadership style. In this study, two teachers were each asked to use two different approaches (person-oriented and task-oriented) to present the same lesson. This procedure violates Fiedler's assumption regarding LPC in the same way as the Johnson and Ryan study. Task structure was measured in terms of the lesson content rather than the task of teaching the content. No Fiedlerian instruments were used to measure situational favourableness. This study was also rejected by Strube and Garcia.

Tucker, J.H.
1977 Leadership in Autonomous Group Environments, Task/Person Orientation and Interpersonal Competence. Unpublished Doctoral dissertation, Georgia State University (School of Arts and Sciences).

In autonomous, leaderless groups task structure was assumed to be unstructured but the assumption was not validated. Situational favourableness was equated with the amount of influence provided for each individual in the group with specific (manipulated) environmental conditions. Position power was deemed to be inapplicable. The level of the GAS scores for groups in Octant VIII was too high to be considered "poor" (scores≥5.7). The treatment of these variables renders the assignment of groups to octants uncertain. Moreover, individual competency was used as the effectiveness measure.

Utz, R.D.
1980 Fiedler's Situational Leadership Theory as a Measurement of Leadership Variables in Public Middle and Junior High Schools. Unpublished Doctoral dissertation, University of Michigan.

Task structure was operationally defined by teacher performance. Position power was assessed by the school principal (leader). The GAS instrument was completed by the principal to assess leader-member relations and by the teachers (group members) to assess the "Principal Atmosphere" (member-member relations?). The effectiveness of the principal was assessed by the teachers but the measure was not correlated with the principal's LPC score. This study, which did not assign schools to octants, does not constitute a valid test of the Contingency Model.

Vecchio, R.P.
1977 An empirical examination of the validity of Fiedler's Model of Leadership Effectiveness. Organizational Behavior and Human Performance, 19(1):180-206.

This study was not included in the data base for the purposes of the meta-analysis. The main reason for its rejection was the way in which the leader-member relations variable was manipulated. This study was also rejected by Strube and Garcia (1981). However, there was no reason to reject the LPC scores reported in this study. To record these scores for future use in updating Posthuma's (1970) normative values, the Vecchio study was included in the "400 Series" (see Appendix D).

APPENDIX F

SUPPLEMENTARY TABLES AND ANALYSES

SUPPLEMENTARY TABLES AND ANALYSES

This appendix presents tables and analyses which supplement the material reported in Chapters 7, 8, and 9.

The first part of this appendix, relevant to Chapter 7, reports some interesting findings which emerged during the process of selecting the "rho + r" correlation listing. The data in Table F.1 (p. 349) compare the median correlations generated from each of the four correlations listings.

The second part of this appendix supplements the median correlation analysis of the task group types reported in Chapter 7. The display in Table F.2 (p. 351) compares the median correlations generated from the "r+rho" correlation listing with those from the "rho+r" correlation listing. The display in Table F.3 (p. 352) shows the median correlations generated from the "rho only" and "r only" correlation listings. Following Table F.3 is an analysis, by task group type, of the direction and magnitude of the obtained medians from the r and rho correlation listings (pp. 353 and 354).

The third part of this appendix supplements the box-and-whisker analysis reported in Chapter 8. The list of figures, each of which is preceded by a brief analysis, is shown on page 355.

The last part of this appendix supplements the analysis of methodological variability reported in Chapter 9 (p.224). Table F.4 (p. 361) shows the overlap in GAS scores used to designate groups as having "good" or "poor" leader-member relations.

THE COMPARISON OF THE OBTAINED MEDIANS FROM THE FOUR CORRELATION LISTINGS

The resolution of the unanticipated methodological problem regarding the four correlation listings revealed that the median correlations varied in both magnitude and direction depending upon which summary statistic had been used in the primary study. This variation appeared not only across the listings but also between the listings and Fiedler's predicted medians for each octant. As shown in Table F.1, the median correlations resulting from the "r + rho" and "rho + r" listings are identical in both magnitude and direction for five out of eight octants (II, IV, V, VI, and VIII). In the remaining three octants (I, III, and VII), the difference between the median values did not exceed 0.035. However, when the median correlations generated from the "r only" and the "rho only" listings were examined, neither set of medians consistently yielded the stronger results across all octants. In five octants (I, II, III, VI, and VII), the magnitude of all the medians from the "r only" listing was greater than that of the corresponding medians from the "rho only" listing. The "rho only" medians yielded the stronger values in the remaining three octants (IV, V, and VIII).

TABLE F.1: COMPARISON OF MEDIAN CORRELATIONS IN EACH LISTING

CORR'N	OCTANT							
LISTING	I	II	III	IV	V	VI	VII	VIII
r+rho	-.130	-.010	-.045	.055	.245	.070	-.070	.090
rho+r	-.165	-.010	-.050	.055	.245	.070	-.038	.090
rho	-.120	-.020	.035	.265	.245	.070	.170	-.325
r	-.315	-.100	-.097	-.070	.045	-.200	-.185	.150
Fiedler[1]	-.52	-.58	-.33	.47	.42	(+)[2]	.05	-.43

1. The median correlations predicted by Fiedler's Contingency Model.

2. Fiedler (1967) predicted only direction for Octant VI.

With respect to direction, the "r + rho" and "rho + r" medians were consistent with one another across all octants. However, the signs of the median correlations in Octants VII and VIII were opposite to those predicted by the Contingency Model. There was less consistency in the signs of the medians yielded by the "r only" and "rho only" listings. Those based on the "rho only" were directionally congruent with Fiedler's predicted values in seven octants, the exception being Octant III. The "r only" medians were congruent in only four octants (I, II, III, and V). If both magnitude and directional congruency are considered jointly, it would appear that the Pearson r correlation ("r only" listing) yields results more closely approximating Fiedler's predicted medians in Octants I, II, and III. By contrast, it is the Spearman rho correlation ("rho only" listing) which yields the more correspondent results in Octants IV, V, VII, and VIII. Although this finding casts some doubt on Fiedler's recommendation regarding the use of Spearman rho, it offers only limited support for the use of the Pearson r as the preferred statistic.

TABLE F.2: COMPARISON OF MEDIAN CORRELATIONS FROM "r+rho" AND "rho+r" LISTS FOR EACH TASK GROUP TYPE[1]

CORR'N GROUP	TGT	OCTANT I	II	III	IV	V	VI	VII	VIII
r+rho	1	-.035	-.010	-.100	-.065	.050	.070	-.185	.165
rho+r	1	-.040	-.010	-.175	-.065	-.345	.070	-.140	.165
r+rho	4	-.125	-.275	.140	.125	.280	.115	.170	-.332
rho+r	4	-.120	-.275	.145	.125	.265	.115	.090	-.332
r+rho	8	-.310	.020	-.555	.240	.050	-.400	-.365	-.285
rho+r	8	-.310	-.180	-.555	.165	.050	-.425	-.365	-.285
r+rho	2	-.320		-.040		.505		.260	
rho+r	2	-.320		-.040		.505		.260	
Fiedler[2]		-.52	-.58	-.33	.47	.42	(+)[3]	.05	-.43
r+rho n	1	26	8	13	18	9	3	12	16
rho+r n	1	25	8	12	18	8	3	11	16
r+rho n	4	10	4	13	12	7	6	7	11
rho+r n	4	11	4	14	12	8	6	8	11
r+rho n	8	11	2	12	2	12	2	12	2
rho+r n	8	11	2	12	2	12	2	12	2
r+rho n	2	5	0	6	0	4	0	4	0
rho+r n	2	5	0	6	0	4	0	4	0
Fiedler n		8	3	12	10	6	0	12	12

1. The n for each within octant median is shown in the bottom rows of the table.

2. The median correlations predicted by Fiedler's Contingency Model.

3. Fiedler (1967) predicted only direction for Octant VI.

TABLE F.3: COMPARISONS OF MEDIAN CORRELATIONS FROM "rho only" AND "r only" LISTS FOR EACH TASK GROUP TYPE[1]

CORR'N GROUP	TGT	OCTANT I	II	III	IV	V	VI	VII	VIII
rho	1	.050	-.010	-.280	.090	.210	.070	.300	.110
r		-.125	-.550	-.075	-.160	-.260		-.190	.165
rho	4	-.120	-.275	.190	.335	.280	.100	.170	-.330
r		.030		.040	-.045	.040	.790	-.070	-.380
rho	2	-.315		.050		.505		.245	
r		-.360		-.050				.260	
rho	8	-.170	-.500	.085	.200	-.320	-.650	-.010	
r		-.470	.020	-.565	.240	.230	-.400	-.465	-.285
Fiedler[2]		-.52	-.58	-.33	.47	.42	(+)[3]	.05	-.43
rho n	1	12	6	7	3	7	3	6	2
r n		16	3	8	15	3	0	7	14
rho n	4	9	4	8	8	7	5	7	8
r n		2	0	6	4	1	1	1	3
rho n	2	4	0	3	0	4	0	2	0
r n		1	0	3	0	0	0	2	0
rho n	8	4	1	4	1	4	1	2	0
r n		9	2	10	2	10	2	10	2
Fiedler n		8	3	12	10	6	0	12	12

1. The n for each within octant median is shown in the bottom rows of the table.

2. The median correlations predicted by Fiedler's Contingency Model.

3. Fiedler (1967) predicted only direction for Octant VI.

ANALYSIS OF MEDIAN CORRELATIONS FROM "rho only" AND "r only" LISTS FOR TASK GROUP TYPES

As indicated in Chapter 7, a more detailed analysis of the "rho only" and "r only" correlation listings is presented. The results from Table F.3 are summarized for each task group type in terms of the direction and magnitude of the obtained medians. These criteria are used to indicate which summary statistic yielded the greater frequency of support for the octants tested. The words "wrong" and "correct" denote, respectively, that the correlation sign is opposite to prediction and in the direction predicted by Fiedler.

Task Group 1

1. Direction: 6 out of 8 Spearman rho signs were "correct" while 4 out of 7 Pearson r signs were "wrong"

2. Magnitude: 5 out of 7 r values were higher than the corresponding values of rho while 2 out of 8 rho values were stronger than those of r

3. Conclusion: the greater frequency of support for the octants tested is associated with the rho statistic.

Task Group 4

1. Direction: 6 out of 8 Spearman rho signs were "correct" while 4 out of 7 Pearson r signs were "wrong"

2. Magnitude: 5 out of 8 values of rho were stronger than the corresponding values of r while 2 out of 7 r values were higher than those of rho

3. Conclusion: the greater frequency of support for the octants tested is associated with the rho statistic

Task Group 2

1. Direction: all 3 signs of r were "correct" while 1 out of 4 signs of rho were "wrong"

2. Magnitude: 2 out of 3 values of r were higher than the corresponding values of rho; one value of r and rho had the same strength

3. Conclusion: the greater frequency of support for the octants tested is associated with the r statistic

Task Group 8

1. Direction: 5 out of 8 signs of r were in the "correct" while 4 out of 8 rho signs were "wrong"

2. Magnitude: 4 out of 8 values of r were higher than the corresponding values of rho while 3 out of 7 rho values were stronger than those of r

3. Conclusion: the greater frequency of support for the octants tested is associated with the r statistic

When the "r only" and "rho only" medians were compared in terms of "correct" direction and stronger magnitude across the octants tested, an interesting difference emerged. A greater percentage of rho than r signs were in the "correct" direction (19/28 or 68% and 14/25 or 56%, respectively). By contrast, a higher proportion of r values were stronger than rho values (13/25 or 52% and 11/26 or 42%). In two cases, empty cells precluded a comparison of magnitude. When the two dimensions of support are considered jointly, there is virtually no difference in the outcome. Support was yielded by 54% (27/50) of the Pearson r values compared with 52% (28/54) of the Spearman rho values. The result is very similar to the finding of 54% support generated by the earlier analysis based only on the directional congruency criterion.

SUPPLEMENTARY ANALYSIS: BOX-AND-WHISKER

That these comparisons have been appended should not be regarded as an indication of lesser importance. They were excluded from the text of Chapter 8 because neither coacting task groups (TGT 2) nor nonacting task groups (TGT 8) permitted comparisons in all octants. In Octants II, IV, VI, and VIII, there were no reported correlations for TGT 2 and only two reported correlations in each of these same four octants for TGT 8. The task group type displays within CC:Fiedler and CC:Other have been included to illustrate the comparison between task groups stated by the primary author to be interacting (TGT 1) and task groups whose interacting nature was inferred from the information provided by the primary author (TGT 4). Comparisons between any two variables are possible only in octants showing both a shaded box and an unshaded box.

The list below indicates the Box-and-Whisker figures and supplementary analyses included in this appendix.

Figure F.1: Stated and Inferred Interacting Task Groups Within "CC:Fiedler" (TGT 1 and TGT 4)

Figure F.2: Stated and Inferred Interacting Task Groups Within "CC:Other" (TGT 1 and TGT 4)

Figure F.3: Coacting and Nonacting Task Groups (TGT 2 and TGT 8)

Figure F.4: Coacting and Inferred Interacting Task Groups (TGT 2 and TGT 4)

Figure F.5: Inferred Interacting and Nonacting Task Groups (TGT 4 and TGT 8)

Stated and Inferred Interacting Task Groups (TGT 1 and TGT 4) Within CC:Fiedler

In Octant III, the two task group types are clearly different in spite of the fact that they are both interacting groups within the same comparability category. This suggests that the inference of interaction may be associated with the differences. However, in light of numerous other possible reasons for the differences (e.g., methodological variations in the studies, organizational setting), this statement should not be regarded as definitive. The remaining four octants are probably more similar than they are different.

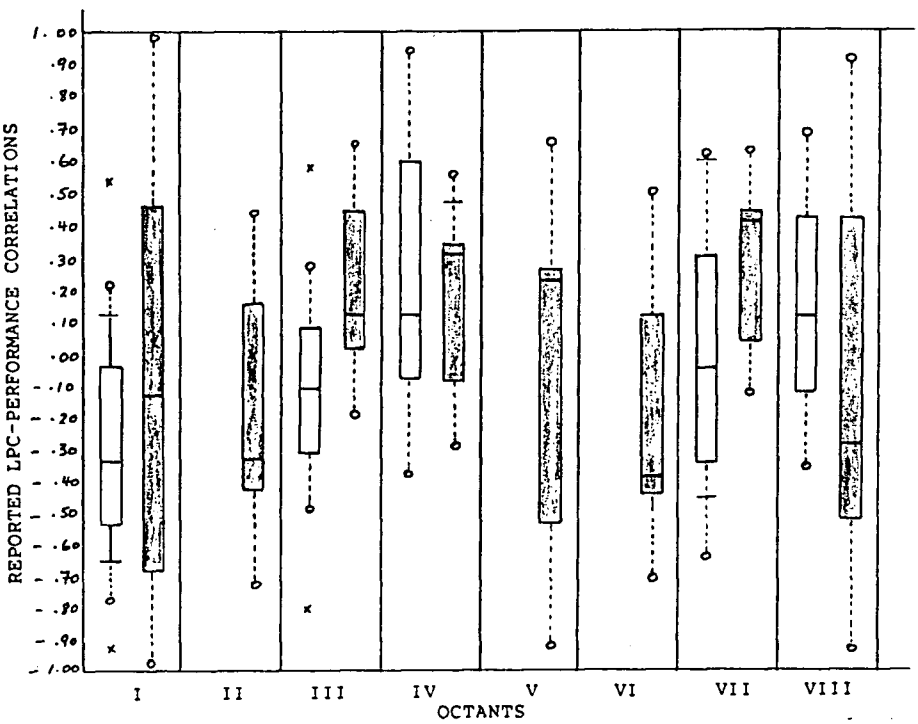


FIGURE F.1: BOX-AND-WHISKER: TASK GROUP TYPE WITHIN COMPARABILITY CATEGORIES

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

CC:F X TGT 1 CC:F X TGT 4

Stated and Inferred Interacting Task Groups (TGT 1 and TGT 4) Within CC:Other

There is a striking difference between the two task group types in Octant VIII despite the fact that both are in CC:Other. Why this difference should be so distinct in only one octant is not at all clear, particularly when the task groups appear to yield quite similar results in Octants I, III, V, and VII.

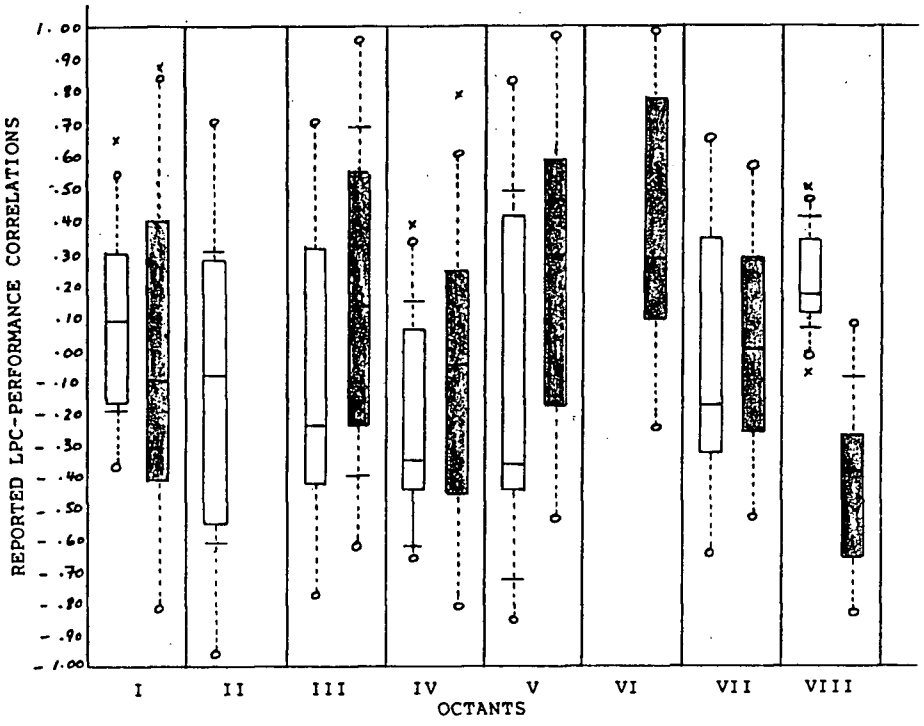


FIGURE F.2: BOX-AND-WHISKER: TASK GROUP TYPES WITHIN COMPARABILITY CATEGORY

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

CC:O X TGT 1 CC:O X TGT 4

Coacting and Nonacting Task Groups (TGT 2 and TGT 8)

In Octant I, very similar results emerge from both task group types. By contrast, the correlations are quite different in Octant VII as indicated by the lack of overlap of the two interquartile ranges. In spite of the overlap between the boxes in Octants III and V, there is a considerable difference between the distributions of correlations. In both octants, the range of values in TGT 8 is about twice that of TGT 2.

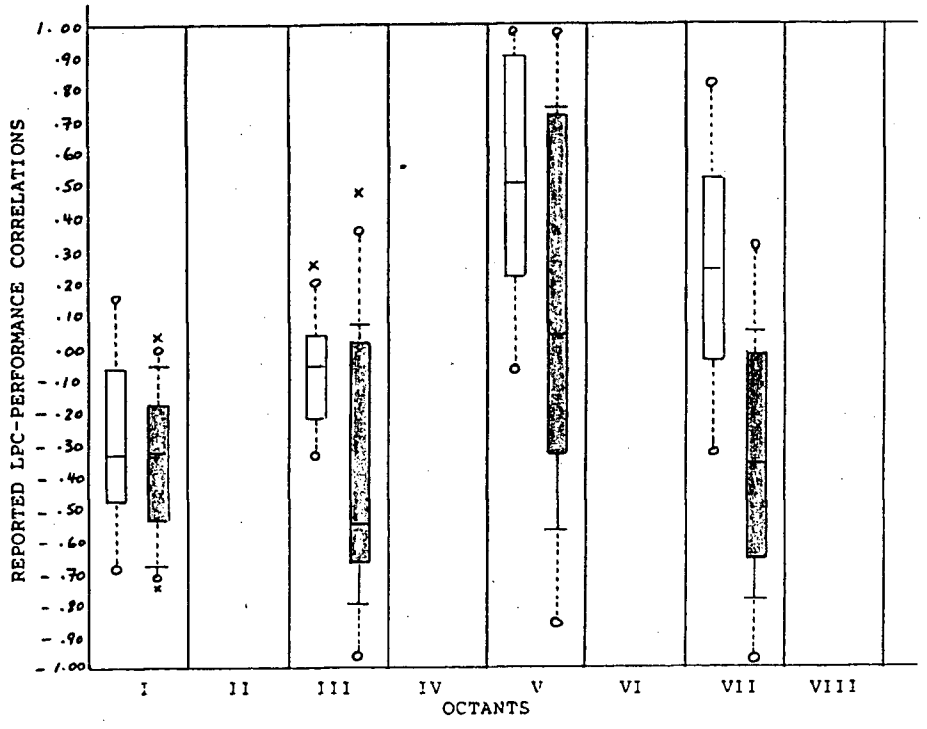


FIGURE F.3: BOX-AND-WHISKER: TASK GROUP TYPES 2 AND 8

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

TGT 2 TGT 8

Coacting and Inferred Interacting Task Groups (TGT 2 and TGT 4)

The distribution of correlations for the two task group types is very similar in Octant VII. In spite of the overlap between the boxes in Octant V, the fact that the majority of the correlations fall above the 0.00 for both distributions suggests more similarities than differences between TGT's 2 and 4. However, the groups may be different in Octant I because approximately half the TGT 4 correlations are above 0.00 while all the TGT 2 correlations are below 0.00. Although the distinction is not as clear, this same pattern occurs in Octant III.

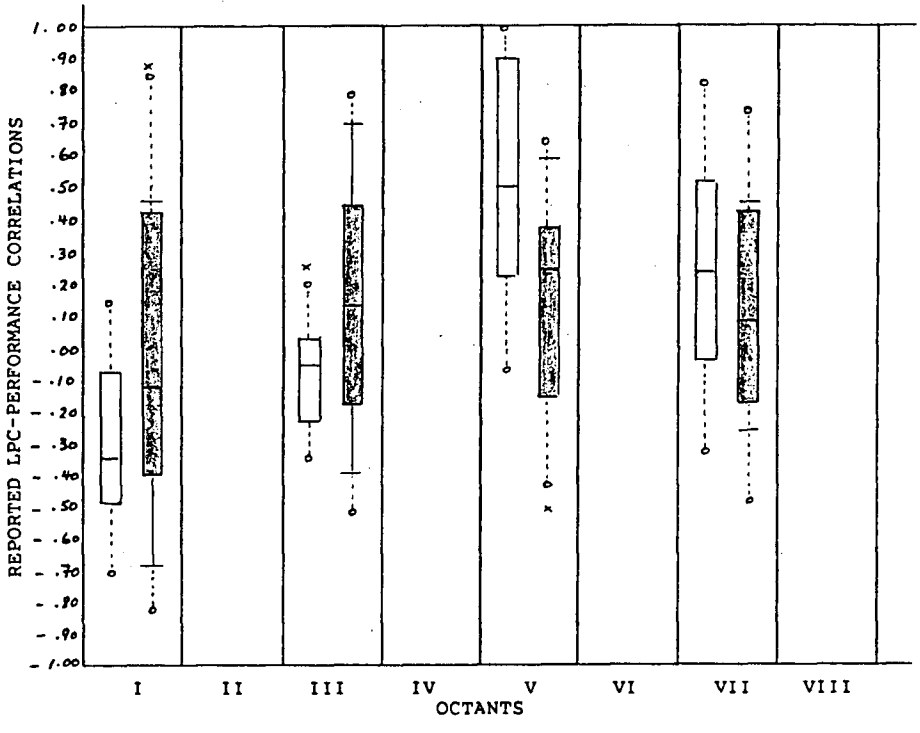


FIGURE F.4: BOX-AND-WHISKER: TASK GROUP TYPES 2 AND 4

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

TGT 2 TGT 4

Inferred Interacting and Nonacting Task Groups (TGT 4 and TGT 8)

The only octant which is more similar than different is Octant V. In Octants I, III, and VII, the top third of the interquartile range for TGT 8 overlaps only with the bottom third of the interquartile range for TGT 4. The limited amount of overlap between these pairs of boxes casts some doubt on the similarity of these two task group types.

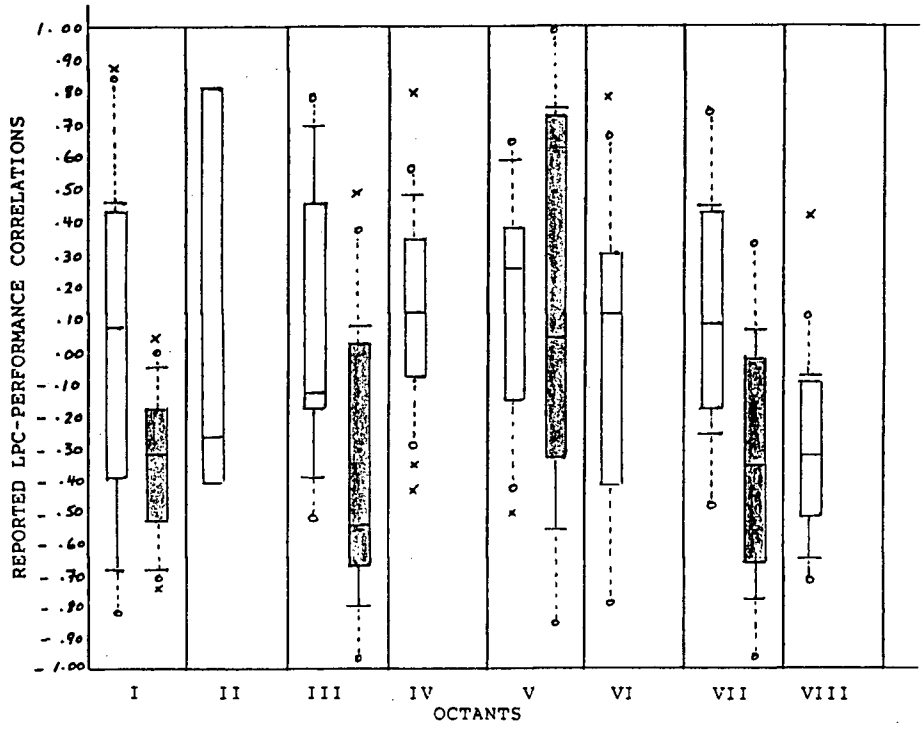


FIGURE F.5: BOX-AND-WHISKER: TASK GROUP TYPES 4 AND 8

——— (whisker): solid vertical line = actual values exist
------ (whisker): broken vertical line = actual values do not exist (i.e., "theoretical" values)
— (adjacent value): solid crossbar shows value closest to but inside inner fence
x (outliers): actual values between inner and outer fences and beyond outer fence
o (inner fence): "theoretical" value at 1.5 interquartile range
• (inner fence): actual value at 1.5 interquartile range
-----o: no solid crossbar (—) indicates the P_{75} = adjacent value and/or P_{25} = adjacent value

TGT 4 TGT 8

TABLE F.4: DESIGNATION OF GAS SCORES AS INDICATING GOOD OR POOR LEADER-MEMBER RELATIONS BASED ON LEADER PERCEPTION

Good LMR		Poor LMR	
69		67	
69		67	
68		----	(c)
68		65	
	(a)	65	
67		65	
66		65	
66		65	
66		64	
66		64	
66		64	
66		64	
66		64	
----	(b)	64	
65		63	
65		63	
65		63	
65		62	
64		61	
64		60	
63		60	
62		50	
61			
58		44	
55			
54			
45			
n=25 Samples (from 19 Studies)		n=22 Samples (from 16 Studies)	

(a) if the item mean of 6.6 is used, then 80% (20/25) of the scores designating good leader-member relations are below the normative mean (6.6 raised to 67)

(b) if the normative mean of 66 is used, then 50% (13/25) of the scores designating good leader-member relations are below the normative value

(c) if the normative mean of 66 is used, then all but two scores designating poor leader-member relations are below the normative value

APPENDIX G

ALLOCATION OF PRIMARY STUDIES TO DESCRIPTOR CATEGORIES

ALLOCATION OF PRIMARY STUDIES TO DESCRIPTOR CATEGORIES

The first three tables in this appendix show the distribution of studies, samples, and octant tests from different perspectives. Table G.1 indicates the frequencies within each study source when the comparability categories were crosstabulated with the study settings.

The data displayed in Table G.2 summarize the frequencies for each study source, comparability category, and study setting separately and for the crosstabulation of study source with study setting.

In Table G.3, the three sources have been combined and the frequencies partitioned by study setting and task group type within each comparability category.

Table G.4 focusses only on the frequency of octant tests. The data, again partitioned by study setting and task group type within each comparability category, show the frequency with which each octant has been tested in the studies which comprise the data base for this meta-analysis. The frequencies clearly reveal the effect of partitioning on the cell sizes.

TABLE G.1: DISTRIBUTION OF STUDIES, SAMPLES, AND OCTANT TESTS BY SOURCE OF STUDY

Source	CC	SS	Studies	Samples	Tests
Journal	F	RL	1	2	8
		L	3	3	26
	O	RL	4	4	29
		L	2	2	4
Dissertation	F	RL	5	10	43
		L	1	1	6
	O	RL	15	15	81
		L	5	8	24
Technical Reports	F	RL	0	0	0
		L	0	0	0
	O	RL	1	3	26
		L	1	1	2
Totals	F+O	RL+L	38	49	249

Abbreviations: CC = comparability category; F = Fiedler
SS = study setting; O = other
RL = real-life; L = laboratory

TABLE G.2: FREQUENCY SUMMARIES FOR STUDY SOURCE, COMPARABILITY CATEGORY, AND STUDY SETTING

Source	Studies	Samples	Tests
Journal	10	11	67
Dissertation	26	34	154
Technical Report	2	4	28
Fiedler	10	16	83
Other	28	33	166
Real Life	26	34	187
Laboratory	12	15	62
Journal: Real Life	5	6	37
Journal: Laboratory	5	5	30
Dissertation: Real Life	20	25	124
Dissertation: Laboratory	6	9	30
Technical Report: Real Life	1	3	26
Technical Report: Laboratory	1	1	2

TABLE G.3: FREQUENCIES WITHIN EACH COMPARABILITY CATEGORY BY STUDY SETTING AND TASK GROUP TYPE

CC	SS	TGT	Studies	Samples	Tests	Total Tests by SS
Fiedler[1]	RL	1	5	8	41	51
		2	3	4	10	
	L	4	4	4	32	32
	Subtotals		12[3]	16	83	
Other[2]	RL	1	8	8	58	136
		4	5	5	22	
		2	3	3	9	
		8	4	6	47	
	L	1	1	1	2	30
		4	6	9	20	
		8	1	1	8	
	Subtotals		28	33	166	
Totals			40[3]	49	249	249

Abbreviations: CC = comparability category
SS = study setting
RL = real-life
L = laboratory
TGT = task group type

1. CC: Fiedler: There were no cases in RL 4, RL 8, L 1, L 2, L 8.
2. CC: Other: There were no cases in L 2.
3. Studies: n=12 and n=40 because two studies, which used two different task group types, were counted twice.

TABLE G.4: FREQUENCY OF OCTANT TESTS PARTITIONED BY COMPARABILITY CATEGORY, STUDY SETTING, AND TASK GROUP TYPE

CC	SS	TGT	Octant I	II	III	IV	V	VI	VII	VIII	Totals RL	L
	RL	1	15	1	7	5	1	1	7	4	41	
	L	4	5	3	4	6	3	3	3	5		32
F[1]	RL	2	4		2		3		1		10	
	RL	Sub-	19	1	9	5	4	1	8	4	51	
	L	total	5	3	4	6	3	3	3	5		32
	RL	1	10	6	5	12	7	2	4	12	58	
	L	1		1		1						2
	RL	4	4		5	3	4		4	2	22	
O[2]	L	4	2	1	5	3	1	3	1	4		20
	RL	2	1		4		1		3		9	
	RL	8	9	2	10	2	10	2	10	2	47	
	L	8	2		2		2		2			8
	RL	Sub-	24	8	24	17	22	4	21	16	136	
	L	total	4	2	7	4	3	3	3	4		30
	RL	Total	43	9	33	22	26	5	29	20	187	
	L	Total	9	5	11	10	6	6	6	9		62
F+O	RL	Rank[3]	1	7	2	5	4	8	3	6		
	L	Rank	3	8	1	2	5	5	5	3		
	RL+L	Rank	1	7	2	4	4	8	3	6		

Abbreviations: CC = comparability category; F = Fiedler
SS = study setting; O = other
RL = real-life; L = laboratory
TGT = task group type

1. CC: Fiedler: There were no cases in RL 4, RL 8, L 1, L 2, L 8.
2. CC: Other: There were no cases in L 2.
3. Rank: the rank-order testing frequency for the eight octants.

APPENDIX H

CORRELATION LISTINGS

TABLE H.1: REPORTED "rho+r" CORRELATIONS FROM ALL OCTANT STUDIES

I	II	III	IV	V	VI	VII	VIII
-0.64	-0.10*	-0.29	-0.05*	0.21	-0.24	0.62	0.38*
-0.51	-0.41	0.60	-0.01*	0.25	-0.39	0.30	-0.10*
-0.59*	0.18	-0.80	0.14*	-0.52	-0.43	-0.30	0.43*
-0.34*	-0.32	0.10*	0.24*	0.28	0.13	-0.32*	0.09*
0.06*	0.81	-0.10*	0.05*	0.90	0.30	0.32*	-0.33
0.14*	0.25	-0.05*	0.33	0.49	0.10	-0.36*	0.44
-0.44*	0.29	-0.18*	-0.08	0.60	0.07	-0.14*	-0.33
-0.92*	-0.02	0.16*	0.35*	0.38	0.11	0.43	-0.10
0.57*	0.31	0.46	0.34	0.31	-0.65	0.45	-0.51
-0.01*	-0.50	0.02	0.48	0.50	-0.20*	0.08	0.27*
-0.13*	-0.60	0.10	-0.33*	0.42	0.79*	0.30	0.16*
-0.43*	0.14*	-0.10*	-0.42*	0.05*		-0.25	0.11*
-0.04*	-0.55*	0.16	-0.07*	0.24		0.17	0.07*
-0.32*	0.11*	0.70	0.16*	-0.33		-0.32	0.15*
0.47		-0.19*	-0.34*	-0.42		0.30	0.36*
-0.13		0.03*	-0.63*	-0.16		-0.19*	0.12*
-0.43		-0.22*	-0.40*	-0.40		0.18*	0.17*
-0.10*		0.14*	-0.60*	0.04*		0.34*	0.42*
-0.69		-0.40	-0.36*	-0.15		-0.04*	0.51*
-0.48		0.70	-0.16*	-0.17		0.01	-0.28*
-0.06		-0.41	-0.45*	0.75*		0.35	-0.38*
-0.36*		0.35	-0.08*	-0.55*		-0.04	-0.40
-0.22		0.37*	-0.36	0.47*		0.07	-0.59
-0.12		0.07*	0.26	0.67*		-0.09	-0.09
0.89		-0.05*	0.12*	0.73*		-0.66*	0.31
-0.40*		0.27	0.80	0.22*		-0.15*	-0.09
0.31		0.08	0.07	0.87		-0.01*	0.23*
0.33*		-0.25	0.09	0.80*		-0.53*	-0.65*
0.06*		-0.24	0.20	-0.73*		-0.37*	-0.80*
0.04		0.57	0.13*	0.32		-0.75*	
0.29*		-0.28	0.62	-0.36		-0.70*	
-0.05		0.50*	0.40*	-0.43		-0.03*	
-0.17		0.03				-0.47*	
-0.17		-0.11				0.53	
-0.15		-0.78*				-0.21*	
0.44		-0.80*					
-0.18		-0.42*					
0.06		-0.57*					
-0.02		-0.56*					
-0.21		-0.66*					
-0.47*		-0.21					
-0.31*		-0.69*					
-0.50*		0.00*					

(cont'd p. 2)

* r

I	III (cont'd)
-0.52*	0.05
-0.71*	
-0.47	
-0.32	
-0.66*	
0.33	
0.24	
0.67	
0.10	

Descriptive Statistics

	I	II	III	IV	V	VI	VII	VIII
Mean	-0.14	-0.03	-0.07	0.01	0.17	-0.04	-0.04	-0.02
Median	-0.17	-0.01	-0.05	0.06	0.25	0.07	-0.04	0.09
Mode	0.06	-0.60	-0.80	-0.36	-0.73	-0.65	0.30	-0.33
SD	0.38	0.41	0.40	0.35	0.47	0.40	0.36	0.37
SE	0.05	0.11	0.06	0.06	0.08	0.12	0.06	0.07
Variance	0.15	0.17	0.16	0.12	0.27	0.16	0.13	0.13
Kurtosis	0.04	-0.31	-0.43	-0.44	-1.06	0.60	-0.73	-0.76
Skewness	0.48	0.25	-0.00	0.05	-0.22	0.51	-0.17	-0.49
Range	1.81	1.41	1.50	1.43	1.63	1.44	1.37	1.31
Minimum	-0.92	-0.60	-0.80	-0.63	-0.73	-0.65	-0.75	-0.80
Maximum	0.89	0.81	0.70	0.80	0.90	0.79	0.62	0.51
# Cases	52	14	44	32	32	11	35	29
+	16	7	21	17	21	6	15	16
-	36	7	23	15	11	5	20	13

TABLE H.2: REPORTED "r+rho" CORRELATIONS FROM ALL OCTANT STUDIES

I	II	III	IV	V	VI	VII	VIII
-0.64	-0.10*	-0.29	-0.05*	0.21	-0.24	0.62	0.38*
-0.51	-0.41	0.60	-0.01*	0.25	-0.39	0.30	-0.10*
-0.59*	0.18	-0.80	0.14*	-0.52	-0.43	-0.30	0.43*
-0.34*	-0.32	0.10*	0.24*	0.28	0.13	-0.32*	0.09*
0.06*	-0.59*	-0.10*	0.05*	0.90	0.30	0.32*	-0.33
0.14*	0.81	-0.05*	0.33	0.49	0.10	-0.36*	0.44
-0.44*	0.25	-0.18*	-0.08	0.60	0.07	-0.14*	-0.33
-0.92*	0.29	0.16*	0.35	0.38	0.11	0.43	-0.10
0.57*	-0.02	0.46	0.34	0.31	-0.60*	0.45	-0.51
-0.01*	0.31	0.02	0.48	0.04*	-0.20*	0.08	0.27*
-0.13*	-0.10	0.10	-0.33*	-0.26*	0.79*	0.30	0.16*
-0.43*	0.14*	-0.10*	-0.42*	0.50		-0.25	0.11*
-0.04*	-0.55*	-0.33*	-0.07*	0.42		0.17	0.07*
-0.32*	0.11*	0.16	0.16*	0.05*		-0.07*	0.15*
0.47		0.70	-0.34*	0.24		-0.13*	0.36*
-0.13		-0.19*	-0.63*	-0.20*		-0.32	0.12*
-0.43		0.03*	-0.40*	-0.34*		0.30	0.17*
-0.10		-0.22*	-0.60*	0.04*		-0.19*	0.42*
-0.69		0.14*	-0.36*	-0.15		0.18*	0.51*
-0.48		-0.40	-0.16*	-0.17		0.34*	-0.28*
-0.06		0.53*	-0.45*	0.75*		-0.04*	-0.38*
-0.36*		0.70	-0.08*	-0.55*		-0.04	-0.40
-0.35*		-0.23*	-0.36	0.47*		0.07	-0.59
-0.22		-0.41	0.26	0.67*		-0.09	-0.09
-0.12		0.35	0.12*	0.73*		-0.66*	0.31
0.89		0.37*	0.80	0.22*		-0.15*	-0.09
-0.40*		0.07*	0.07	0.87		-0.01*	0.23*
0.46*		-0.05*	0.09	0.80*		-0.53*	-0.65*
-0.10*		0.27	0.35*	-0.73*		-0.37*	-0.80*
0.31		0.26*	0.13*	0.32		-0.75*	
0.33		0.31*	0.62	-0.36		-0.70*	
0.06		0.50*	0.40*	-0.43		-0.03*	
0.04		0.03				-0.47*	
0.29*		-0.11				0.53	
-0.05		-0.78*				-0.21*	
-0.14*		-0.80*					
-0.09*		-0.42*					
0.06*		-0.57*					
-0.02		-0.56*					
-0.21		-0.66*					
-0.47*		-0.21					
-0.31*		-0.69*					

(cont'd p. 2)

* r

I	III (cont'd)
-0.50*	0.00*
-0.52*	0.05
-0.71*	
-0.47	
-0.32	
-0.66*	
0.33*	
0.24	
0.67	
0.10	

Descriptive Statistics

	I	II	III	IV	V	VI	VII	VIII
Mean	-0.14	0.00	-0.05	0.02	0.18	-0.03	-0.06	-0.02
Median	-0.13	-0.01	-0.05	0.06	0.25	0.07	-0.07	0.09
Mode	0.06	-0.10	-0.80	-0.36	0.04	-0.60	0.30	-0.33
SD	0.39	0.38	0.40	0.36	0.45	0.39	0.35	0.36
SE	0.05	0.10	0.06	0.06	0.08	0.12	0.06	0.07
Variance	0.15	0.15	0.16	0.13	0.20	0.15	0.13	0.13
Kurtosis	-0.00	0.17	-0.55	-0.53	-0.86	0.62	-0.61	-0.75
Skewness	0.48	0.22	-0.10	0.03	-0.26	0.61	-0.04	-0.49
Range	1.81	1.40	1.50	1.43	1.63	1.39	1.37	1.31
Minimum	-0.92	-0.59	-0.80	-0.63	-0.73	-0.60	-0.75	-0.80
Maximum	0.89	0.81	0.70	0.80	0.90	0.79	0.62	0.51
# Cases	52	14	44	32	32	11	35	29
+	16	7	22	17	22	6	13	16
-	36	7	22	15	10	5	22	13

TABLE H.3: REPORTED "rho" CORRELATIONS FROM ALL OCTANT STUDIES

I	II	III	IV	V	VI	VII	VIII
-0.64	-0.10	-0.29	0.33	0.21	-0.24	0.62	-0.33
-0.51	-0.41	0.60	-0.08	0.25	-0.39	0.30	0.44
0.47	0.18	-0.80	0.35	-0.52	-0.43	-0.30	-0.33
-0.13	-0.32	0.46	0.34	0.28	0.13	0.43	-0.10
-0.43	0.81	0.02	0.48	0.90	0.30	0.45	-0.51
-0.10	0.25	0.10	-0.36	0.49	0.10	0.08	-0.40
-0.69	0.29	0.16	0.26	0.60	0.07	0.30	-0.59
-0.48	-0.02	0.70	0.80	0.38	0.11	-0.25	-0.09
-0.06	0.31	-0.40	0.07	0.31	-0.65	0.17	0.31
-0.22	-0.50	0.70	0.09	0.50		-0.32	-0.09
-0.12	-0.60	-0.41	0.20	0.42		0.30	
0.89		0.35	0.62	0.24		-0.04	
0.31		0.27		-0.33		0.01	
0.33		0.08		-0.42		0.35	
0.06		-0.25		-0.16		0.07	
0.04		-0.24		-0.40		-0.09	
-0.05		0.57		-0.15		0.53	
-0.17		-0.28		-0.17			
-0.17		0.03		0.87			
-0.15		-0.11		0.32			
0.44		-0.21		-0.36			
-0.18		0.05		-0.43			
-0.02							
-0.21							
-0.47							
-0.32							
0.24							
0.67							
0.10							

Descriptive Statistics

	I	II	III	IV	V	VI	VII	VIII
Mean	-0.05	-0.01	0.05	0.26	0.13	-0.11	0.15	-0.17
Median	-0.12	-0.02	0.04	0.27	0.25	0.07	0.17	-0.33
Mode	-0.17	-0.60	0.70	-0.36	-0.52	-0.65	0.30	-0.33
SD	0.38	0.43	0.40	0.31	0.43	0.32	0.29	0.34
SE	0.07	0.13	0.09	0.09	0.09	0.11	0.07	0.11
Variance	0.14	0.18	0.16	0.10	0.19	0.11	0.08	0.11
Kurtosis	0.29	-0.35	-0.49	0.52	-1.08	-1.21	-0.99	-0.24
Skewness	0.60	0.34	0.01	-0.25	0.06	-0.49	-0.21	0.74
Range	1.58	1.41	1.50	1.16	1.42	0.95	0.94	1.03
Minimum	-0.69	-0.60	-0.80	-0.36	-0.52	-0.65	-0.32	-0.59
Maximum	0.89	0.81	0.70	0.80	0.90	0.30	0.62	0.44
# Cases	29	11	22	12	22	9	17	10
+	10	5	13	10	13	5	12	2
-	19	6	9	2	9	4	5	8

TABLE H.4: REPORTED "r" CORRELATIONS FROM ALL OCTANT STUDIES

I	II	III	IV	V	VI	VII	VIII
-0.59	-0.59	0.10	-0.05	0.04	-0.60	-0.32	0.38
-0.34	-0.10	-0.10	-0.01	-0.26	-0.20	0.32	-0.10
0.06	0.14	-0.05	0.14	0.05	0.79	-0.36	0.43
0.14	-0.55	-0.18	0.24	-0.20		-0.14	0.09
-0.44	0.11	0.16	0.05	-0.34		-0.07	0.27
-0.92		-0.10	-0.33	0.04		-0.13	0.16
0.57		-0.33	-0.42	0.75		-0.19	0.11
-0.01		-0.19	-0.07	-0.55		0.18	0.07
-0.13		0.03	0.16	0.47		0.34	0.15
-0.43		-0.22	-0.34	0.67		-0.04	0.36
-0.04		0.14	-0.63	0.73		-0.66	0.12
-0.32		0.53	-0.40	0.22		-0.15	0.17
-0.36		-0.23	-0.60	0.80		-0.01	0.42
-0.35		0.37	-0.36	-0.73		-0.53	0.51
-0.40		0.07	-0.16			-0.37	-0.28
0.46		-0.05	-0.45			-0.75	-0.38
-0.10		0.26	-0.08			-0.70	0.23
0.29		0.31	0.12			-0.03	-0.65
-0.14		0.50	0.35			-0.47	-0.80
-0.09		-0.78	0.13			-0.21	
0.06		-0.80	0.40				
-0.47		-0.42					
-0.31		-0.57					
-0.50		-0.56					
-0.52		-0.66					
-0.71		-0.69					
-0.66		0.0					
0.33							

Descriptive Statistics

	I	II	III	IV	V	VI	VII	VIII
Mean	-0.21	-0.20	-0.13	-0.11	0.12	-0.00	-0.21	0.07
Median	-0.32	-0.10	-0.10	-0.07	0.05	-0.20	-0.19	0.15
Mode	0.06	-0.59	-0.10	-0.63	0.04	-0.60	-0.75	-0.80
SD	0.36	0.35	0.38	0.30	0.51	0.72	0.31	0.36
SE	0.07	0.16	0.07	0.07	0.14	0.41	0.07	0.08
Variance	0.13	0.12	0.14	0.09	0.26	0.51	0.10	0.13
Kurtosis	-0.25	-3.04	-0.72	-1.05	-1.17		-0.47	0.86
Skewness	0.36	-0.33	-0.18	-0.10	-0.07	1.14	-0.01	-1.19
Range	1.49	0.73	1.33	1.03	1.53	1.39	1.09	1.31
Minimum	-0.92	-0.59	-0.80	-0.63	-0.73	-0.60	-0.75	-0.80
Maximum	0.57	0.14	0.53	0.40	0.80	0.79	0.34	0.51
# Cases	28	5	27	21	14	3	20	19
+	7	2	11	8	9	1	3	14
-	21	3	16	13	5	2	17	5

LIST OF AVAILABLE CORRELATION LISTINGS

The four listings of reported LPC-performance correlations, grouped according to several different variables, are available. In addition to the correlations, each listing contains a set of descriptive statistics. Photocopies may be obtained by writing to the address shown at the end of the list.

Single Variable Groupings of Correlations

Variable(s)	Reference Number
Comparability Category (CC:F, CC:O)	4(24):63
Study Setting (Real Life, Lab)	4(25):64
Stated Interacting Task Groups (TGT 1)	4(26):65
Inferred Interacting Task Groups (TGT 4)	4(27):66
Coacting Task Groups (TGT 2)	4(28):67
Nonacting Task Groups (TGT 8)	4(29):68
TGT 1 by Real Life	4(30):69
TGT 4 by Real Life	4(31):70
TGT 2 by Real Life	5(1):71
TGT 8 by Real Life	5(2):72
TGT 4 by Laboratory	5(3):73
TGT 1+4 by Laboratory	5(4):74
TGT 1+4+8 by Laboratory	5(5):75
CC:Fiedler by Real Life	5(6):76
CC:Fiedler by Laboratory	5(7):77
CC:Other by Real Life	5(8):78
CC:Other by Laboratory	5(9):79
Leadership Effectiveness by Group Performance	5(10):80
Leadership Effectiveness by Leader Performance	5(11):81
Groups as Unit of Analysis	5(12):82
Leaders as Unit of Analysis	5(13):83
Group Performance by Groups	5(14):84
Leader Performance by Leaders	5(15):85
Military Studies	5(16):86
University/ROTC Studies	5(17):87
School-based Studies	5(18):88
Business and Industrial Studies	5(19):89
Hospital Studies	5(20):90
Journals	5(21):91
Technical Reports	5(22):92
Dissertations	5(23):93
Psychology	5(24):94
Education + Educational Administration	5(25):95
Education other than Educational Administration	5(26):96
Educational Administration only	5(27):97
Commerce and Business Administration	5(28):98
Researchers Associated with Fiedler	5(29):99
Independent Researchers	5(30):100

The University of British Columbia
Faculty of Education
Dep't of Administrative, Adult, and Higher Education
6298 Biological Sciences Road
Vancouver, B.C.
V6T 1Z5

APPENDIX I

CODING INSTRUMENT

CODING INSTRUMENT

Date Coded: _____

I DOCUMENT INFORMATION (all documents)

(1)* 1 Document ID Number _____ → [1 2 3]

(1) 2 Card Number 1 _____ → [4]

(1) 3 Year _____ → [5 6]

(1) 4 Author(s) _____ → [7]

_____

_____

_____

(1) 5 Source _____ → [8]

(1) 6 Type _____ → [9 10]

(1) 7 Journal Name _____ → [11 12]

(2) 8 Author Background _____ → [13]

(3) 9 Fiedler Associate(s) _____ → [14]

(3) 10 Country _____ → [15]

II PURPOSE(S) OF STUDY (11,17,18 only)

(3) 11.1 _____

_____

_____

_____

_____

_____

_____

III HYPOTHESES/RESEARCH QUESTIONS (11,17,18 only)

(3) 12.1 _____

_____

_____

_____

_____

_____

*bracketed numbers refer to page in Code Book

III (cont'd)

IV STUDY CHARACTERISTICS (11,17,18 only)

A RESEARCH DESIGN

(4) 13. class of study ____ [16]

(4) 14. study setting ____ [17]

(4) 15. organizational setting ____ [18][19]

(5) 16. task group type ____ [20]

(5) 17. study design ____ [21]

(5) 18. study duration ____ [22][23]

(5) 19. octant(s) focus ____ [24][25]

(6) 20. VARIABLE(S) CONSIDERED IN THIS STUDY

20.01 leadership effectiveness (LE) ____ [26]

20.02 group performance (GP) ____ [27]

20.03 leadership style (LPC score) ____ [28]

20.04 leader-member relations (LMR) ____ [29]

20.05 task structure (TS) ____ [30]

20.06 position power (PP) ____ [31]

20.07 experience ____ [32]

20.08 training ____ [33]

20.09 experience and training → [34]

20.10 intelligence → [35]

20.11 motivation (not LPC) → [36]

20.12 intelligence and motivation (not LPC) → [37]

20.13 some other combination of 07 - 12 → [38]

LPC INTERPRETATION

20.14 motivation (needs disposition: Fiedler) → [39]

20.15 goal hierarchy (Fiedler) → [40]

20.16 differentiation ability (Foa) → [41]

20.17 cognitive complexity (Mitchell) → [42]

20.18 value attitude (Rice) → [43]

20.19 some combination of 14 - 18 → [44]

20.20 some other interpretation of LPC → [45]

20.21 stress → [46]

20.22 cultural homogeneity/heterogeneity → [47]

20.23 sex differences → [48]

20.24 emergent leadership → [49]

20.25 organizational complexity → [50]

20.26 teaching effectiveness → [51]

20.27 school effectiveness → [52]

20.28 leader behaviour → [53]

20.29 job satisfaction → [54]

20.30 task performance (only part of a job) → [55]

20.31 job performance (overall work) [56]

20.32 group atmosphere (other than GAS/LMR) [57]

20.33 other variable(s) used [58]

(7) 21. DEPENDENT VARIABLE(S)

21.01 leadership effectiveness (LE) [59]

21.02 group performance (GP) [60]

21.03 leadership style (LPC) [61]

21.04 leader - member relations (LMR) [62]

21.05 task structure (TS) [63]

21.06 position power (PP) [64]

21.07 experience [65]

21.08 training [66]

21.09 experience and training [67]

21.10 intelligence [68]

21.11 motivation (not LPC) [69]

21.12 intelligence and motivation (not LPC) [70]

21.13 some other combination of 07 - 12 [71]

21.14 LPC interpretation [72]

21.15 stress [73]

21.16 cultural homogeneity/heterogeneity [74]

21.17 sex differences [75]

21.18 emergent leadership [76]

21.19 organizational complexity [77]

21.20 organizational change ____ → ____ 78

21.21 teaching effectiveness ____ → ____ 79

(7) Document ID Number ____ START NEW CARD → 1 2 3

Card Number 2 ____ → 2 4

21.22 school effectiveness ____ → ____ 5

21.23 leader behaviour ____ → ____ 6

21.24 job satisfaction ____ → ____ 7

21.25 task performance (part of job) ____ → ____ 8

21.26 job performance (overall work) ____ → ____ 9

21.27 group atmosphere (other than GAS/LMR) ____ → ____ 10

21.28 other variable(s) so designated ____ → ____ 11

(8) 22 INDEPENDENT VARIABLE(S)

22.01 leadership effectiveness (LE) ____ → ____ 12

22.02 group performance (GP) ____ → ____ 13

22.03 leadership style (LPC) ____ → ____ 14

22.04 leader - member relations (LMR) ____ → ____ 15

22.05 task structure (TS) ____ → ____ 16

22.06 position power (PP) ____ → ____ 17

22.07 experience ____ → ____ 18

22.08 training ____ → ____ 19

22.09 experience and training ____ → ____ 20

22.10 intelligence ____ → ____ 21

22.11 motivation (not LPC) ____ → ____ 22

22.12 intelligence and motivation (not LPC) → 23

22.13 some other combination of 07 - 12 → 24

22.14 LPC interpretation → 25

22.15 stress → 26

22.16 cultural homogeneity/heterogeneity → 27

22.17 sex differences → 28

22.18 emergent leadership → 29

22.19 organizational complexity → 30

22.20 organizational change → 31

22.21 teaching effectiveness → 32

22.22 school effectiveness → 33

22.23 leader behaviour → 34

22.24 job satisfaction → 35

22.25 task performance (part of a job) → 36

22.26 job performance (overall work) → 37

22.27 group atmosphere (other than GAS/LMR) → 38

22.28 other variable(s) so designated → 39

(9) 23 MODERATING VARIABLE(S)

23.01 experience → 40

23.02 training → 41

23.03 experience and training → 42

23.04 leader intelligence → 43

23.05 group member intelligence → 44

23.06 leader motivation (not LPC) → 45

23.07 group member motivation (not LPC) → 46

23.08 some other combination of 01 - 07 → 47

23.09 stress → 48

23.10 cultural difference(s) → 49

23.11 sex differences → 50

23.12 job satisfaction of leader → 51

23.13 job satisfaction of group members → 52

23.14 job satisfaction of both l and gm's → 53

23.15 organizational complexity → 54

23.16 organizational change → 55

23.17 other variable(s) so designated → 56

(9) 24 Sample selection: leaders → 57 58

(10) 25 Sample selection: group members → 59 60

(10) 26 Sample selection: S's equal status → 61

(11) 27 Hierarchical distinction → 62

IV B. INSTRUMENTATION-MEASUREMENT

(11) 28 LPC scale → 63

(11) 29 LMR → 64 65

(12) 30 TS → 66 67

(12) 31 PP → 68

(12) 32 SC/SF → 69

(13) 33 OTHER INSTRUMENTS

33.1 Leader behaviour _____ → 70

33.2 Additional instruments _____ → 71

IV C. RESEARCH PROCEDURES

(13) 34 TREATMENT OF LPC

34.01 Completed by _____ → 72

34.02 Derivation of individual scores _____ → 73

34.03 Division of sample scores _____ → 74 75

34.04 Sample $\bar{X}$ (PV) _____ → 76

34.05 Sample $\bar{X}$ (RD or 0) _____ → 77 78 79 80

(15) Document ID Number _____ START NEW CARD → 1 2 3

Card Number 3 _____ → 3 4

34.06 Sample SD (PV) _____ → 5

34.07 Sample SD (RD or 0) _____ → 6 7 8 9

34.08 High LPC designation (PV) _____ → 10

34.09 High LPC designation (RD or 0) _____ → 11 12 13 14

34.10 Low LPC designation (PV) _____ → 15

34.11 Low LPC designation (RD or 0) _____ → 16 17 18 19

34.12 Middle LPC designation (PV) _____ → 20

34.13 Middle LPC designation (RD or 0) _____ → 21 22 23 24

34.14 LPC test - retest reliability (PV) _____ → 25

34.15 LPC test - retest reliability (RD or 0) _____ → 26 27 28

34.16 Other LPC data reported _____ → 29

(18) 35 TREATMENT OF GAS (i.e. LMR)

35.01 Completed by ____ 30

35.02 Division of sample scores ____ 31

35.03 Sample $\bar{X}$ (PV) ____ 32

35.04 Sample $\bar{X}$ (RD or 0) ____ 33 34 35 36

35.05 Sample SD (PV) ____ 37

35.06 Sample SD (RD or 0) ____ 38 39 40 41

35.07 Good LMR designation (PV) ____ 42

35.08 Good LMR designation (RD or 0) ____ 43 44 45 46

35.09 Poor LMR designation (PV) ____ 47

35.10 Poor LMR designaton (RD or 0) ____ 48 49 50 51

35.11 Other GAS data reported ____ 52

35.12 Other LMR data reported ____ 53

(20) 36. "EMPIRICAL" TREATMENT OF TS

36.1 Completed by ____ 54

36.2 Division of sample scores ____ 55

(21) 37. "DEFINITIONAL" TREATMENT OF TS

37.1 Basis for defining degree of structure ____ 56

37.2 Degree of structure evaluated by ____ 57

(22) 38. "CONCEPTUAL" TREATMENT OF TS

38.1 Source(s) of information ____ 58 59

(23) 38.2 Conceptual equivalents

38.2.1 Experience

leaders ——— [60]

group members ——— [61]

leaders/group members ——— [62]

all subjects = status ——— [63]

38.2.2 Training

leaders ——— [64]

group members ——— [65]

leaders/group members ——— [66]

all subjects = status ——— [67]

38.2.3 Experience and Training

leaders ——— [68]

group members ——— [69]

leaders/group members ——— [70]

all subjects = status ——— [71]

38.2.4 Intelligence

leaders ——— [72]

group members ——— [73]

leaders/group members ——— [74]

all subjects = status ——— [75]

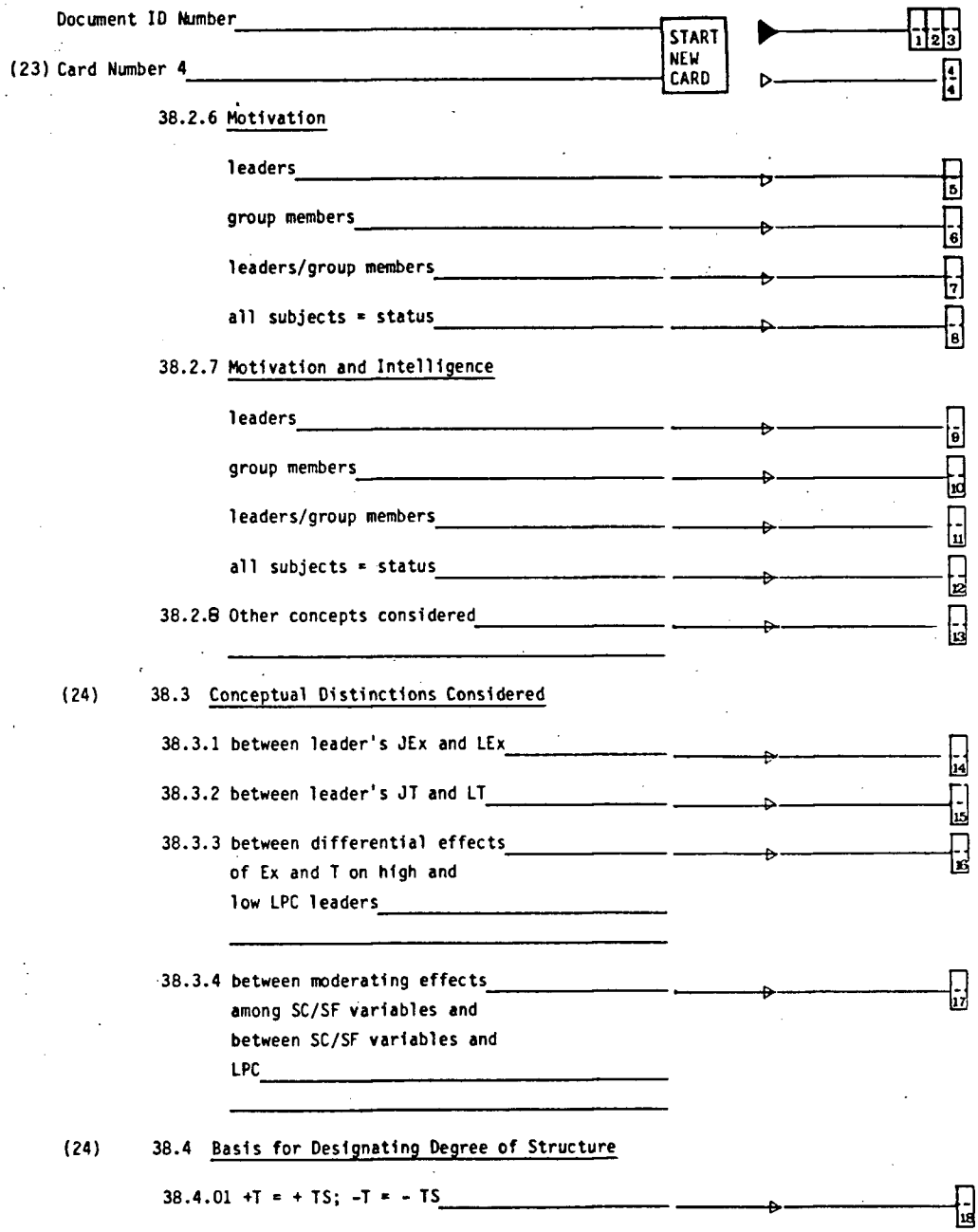
38.2.5 Intelligence and Experience

leaders ——— [76]

group members ——— [77]

leaders/group members ——— [78]

all subjects = status ——— [79]


Document ID Number
START NEW CARD
1 2 3
(23) Card Number 4
4
38.2.6 Motivation
leaders
5
group members
6
leaders/group members
7
all subjects = status
8
38.2.7 Motivation and Intelligence
leaders
9
group members
10
leaders/group members
11
all subjects = status
12
38.2.8 Other concepts considered
13
(24) 38.3 Conceptual Distinctions Considered
38.3.1 between leader's JEx and LEx
14
38.3.2 between leader's JT and LT
15
38.3.3 between differential effects of Ex and T on high and low LPC leaders
16
38.3.4 between moderating effects among SC/SF variables and between SC/SF variables and LPC
17
(24) 38.4 Basis for Designating Degree of Structure
38.4.01 +T = + TS; -T = - TS
18

38.4.02 +JT = + TS; -JT = - TS ——▷ [19]

38.4.03 + LT = + TS; - LT = - TS ——▷ [20]

38.4.04 + JT + (+LT) = + TS; - JT + (-LT) = - TS ——▷ [21]

38.4.05 + Ex = + TS; - Ex = - TS ——▷ [22]

38.4.06 + JEx = + TS; - JEx = - TS ——▷ [23]

38.4.07 + LEx = + TS; - LEx = - TS ——▷ [24]

38.4.08 + JEx + (+LEx) = + TS; - JEx + (- LEx) = - TS ——▷ [25]

38.4.09 + T + (+ Ex) = + TS; - T + (- Ex) = - TS ——▷ [26]

38.4.10 + IQ + (+ Ex) = + TS; - IQ + (+ Ex) = - TS ——▷ [27]

38.4.11 Some other combinations of variables ——▷ [28]

(25) 39. "EMPIRICAL" TREATMENT OF PP

39.1 Completed by ——▷ [29]

39.2 Division of sample scores ——▷ [30]

(26) 40. "DEFINITIONAL" TREATMENT OF PP

40.1 Basis for defining degree of power ——▷ [31]

40.2 Degree of power evaluated by ——▷ [32]

(27) 41. Basis for Determining Degree of SC/SF ——▷ [33]

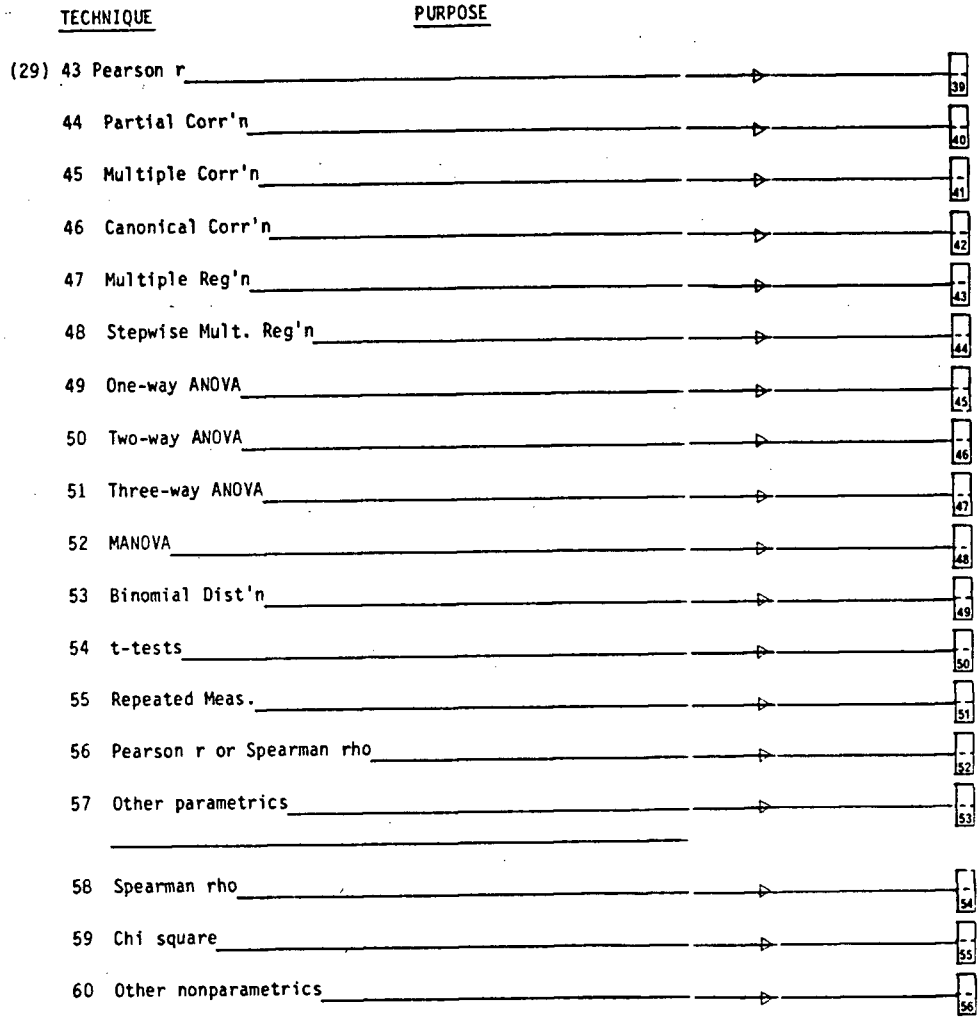
(27) 42. TREATMENT OF LEADERSHIP EFFECTIVENESS (LE)

42.1 Basis for judging LE ——▷ [34]

42.2 Quality of LE rated by ——▷ [35][36]

42.3 Division of sample scores ——▷ [37][38]

IV D. STATISTICAL TECHNIQUES



V SAMPLE CHARACTERISTICS (11,17,18 only)

(30) A. TOTAL SAMPLE

61 Number of subjects

61.1 PV ——— 57

61.2 RD or 0 ——— 58 59 60 61

62 Number of males

62.1 PV ——— 62

62.2 RD or 0 ——— 63 64 65 66

63 Number of females

63.1 PV ——— 67

63.2 RD or 0 ——— 68 69 70

64 Number of groups

64.1 PV ——— 71

64.2 RD or 0 ——— 72 73 74

65 Number of members per group

65.1 PV ——— 75

65.2 RD or 0 ——— 76 77

66 $\bar{X}$ LPC: all subjects

66.1 PV ——— 78

Document ID Number ——— START NEW CARD ▶ 1 2 3

(31) Card Number 5 ——— ▷ 5 4

66.2 RD or 0; Range: ____ to ____ 5 6 7 8

67 $\bar{X}$ LPC: males

67.1 PV ——— 9

67.2 RD or 0; Range: ____ to ____ 10 11 12 13

68 $\bar{X}$ LPC: females

68.1 PV ——— 14

68.2 RD or 0; Range: ___ to 

69 Number of high LPC subjects

69.1 PV 

69.2 RD or 0 

70 Number of high LPC males

70.1 PV 

70.2 RD or 0 

71 Number of high LPC females

71.1 PV 

71.2 RD or 0 

72 Number of middle LPC subjects

72.1 PV 

72.2 RD or 0 

73 Number of middle LPC males

73.1 PV 

73.2 RD or 0 

74 Number of middle LPC females

74.1 PV 

74.2 RD or 0 

75 Number of low LPC subjects

75.1 PV 

75.2 RD or 0 

76 Number of low LPC males

76.1 PV 

76.2 RD or 0 

77 Number of low LPC females

77.1 PV ——— 49

77.2 RD or 0 ——— 50 51

78 $\bar{X}$ Age: all subjects

78.1 PV ——— 52

78.2 RD or 0; Range: ___ to ——— 53 54 55 56

79 $\bar{X}$ IQ: all subjects

79.1 PV ——— 57

79.2 RD or 0; Range: ___ to ——— 58 59 60 61 62

80 SD:IQ all subjects

80.1 PV ——— 63

80.2 RD or 0 ——— 64 65 66 67

81 $\bar{X}$ Years of training: all subjects

81.1 PV ——— 68

81.2 RD or 0; Range: ___ to ——— 69 70 71 72

82 $\bar{X}$ Years of Training: males

82.1 PV ——— 73

82.2 RD or 0; Range ___ to ——— 74 75 76 77

83 $\bar{X}$ Years of training: females

83.1 PV ——— 78

Document ID Number ——— START NEW CARD ——— 1 2 3

(33) Card Number 6 ——— START NEW CARD ——— 6 4

83.2 RD or 0; Range ___ to ——— 5 6 7 8

84 $\bar{X}$ Years of experience: all subjects

84.1 PV _____ 9

84.2 RD or 0; Range ___ to _____ 10 11 12 13

85 $\bar{X}$ Years of Experience: males

85.1 PV _____ 14

85.2 RD or 0; Range ___ to _____ 15 16 17 18

86 $\bar{X}$ Years of experience: females

86.1 PV _____ 19

86.2 RD or 0; Range ___ to _____ 20 21 22 23

(33) 87 Type of subjects _____ 24 25

IV B. LEADER SUBJECTS ONLY

(35) 88 Total number of leaders

88.1 PV _____ 26

88.2 RD or 0 _____ 27 28 29

89 Number of males

89.1 PV _____ 30

89.2 RD or 0 _____ 31 32 33

90 Number of females

90.1 PV _____ 34

90.2 RD or 0 _____ 35 36

91 Number of leaders per octant

91.1 PV _____ 37

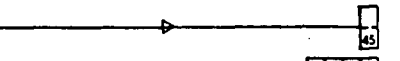
91.2 RD or 0 _____ 38 39

92 $\bar{X}$ LPC score: all leaders

92.1 PV _____ 40

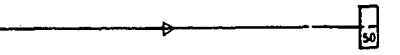
92.2 RD or 0; Range ___ to _____ 41 42 43 44

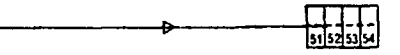
93 $\bar{X}$ LPC score: males

93.1 PV 

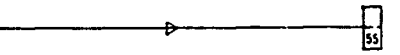
93.2 RD or 0; Range ___ to 

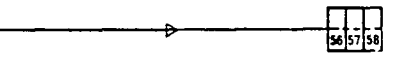
94 $\bar{X}$ LPC score: females

94.1 PV 

94.2 RD or 0; Range ___ to 

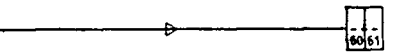
95 Number of high LPC males

95.1 PV 

95.2 RD or 0 

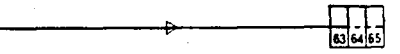
96 Number of middle LPC males

96.1 PV 

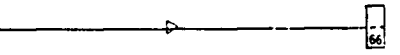
96.2 RD or 0 

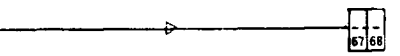
97 Number of low LPC males

97.1 PV 

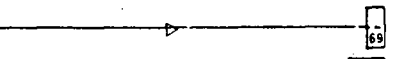
97.2 RD or 0 

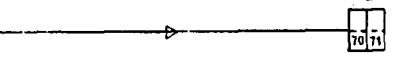
98 Number of high LPC females

98.1 PV 

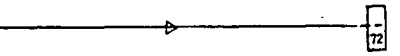
98.2 RD or 0 

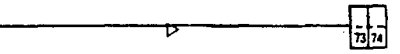
99 Number of middle LPC females

99.1 PV 

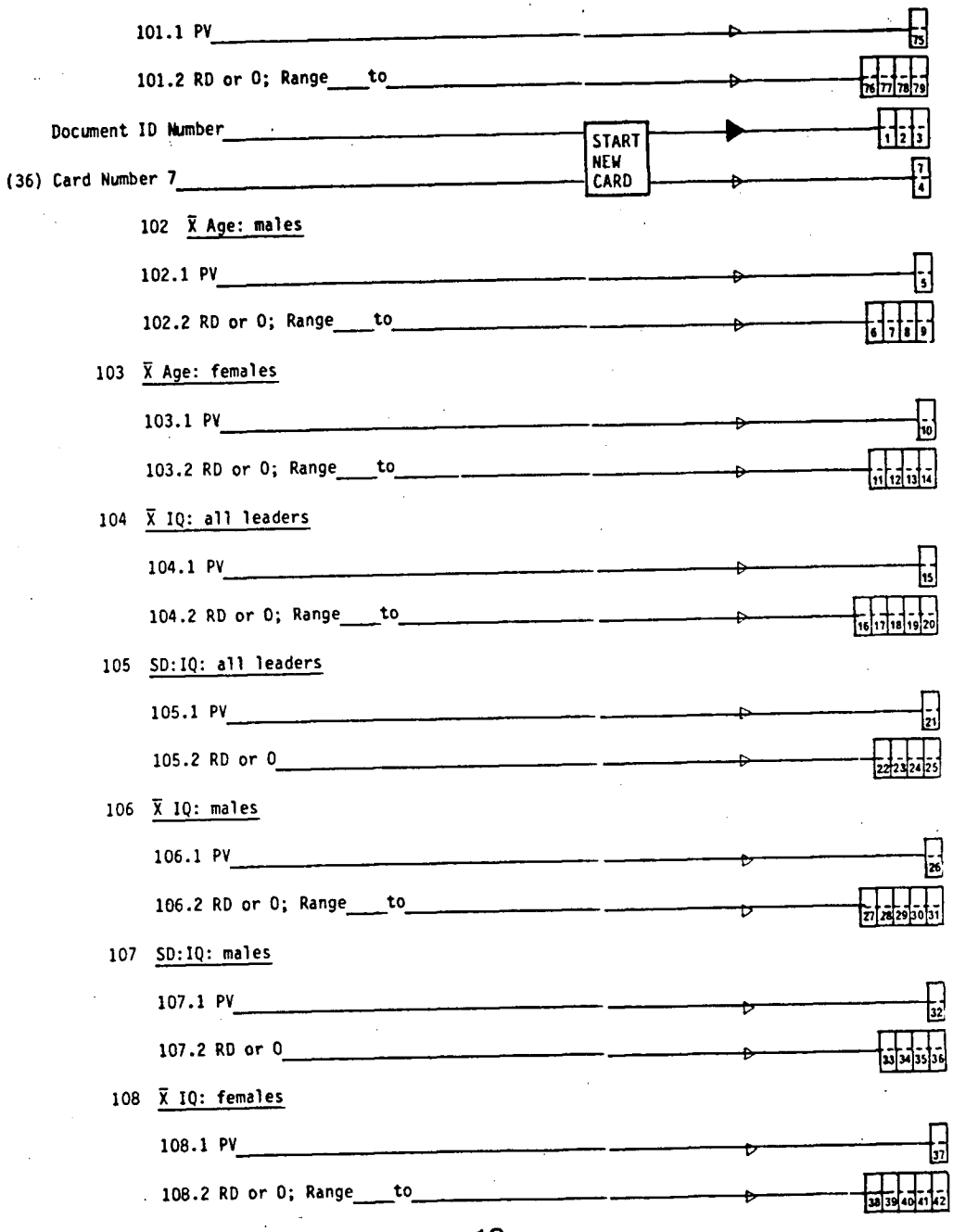
99.2 RD or 0 

100 Number of Low LPC females

100.1 PV 

100.2 RD or 0 

101 $\bar{X}$ Age: all leaders


101.1 PV
75
101.2 RD or 0; Range___to
76 77 78 79
Document ID Number
1 2 3
START NEW CARD
(36) Card Number 7
7 4
102 X̄ Age: males
102.1 PV
5
102.2 RD or 0; Range___to
6 7 8 9
103 X̄ Age: females
103.1 PV
10
103.2 RD or 0; Range___to
11 12 13 14
104 X̄ IQ: all leaders
104.1 PV
15
104.2 RD or 0; Range___to
16 17 18 19 20
105 SD:IQ: all leaders
105.1 PV
21
105.2 RD or 0
22 23 24 25
106 X̄ IQ: males
106.1 PV
26
106.2 RD or 0; Range___to
27 28 29 30 31
107 SD:IQ: males
107.1 PV
32
107.2 RD or 0
33 34 35 36
108 X̄ IQ: females
108.1 PV
37
108.2 RD or 0; Range___to
38 39 40 41 42

109 SD:IQ: females

109.1 PV ____ 43

109.2 RD or 0 ____ 44 45 46 47

110 $\bar{X}$ Years of training: all leaders

110.1 PV ____ 48

110.2 RD or 0; Range ___ to ____ 49 50 51 52

111 $\bar{X}$ Years of training: males

111.1 PV ____ 53

111.2 RD or 0; Range ___ to ____ 54 55 56 57

112 $\bar{X}$ Years of training: females

112.1 PV ____ 58

112.2 RD or 0; Range ___ to ____ 59 60 61 62

113 $\bar{X}$ Years of experience: all leaders

113.1 PV ____ 63

113.2 RD or 0; Range ___ to ____ 64 65 66 67

114 $\bar{X}$ Years of experience: males

114.1 PV ____ 68

114.2 RD or 0; Range ___ to ____ 69 70 71 72

115 $\bar{X}$ Years of experience: females

115.1 PV ____ 73

115.2 RD or 0; Range ___ to ____ 74 75 76 77

(33) 116 Type of leader subjects ____ 78 79

Document ID Number ____ START NEW CARD ____ 1 2 3

(38) Card Number 8 ____ 8 4

V C. GROUP MEMBER SUBJECTS ONLY

(38) 117 Total number of group members

117.1 PV ——— 5

117.2 RD or 0 ——— 6 7 8

118 Number of males

118.1 PV ——— 9

118.2 RD or 0 ——— 10 11 12

119 Number of females

119.1 PV ——— 13

119.2 RD or 0 ——— 14 15

120 Number of members per group

120.1 PV ——— 16

120.2 RD or 0 ——— 17 18

121 Number of members per octant

121.1 PV ——— 19

121.2 RD or 0 ——— 20 21

122 $\bar{X}$ LPC score: all group members

122.1 PV ——— 22

122.2 RD or 0; Range ___ to ——— 23 24 25 26

123 $\bar{X}$ LPC score: males

123.1 PV ——— 27

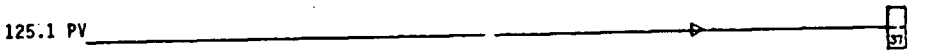
123.2 RD or 0; Range ___ to ——— 28 29 30 31

124 $\bar{X}$ LPC score: females

124.1 PV ——— 32

124.2 RD or 0; Range ___ to ——— 33 34 35 36

125 Number of high LPC males

125.1 PV 

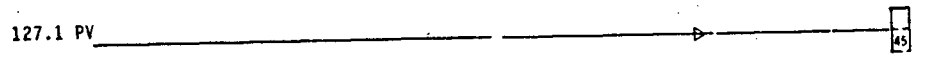
125.2 RD or 0 

126 Number of middle LPC males

126.1 PV 

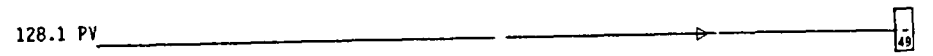
126.2 RD or 0 

127 Number of low LPC males

127.1 PV 

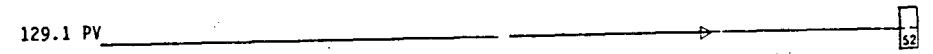
127.2 RD or 0 

128 Number of high LPC females

128.1 PV 

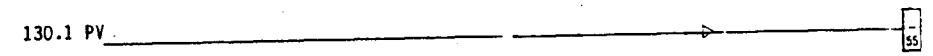
128.2 RD or 0 

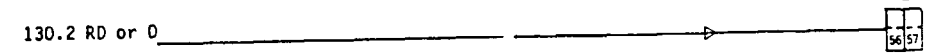
129 Number of middle LPC females

129.1 PV 

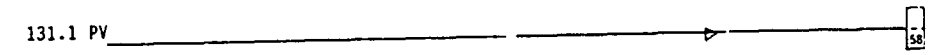
129.2 RD or 0 

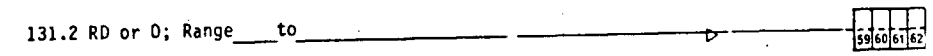
130 Number of low LPC females

130.1 PV 

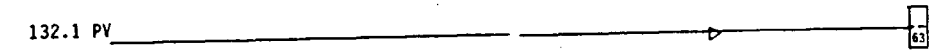
130.2 RD or 0 

131 $\bar{X}$ Age: all group members

131.1 PV 

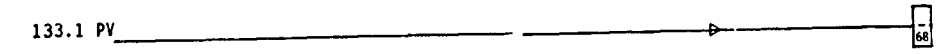
131.2 RD or 0; Range ___ to ___ 

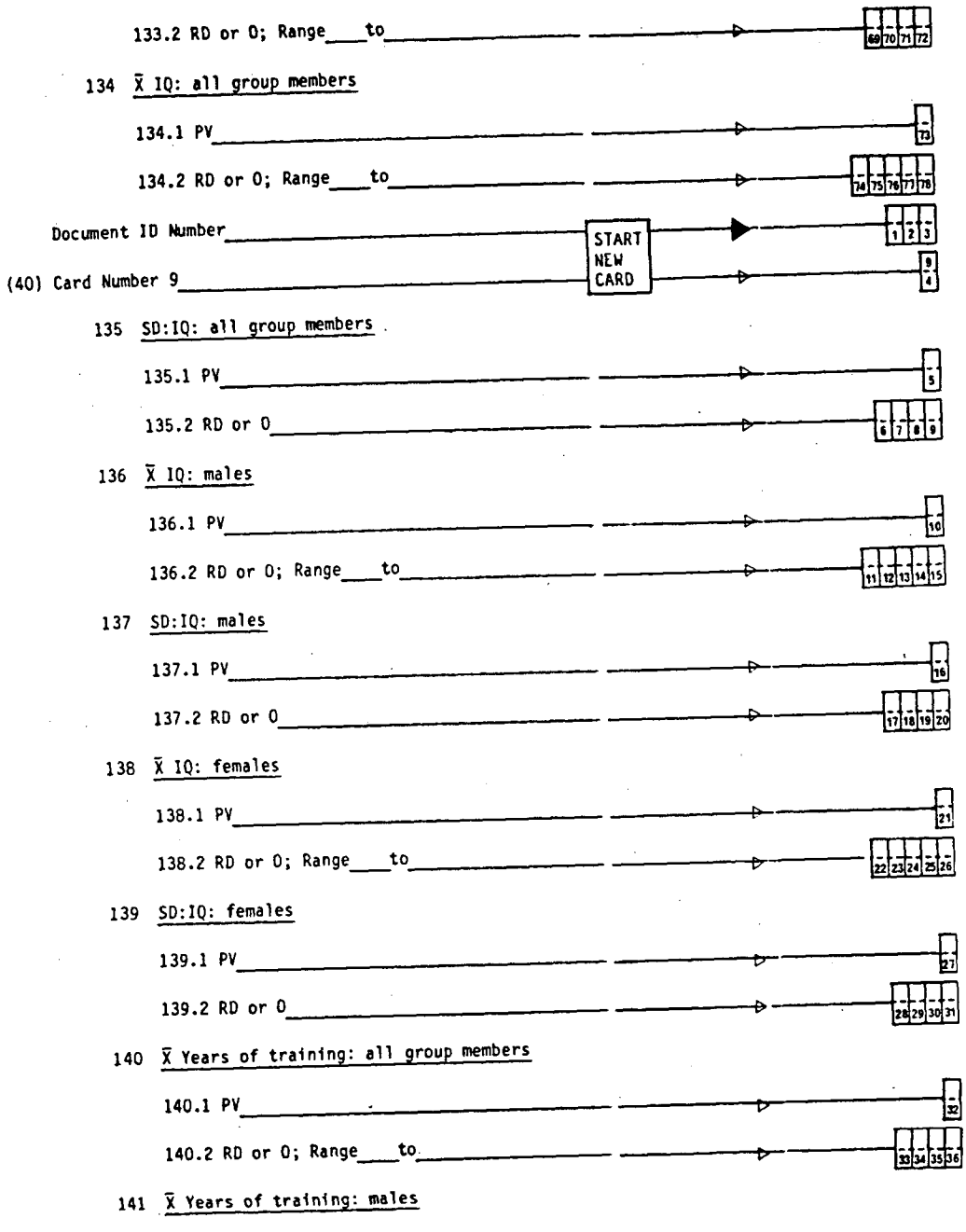
132 $\bar{X}$ Age: males

132.1 PV 

132.2 RD or 0; Range ___ to ___ 

133 $\bar{X}$ Age: females

133.1 PV 


133.2 RD or 0; Range___to
69 70 71 72
134 X̄ IQ: all group members
134.1 PV
73
134.2 RD or 0; Range___to
74 75 76 77 78
Document ID Number
1 2 3
START NEW CARD
(40) Card Number 9
80
4
135 SD:IQ: all group members
135.1 PV
5
135.2 RD or 0
6 7 8 9
136 X̄ IQ: males
136.1 PV
10
136.2 RD or 0; Range___to
11 12 13 14 15
137 SD:IQ: males
137.1 PV
16
137.2 RD or 0
17 18 19 20
138 X̄ IQ: females
138.1 PV
21
138.2 RD or 0; Range___to
22 23 24 25 26
139 SD:IQ: females
139.1 PV
27
139.2 RD or 0
28 29 30 31
140 X̄ Years of training: all group members
140.1 PV
32
140.2 RD or 0; Range___to
33 34 35 36
141 X̄ Years of training: males

141.1 PV ——— 37

141.2 RD or 0; Range ___ to ——— 38 39 40 41

142 $\bar{X}$ Years of training: females

142.1 PV ——— 42

142.2 RD or 0; Range ___ to ——— 43 44 45 46

143 $\bar{X}$ Years of experience: all group members

143.1 PV ——— 47

143.2 RD or 0; Range ___ to ——— 48 49 50 51

144 $\bar{X}$ Years of experience: males

144.1 PV ——— 52

144.2 RD or 0; Range ___ to ——— 53 54 55 56

145 $\bar{X}$ Years of experience: females

145.1 PV ——— 57

145.2 RD or 0; Range ___ to ——— 58 59 60 61

(33) 146 Type of group member subjects ——— 62 63

VII FINDINGS (11,17,18 only) Document ID

A. SUBSTANTIVE B. STATISTICAL

147.1

VII CONCLUSIONS (11,17,18 only) Document ID____

148.1 _____

(43) 149 Extent of Congruency with Model

(43) 150 Extent of Congruency with Instruments

VIII CRITICAL ANALYSIS OF STUDY (11,17,18 only)

(44) A. THREATS TO VALIDITY (TV SCORE)

151 sampling bias absent ____ [66]

152 sample attrition satisfactory ____ [67]

153 outcomes unaffected by testing procedures ____ [68]

154 appropriate statistics ____ [69]

155 appropriate use of F's instruments ____ [70]

156 "effect" of instrument change reported ____ [71]

157 reliability of changed instrument reported ____ [72]

158 validity of changed instrument reported ____ [73]

159 outcomes unaffected by conceptualization of a variable ____ [74]

TV SCORE ($\bar{X}$) → $\Sigma X =$ $n_X =$ $\bar{X}_{TV} =$

(45) B. METHODOLOGICAL ADHERENCE (MA SCORE) Document ID___

160 LMR definition, measurement, assumption ——▷ [75]

161 TS definition, measurement/conception, assumption ——▷ [76]

162 PP definition, measurement, assumption ——▷ [77]

163 SC/SF matched with LPC score ——▷ [78]

164 LE judged on basis of GP ——▷ [79]

165 leadership style and behaviour distinguished ——▷ [80]

ΣY =

MA SCORE (Y) ⟶ n_y =

$\bar{Y}_{MA}$ =

IX FURTHER COMMENTS, OBSERVATIONS, QUESTIONS

166.1 _____