ORGANIZATIONAL BEHAVIOR AND HUMAN PERFORMANCE 9, 339–355 (1973)

The Contingency Model of Leadership Effectiveness: An Evaluation

AHMED SAKR ASHOUR[1]
California State University, Fresno

The contingency model of leadership effectiveness and its supporting research are evaluated from empirical, methodological, and theoretical perspectives. The cumulative empirical evidence indicates that the major hypothesis of the model is not conclusively supported. Numerous methodological shortcomings are pointed out. The theoretical scheme underlying the model is analyzed, and a new scheme is proposed. The analysis suggests that identification of critical behaviors in the leadership situation should be a starting point for future research on leadership effectiveness.

The contingency model of leadership effectiveness was inductively developed by Fiedler (1964, 1967) to account for findings of 15 years of research on interacting groups. Although the model has stimulated numerous studies, its validity was recently questioned (Graen, Alvares, Orris & Martella, 1970; Graen, Orris & Alvares, 1971a; Graen, Orris & Alvares, 1971b).

The contingency model postulates that performance of interacting groups is dependent on the interaction of leadership style and situational favorableness. According to the model, "task oriented" leaders perform more effectively in very favorable and very unfavorable situations, while "relations oriented" leaders perform more effectively in situations intermediate in favorableness.

The leadership style variable is measured using an instrument called "esteem for least preferred coworker" (LPC). The instrument asks the respondent to think of all coworkers he has ever had. He is then asked to describe the one person that he can least work well with, that is, the person he least preferred as a coworker. The description is made by rating that person on eight-point bipolar adjective scales (usually from 16 to 24 items). Examples of these bipolar adjectives are friendly–unfriendly, supportive–hostile, and cooperative–uncooperative. The LPC score is obtained by totaling up the item values, giving a value of eight

[1] Requests for reprints should be sent to the author who is now at the Department of Business Administration, Faculty of Commerce, Alexandria University, Alexandria, Egypt.

Copyright © 1973 by Academic Press, Inc.
All rights of reproduction in any form reserved.

to the favorable pole of the scale. High score indicates favorable description of the least preferred coworker, while low score indicates unfavorable description. Fiedler (1967) interprets the score on this instrument in the following manner. The lower the LPC score the greater the task orientation, and the higher the LPC the greater the relationship orientation. The assumption is that a task oriented person tends to give a less favorable evaluation to his least preferred coworker than a relations oriented person.

The situational favorableness variable is defined as the degree to which the situation itself provides the leader with potential power and influence over the group's behavior (Fieldler, 1967). Operationally, the variable is indexed along three component dimensions. They are leader–member relation, task structure, and position power.

Leader–member relation represents the leader's evaluation of the member's reaction towards him and his reaction toward the role of the leader (Fiedler, 1967). It has been measured using leader's rating of the group atmosphere, members' ratings of group atmosphere, and the degree to which the leader is sociometrically chosen by group members. Leader's rating of group atmosphere is the method used most frequently.

Task structure represents the extent to which the task is programmable and is defined operationally in terms of a task scaling method developed by Shaw (1963). The task is scaled along four dimensions: goal clarity, decision variability, solution specificity, and goal path multiplicity. An eight-point scale is used to rate the task and a mean scale score of 5.0 is used as a cutting point for dichotomizing task structure.

Position power is the extent to which the leader possesses reward, coercive and legitimate power, as well as having special knowledge or skills not available to the group members. It is measured using an 18-item checklist designed to assess the foregoing privileges that the leader has.

Each of the previous situational dimensions is dichotomized into two levels. The possible combinations of these levels create eight situational types that vary in favorableness. The model assumes that a very favorable situation is characterized by good leader–member relation, high task structure, and strong position power. In contrast, a very unfavorable situation is characterized by poor leader–member relation, low task structure, and weak position power.

Leadership effectiveness, the criterion the model attempts to predict, is defined in terms of performance by the group of its major assigned task. The measure most frequently used is rating of group performance.

Fiedler (1967) summarized the studies upon which the model was built and used the median of the correlations obtained as the point predictions

TABLE 1
MEDIAN CORRELATIONS BETWEEN LPC AND GROUP EFFECTIVENESS IN THE EIGHT SITUATIONAL OCTANTS OF STUDIES USED TO DEVELOP THE MODEL

	Situational classification				
Octant	Leader–member relation	Task structure	Position power	Median rho	Number of correlations
I	good	high	strong	−.52	8
II	good	high	weak	−.58	3
III	good	low	strong	−.33	12
IV	good	low	weak	.47	10
V	poor	high	strong	.42	6
VI	poor	high	weak	—	0
VII	poor	low	strong	.05	12
VIII	poor	low	weak	−.43	12

of the model. Table 1 provides this summary of the relationship between the leaders' LPC measure and group effectiveness in the eight situational octants. According to Fiedler (1967), the order of these octants (from one to eight) represents decreasing degrees of situational favorableness. Therefore, the variability of the correlations observed between octants is due to the difference in situational favorableness. The prediction made is a curvilinear relationship between leadership style and effectiveness depending on characteristics of the situation.

Fiedler's requirement of empirical support for the model is merely obtaining correlations in directions consistent with those in Table 1. Furthermore, he requires that for studies to be an exact test of the model, they should adhere closely to the methodology prescribed by the model.

This paper attempts to evaluate the contingency model using three criteria: (1) empirical validity, (2) methodological rigor, and (3) theoretical adequacy. It is hoped that such an evaluation would point out areas of improvements for future research.

EMPIRICAL VALIDITY

Two recent reviews have attempted to evaluate the evidence on the contingency model The review by Graen *et al.* (1970) compared mean correlations for the different octants of studies used to develop the model and those conducted to test it. They found significant differences between the two sets of correlations and, therefore, concluded that the predictions made by the model are not supported by evidential results.

In defense of his model, Fiedler (1971) reviewed empirical findings on the model and used the directions of the median correlations as a testing

TABLE 2
SUMMARY OF FIELD AND LABORATORY STUDIES TESTING THE CONTINGENCY MODEL

Study	Octants							
	I	II	III	IV	V	VI	VII	VIII
				Field studies				
Hunt (1967)	−.64		−.80		.21		.30	
	−.51		.60				−.30	
Hill (1969)		−.10	−.29			−.24	.62	
Fiedler *et al.* (1969)		−.21		.00		.67[a]		−.51
O'Brien *et al.* (1969)		−.46		.47		−.45		−.14
				Laboratory experiments				
Belgian Navy	−.72	.37	−.16	.08	.16	.07	.26	−.37
(Fiedler, 1966)	−.77	.50	−.54	.13	.03	.14	−.27	.60
Shima (1968)		−.26		.71[a]				
Mitchell (1969)		.24		.43				
		.17		.38				
Fiedler (1971)		.34		.51				
Chemers & Skrzypek (1972)	−.43	−.32	.10	.35	.28	.13	.08	−.33

[a] $p < .05$.

criterion. By differentiating between field and laboratory studies and finding that the majority of field correlations are in the predicted directions, he concluded that the model is supported by data from field studies but not fully supported by data from laboratory studies. Table 2 represents the basic data that Fiedler used in claiming support for the model.

An examination of the correlations in Table 2 reveals that out of the 51 correlations included, only two were statistically significant, and 11 correlations (excluding octant VI) had either zero value or were not in the predicted directions. Field data, on which Fiedler claims support for his model, had only one statistically significant correlation out of the 19 correlations included! Furthermore, correlations not in the predicted directions or of zero value appear in octants III, IV, VII, and VIII. Clearly, and contrary to Fiedler's conclusions, the empirical evidence does not provide conclusive support for the contingency model.

METHODOLOGICAL RIGOR

The contingency model was developed through the use of certain research procedures. Fiedler has prescribed the use of these procedures for studies testing the model. The model, in this sense, is methodology-bound.

Therefore, an examination of the adequacy of this methodology would point out the problems that might have contributed to the inconclusive results characterizing studies based on the model. Three methodological issues seem to stand out in this respect. They are: (1) criteria for testing the model, (2) measurements and their interpretations, and (3) samples and research settings.

Criteria for Testing the Model

As mentioned, Fiedler merely accepts correlations obtained in the predicted direction as indicating meaningful results regardless of whether or not these correlations were statistically significant. The rationale for this lenient requirement is the difficulty of obtaining homogeneous subjects to assign them to the eight octants, especially when the sample unit is the group rather than the individual. This rationale implies that studies on groups should be exempted from the requirement of statistical reliability. If this would be the case, then what alternative criterion should be used to guard against accepting findings that are the product of chance fluctuation?

Graen *et al.* (1971b) have warned against the danger of rejecting the usual criterion of statistical reliability for "we may be forced into generating empirically most of the sampling distribution of our statistic under the null hypothesis before we can accept that hypothesis (p. 289)." The present writer recognizes that there are situations where the statistical test of significance has to be relaxed and the magnitude and direction of results would serve as a substitute criterion for meaningful evidence such as in clinical prediction and exploratory studies. Apparently the contingency model cannot be classified under these cases. The contingency model is a theoretical formulation that offers specific hypotheses and claims generalizability over a variety of leadership settings. Thus, to be accepted, it has to be judged against the usual criteria of scientific investigation.

In his recent review article, Fiedler (1971) tried to show statistical meaningfulness for the correlations obtained in the different octants; thus, he stated: "The joint probabilities of the correlations in octants I, III, and IV are significant below the .05 level. Also 34 of the 45 correlations are in the predicted direction, a finding significant at the .01 level by binomial test. It should be noted that the number of correlations in the predicted direction would be significant at the .01 level by binomial test even if we included the two Graen *et al.* (1970, 1971) experiments. These results permit the conclusion that we are not dealing with random effects (pp. 139–140)."

However, the preceding procedures do not provide a methodologically

rigorous test of the model. First, the model predicts point estimates of the correlations obtained in the different octants (Table 1); hence, the correlations obtained from studies testing the model should be compared and tested for significance against the prediction that the model makes rather than merely testing them for significant difference from zero. Second, using the binomial test is a misleading procedure to test the significance of correlations. It is possible to obtain coefficients all of which are in the predicted direction and range from .01 to .05, but none is statistically significant. Applying the binomial test as used by Fiedler would indicate that they all conform to the predicted direction when every one of them could be a product of mere chance. Third, it is methodologically unsound to apply different criteria and procedures to different portions of the data just to claim support for the whole theory. Failure to obtain statistical significance for the correlations in the different octants (not only octants I, III, and IV) increases doubts about the validity of the whole model.

Fiedler also requires that studies testing the model should adhere closely to the methodology it prescribes; otherwise, the findings cannot constitute a test of the model. Fiedler used this stringent requirement to cast doubt on studies that have disconfirmed his hypothesis such as laboratory studies that obtained results opposite to what the model predicts (Fiedler, 1971). Clearly, this requirement limits the generalizability of the model and casts doubt on its validity.

It is interesting to note that Fiedler relaxed the methodological conformity requirement for studies that deviated from the prescribed procedure such as the Belgian Navy Study (Fiedler, 1965, 1966, 1967) and O'Brien's study (1969), when these studies obtained results in directions consistent with what the model predicted. This double standard procedure, to be sure, yields support for the theory but does not contribute effectively to the understanding of the phenomenon under investigation.

Measurement

LPC. As mentioned earlier, the LPC is used as the measure of leadership style. The test–retest reliability of that instrument has varied from .70 to as low as .31. This means that the instrument could generate error variance that is as low as 30% and as high as 69%. A serious question could be raised as to whether the LPC score reflects a stable motivational tendency or that it measures mostly chance responses.

As to the interpretation of the reliable variance that the scale generates, Fiedler (1967) noted that the LPC score has been resistant to meaningful interpretations which relate it to personality traits and to consistent behavior patterns. The evidence that Fiedler relied on in interpreting

the score as measuring task orientation and relations orientation is far from clear. The behavioral correlates of a high LPC leader indicate that he is more active in interpersonal relations areas (Bishop, 1964). The low LPC leader becomes more active in task related areas only when the leadership situation increases in difficulty (Bishop, 1964). Therefore, only high LPC scores are interpretable independently of the situation they encounter. The study by Myers (1962) seems to support this contention. Myers found that, relative to high LPC persons, low LPC persons are more sensitive to success or failure in competitive conditions than in noncompetitive conditions. However, in an unpublished study by Paul Ninane (Fiedler, 1967), high LPC leaders showed greater sensitivity to success or failure than low LPC leaders. In the Ninane study the leaders were given random feedback information indicating success or failure in the group task independent of actual success or failure. In attempting to explain the findings of Ninane, Fiedler suggests that the low LPC leader "will obtain his rewards from the intrinsic satisfaction of doing the job, and he will, therefore, be less concerned about others' evaluations of his performance" (Fiedler, 1967, p. 58). Fiedler's interpretation is subject to two criticisms: (1) it ignores the dependency of task awareness of the low LPC persons on stress conditions indicated by Bishop and Myers studies; and (2) it assumes that low LPC leaders develop their reinforcing cues directly from task performance irrespective of external information. If this were true, then low LPC persons would have shown sensitivity to success or failure independently of the external factor of competition.

Other studies have provided evidence that is clearly contrary to Fiedler's interpretation. Thus, Steiner (1959) found that low LPC persons tend to be socially more expansive on three social distance measures and less extrapunitive than high LPC persons. In another study by Nealley (Mitchell, 1970) low LPC persons stated that when they were faced with a hypothetical work situation, they would prefer good interpersonal relations, while high LPC persons stated that they wished to have an efficient group. It is clear then that the interpretation of LPC instrument as measuring task-orientation versus relations-orientation is of questionable validity.

Mitchell (1970) suggested an alternative interpretation of the instrument. He found evidence that low LPC leaders tended to give more stereotyped and cognitively simple responses than high LPC leaders. Similar results were obtained by Schroder and his associates (Fiedler, 1971). The findings of these studies suggest that the instrument could be measuring cognitive complexity.

Another interpretation is to consider the LPC instrument as measur-

ing the halo error of rating. The halo error is a common error characterizing supervisory rating of personnel. When present, it indicates the inability of the rater to differentiate between different traits of the ratee. Under the influence of such error, the internal consistency of the rating is high. A rater free from such a tendency would tend to rate each trait independently, thus producing greater internal rating variance. This interpretation is supported by the fact that the low LPC person describes his least preferred coworker in a uniformly undifferentiated manner. The high LPC person's description involves greater differentiation. The LPC person's description has a considerably greater item variance (a standard deviation of .82 for the high versus .43 for the low LPC person) (Fiedler, 1967).

The halo error and cognitive complexity interpretations seem to have more validity than the leadership style interpretation. Based on these interpretations, the independent measure in the model should be item variance rather than item total.

Situational Classification. The first variable of situational classification is leader–member relation. As mentioned, the measure used most frequently is the leader's rating of group atmosphere. Apparently, the leader's rating of group atmosphere reflects his perception of group climate which might or might not coincide with members' perceptions of that climate. The influence of the leader's perception suggests that the measure of group atmosphere is not independent of the LPC measure. Both are subject to the influence of the respondent's cognitive style. There is some indirect evidence indicating that leader's rating of group atmosphere is a measure of his perception rather than actual group atmosphere. Studies by Mitchell (1970) and O'Brien (1967) indicate low intercorrelation between group atmosphere rating by leader, by members, and the degree to which the leader is sociometrically chosen by group members. The usual procedure of rating the group atmosphere after task completion involves another source of contamination. Studies by McGrath and Julian (1963) and Myers (1962) indicate that leader's perception of group atmosphere is partially influenced by success or failure in the group task. The difficulty of obtaining this measure independently of task-related interaction and of perception casts doubt on the feasibility of using even members' ratings or the sociometric choice measure. It would be hard to ask members to reveal feelings toward a leader without prior experience with him in task related activities. This is an especially acute problem in laboratory settings.

The measure of task structure suggested by Fiedler was originally developed by Shaw (1963). It is based on rating of task instructions on specified dimensions. Graen *et al.* (1971b) criticized Shaw's measure and

suggested the use of a measure developed by Hackman, Jones, and McGrath (1967). The measure of Hackman *et al.* uses ratings of dimensions of group outputs rather than task instructions.

The prescribed procedure of measuring position power involves administering an 18-item checklist that assesses the degree to which the leader possesses reward, coercive and legitimate power, as well as having special knowledge or skills not available to the group members. It seems that the procedures of operationalizing this variable were responsible for its meager influence on group performance. First, in a number of studies this variable was measured by the intuitive judgment of the investigator based upon the expected status and role of the leader. For example, in the study by Meuwese and Fiedler (1967), 162 college seniors enrolled in ROTC programs were divided into 54 three-man groups and the member with the highest ROTC rank in the group was designated as a leader. Graen *et al.* (1970) criticized this procedure for not being validated and for being highly intuitive. Second, the 18-item checklist measures special privileges that the leader has in the group. These privileges might or might not reflect the leader's power. Mere control of incentives and sanctions does not provide sufficient indication of social power. For social power to exist, what the leader controls must also be of important value to the group members. Apparently Fiedler's measure of power does not take account of members' needs or their valuation of what the leader controls. Third, the checklist includes items that overlap with measures of leader–member relation, suggesting that the measure of position power is not independent of leader–member relation. For example, one of the items measures the degree to which the leader's position is dependent on members (i.e., being elected versus being appointed). This is a measure of leader's acceptability in the group which overlaps with measures of leader–member relation.

In addition to the problems of measuring the individual variables of the situational dimension, the classification of this dimension as proposed by the model poses great difficulty in interpretation. The model proposes that the classification represents degrees of situational favorableness. This evidence does not show, however, that the particular order of situational classification of the three variables taken together corresponds to different degrees of the leader's influence over group members. Specific evidence is needed, if this interpretation is to be accepted, to show that the different degrees of members' compliance to influence attempts by the leader actually correspond to the situational classification proposed by the model (Mitchell, Biglan, Oncken & Fiedler, 1970).

Mitchell *et al.* (1970) referred to another shortcoming of the situational classification. They criticized the absence of an interval measure

of situational favorableness and suggested the use of a measure of leader's potential influence developed by O'Brien (1969) and derived from structural rule theory. O'Brien's measure is an interval scale that includes the same situational variables suggested by the contingency model but treats them as a whole. However, for O'Brien's measure to be interpreted as a favorableness measure, evidence still has to be shown that the measure has high positive correlation to members' compliance to influence attempts by the leader.

The proposed moderating effect of the situational variable creates an additional difficulty. Korman (1971) has pointed out that contingency hypotheses require *a priori* specification of the "critical values" of situational variables. He argues that the moderating effect of these variables require a specification of the values at which the effect of the variables shifts in direction. Without *a priori* specification of these values, different studies might choose different values. With the lack of comparability of the situational values, contradictory findings are likely to be obtained. This increases the chance of making incorrect conclusions about the truthfulness of the contingency hypotheses.

Samples and Research Settings

The contingency model was initially inferred from a wide variety of settings that included a rather broad sampling of actors, behaviors, and contexts. The samples out of which the data were generated included high school basketball teams (Fiedler, 1954), infantry squads (Havron *et al.*, 1951), B-29 bomber crews and army tank crews (Fiedler, 1955), open-hearth steel shop crews (Cleven & Fiedler, 1956), company managers and chairmen of the boards of directors of small cooperatively owned corporations (Godfry *et al.*, 1959), antiaircraft artillery crews (Hutchins & Fiedler, 1960), Dutch University students (Fiedler *et al.*, 1961), church groups (Fiedler, Bass & Fiedler, 1967), Navy ROTC groups (Anderson & Fiedler, 1964), sales display teams and service station managers (Hawkins, 1965), and ROTC groups (Meuwese, 1964).

In a way, the broad sampling of settings and populations would shed light on the generalizability of the model, if the full range of the variables were investigated in each setting. However, the data used to generate the model were based on samples that differed from octant to octant. Table 3 shows that none of the studies investigated the full range of situational types simultaneously. With the exception of the ROTC study, no sample investigated more than two octants. The limited representation of the octants raises the possibility that the data obtained could merely be a function of the populations sampled rather than reflecting true differences between octants.

TABLE 3
POPULATIONS AND SETTINGS SAMPLED FOR THE DIFFERENT OCTANTS IN THE STUDIES THAT WERE USED TO DEVELOP THE MODEL

Populations and settings	Octants							
	I	II	III	IV	V	VI	VII	VIII
B-29 bomber crews	X				X			
Army tank crew	X				X			
Antiaircraft crews	X				X			
Infantry squads	X				X			
Open-hearth steel shop crews	X				X			
Company managers	X				X			
High school basketball teams		X						
Student surveying parties		X						
ROTC groups			X	X			X	X
Navy ROTC groups			X				X	
Dutch university students				X				X
Church groups				X				X
Mental health groups				X				X
Chairmen of boards of directors of small corporations				X				X

The limited representation of the situational octants continued to characterize later studies. Thus, out of approximately 26 studies testing the model, only three studies (Fiedler, 1966; Chemers & Skrzypek, 1972; Graen *et al.*, 1971a) tested predictions in all eight octants simultaneously. None of the 40 correlations reported in these three studies were statistically significant, and 14 of them were opposite to the predicted direction.

An additional weakness of research on the contingency model has been the use of small samples. The majority of the studies reported by Fiedler (1967) used sample sizes under 10 groups. Studies based on small samples involve greater probability of committing type I and type II errors than those based on large samples. Apparently, the use of small samples has served to perpetuate the ambiguity characterizing results obtained on the contingency model.

THEORETICAL ADEQUACY

The nature of the theoretical scheme underlying research on leadership effectiveness and the contingency model has contributed to the difficulty of obtaining conclusive evidence in this area. An examination of leadership research in the last 25 years indicates that research efforts have been attempting to explain variability in group effectiveness in terms of leadership style. Leadership traits were measured and correlated with measures of group effectiveness. The assumption has been that the in-

fluence of leader's style or traits could be detected in the performance of his group. The contradictory findings obtained on this hypothesis directed the attention of the investigators to the importance of situational variables in moderating this simple relationship. Therefore, Fiedler's model came to propose a curvilinear relationship between leadership style and group effectiveness, a relationship that depends on particular situational conditions.

The model has made specific predictions of the correlations to be expected in the eight situations. However, the correlations by themselves gave little insight into the underlying reasons for the predictions proposed. In attempting to offer an explanation for the correlations proposed, Fiedler (1967) cited data from two studies, the Dutch creativity study (Meuwese & Oonk, 1960), and the ROTC study (Fiedler, 1967). Central to his explanation is the inference that low LPC leaders tend to engage in task related activities while high LPC leaders tend to engage in relations oriented activities, and that the anxiety experienced by leaders in unfavorable conditions tends to intensify these tendencies.

An examination of the Dutch and ROTC studies reveals that these studies do not cover systematically the eight situational octants. Furthermore, the data from the two studies clearly indicate that under favorable conditions the low LPC leaders exhibit the greater percentage of relations oriented comments, while the high LPC leaders exhibit the greater percentage of task oriented comments. The data indicate no clear behavioral difference between high LPC and low LPC leaders in the intermediate conditions. It appears that the interpretation of the LPC instrument is a central part in Fiedler's explanation. Such an interpretation is questionable in the light of the ambiguity characterizing the meaning of the LPC.

Fiedler makes no reference to particular group behaviors and reactions to leaders' behaviors. It is not clear from the model how successful leaders exert their influence on task performance in each of the situational octants. The model does not specify how leaders contribute to performance of structured versus unstructured tasks. How do they influence their groups to engage in task related activities in the extremely unfavorable condition, when the group is initially resentful to that influence? What are the members' behaviors in reaction to the leader's influence attempts in every octant? What are the patterns of interaction between the leader's behaviors and group behaviors in each of the eight octants?

Having no knowledge about the critical behaviors of leaders and members in every octant makes it extremely difficult to assign any behavioral meaning to the correlations proposed. A behavioral explanation would require incorporating leaders' activities and group activities in the formal hypothesis suggested by the model.

It seems that the appealing simplicity of correlating leaders' traits with outcomes of group performance is greatly responsible for the difficulty of interpreting the predictions suggested by the model. The present writer believes that the temptation of this seemingly simple scheme is also responsible for the lack of conclusive evidence on the predictions proposed.

A modification of this simplified scheme would require inclusion of intervening linkages between the predictors (leader's traits) and the effectiveness criteria (group performance outcomes). Figure 1 represents a proposed scheme that includes these linkages.

The proposed scheme could serve as a conceptual frame of reference to guide research on leadership. Note the complexity indicated by the behavioral linkages intervening between leader's characteristics (traits) and outcomes. It suggests possible moderating influence (dotted arrows) as well as independent effects that different situational variables have on the different linkages in the scheme. The scheme further proposes that we ought to examine the linkages between leader and group characteristics and behaviors and between behaviors and outcomes.

It is important to note that predicting the behavior of an individual (leader) using measures of his characteristics is recognized as being complex enough by researchers in the area of testing (Dunnette, 1963; Ghiselli, 1963). It involves the possibility of interactions between measures of individual traits and different trait combinations, different types of individuals, and different behaviors with the possible filtering effect that situational variables exert on behaviors and behavior outcomes.

Predicting leadership effectiveness is even more complex. It involves extending the prediction of individual's (leader) behaviors to include

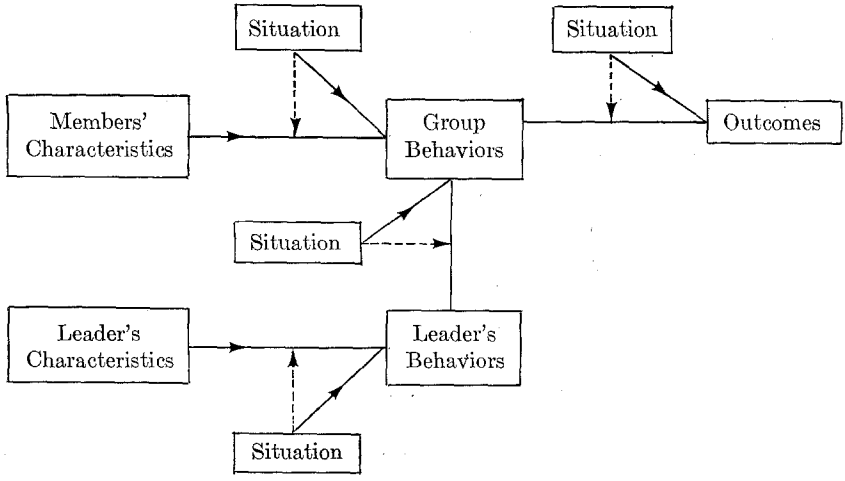

FIG. 1. A proposed conceptual scheme for research on leadership effectiveness.

also their influence on group behaviors and outcomes. Its complexity is derived from extending the prediction and from using two different units of analysis, the individual and the group. Clearly, the contingency model is based on a simplistic scheme that omits essential linkages intervening between leader's traits and group outcomes.

Korman (1971) pointed out another theoretical limitation of the contingency model. He noted that the model focuses only on the static aspects of the leadership phenomenon. The model ignores the long-range influence of the situation on the leader and the group; thus, it limits its scope to only static conditions. According to Korman, this implies a naive assumption that the world "is a static one" whereby the relationships being postulated are invariant in nature.

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The present review evaluated the evidence on the validity of the contingency model, the methodology it prescribes, and the conceptual scheme underlying it.

The empirical evidence bearing on the model is inconclusive. Contradictory results are obtained from studies testing the model, most of which lack statistical significance.

The starting point in improving the model requires examining and monitoring leaders' behaviors and group behaviors in different situational conditons. Research on leadership effectiveness should examine the intervening linkages suggested by the conceptual scheme proposed in Fig. 1, along with the possibility of short-term and long-term influences of different situational variables on the different linkages.

The following are some methodological suggestions for future studies on the contingency model and on leadership effectiveness in general.

1. Larger samples should be used to clarify the lack of statistical significance of results previously obtained. The criterion of statistical reliability should be employed to evaluate the validity of the predictions suggested by the model.

2. Leaders' behaviors and group behaviors should be monitored using the critical incident methodology (Flanagan, 1954; Kirchner & Dunnette, 1957). The examination of behavior using this methodology would enhance our understanding of the connection between critical leadership behaviors, critical group behaviors, and the outcome criteria of these behaviors. This methodology would also direct attention toward predicting only the important elements of behaviors exhibited in the different leadership situations.

3. The LPC could be measuring certain cognitive traits of the leader,

but other cognitive, personality, and motivational variables should be used to predict critical leaders' behaviors. Evidence of high test–retest reliability and of the construct validity of the predictor measures, including the LPC, should be provided prior to using them.

4. Measures of cognitive, personality, and motivational variables of group members should be used, along with leaders' behaviors, as predictors of critical group behaviors.

5. Investigation of the effect of situational variables should include other variables such as group size, member–member relations, other dimensions of group task (examples: content of activities, performance cycle), and incentive scheme. Measures of situational variables and any possible classification of situational types should show evidence of construct validity. The moderating as well as the independent effect of the situational variables should be investigated.

6. It is important to examine causal connections between traits, behaviors, situations, and outcomes. Therefore, causal hypotheses should be tested using experimental settings. Experimental studies should be used to validate causal inferences generated from field data.

7. Longitudinal studies should be designed to examine the long-term interactions between the variables. The use of the critical incident method would facilitate focusing on the dynamic aspects of leader and group behaviors.

REFERENCES

ANDERSON, L. R., & FIEDLER, F. E. The effect of participatory and supervisory leadership on group creativity. *Journal of Applied Psychology,* 1964, **48,** 227–236.

BISHOP, D. W. Relations between task and interpersonal success and group member adjustment. Urbana, IL: Group Effectiveness Research Laboratory, University of Illinois, 1964. Mimeograph.

CHEMERS, M. M., & SKRZYPEK, G. J. Experimental test of the contingency model of leadership effectiveness. *Journal of Personality and Social Psychology,* 1972, **24,** 172–177.

DUNNETTE, M. D. A modified model for test validation and selection research. *Journal of Applied Psychology,* 1963, **47,** 317–323.

FIEDLER, F. E. Assumed similarity measures as predictors of team effectiveness. *Journal of Abnormal and Social Psychology,* 1954, **49,** 381–388.

FIEDLER, F. E. *A theory of leadership effectiveness.* New York: McGraw-Hill, 1967.

FIEDLER, F. E. The influence of leader-keyman relations on combat crew effectiveness. *Journal of Abnormal and Social Psychology,* 1955, **51,** 227–235.

FIEDLER, F. E. A contingency model of leadership effectiveness. *In* L. Berkowitz (Ed.), *Advances in experimental social psychology.* New York: Academic Press, 1964.

FIEDLER, F. E. Engineering the job to fit the manager. *Harvard Business Review,* September, 1965, 115–122.

FIEDLER, F. E. The effect of leadership and cultural heterogeneity on group per-

formance: A test of the contingency model. *Journal of Experimental Social Psychology,* 1966, **2,** 237–264.

FIEDLER, F. E. Validation and extension of the contingency model of leadership effectiveness: A review of empirical findings. *Psychological Bulletin,* 1971, **76,** 128–148.

FIEDLER, F. E., BASS, A. R., & FIEDLER, J. M. The leader's perception of coworkers, group climate, and group creativity: A cross-validation. *In* F. E. Fiedler, *A theory of leadership effectiveness.* New York: McGraw-Hill, 1967.

FIEDLER, F. E., LONDON, P., & NEMO, R. S. Hypnotically induced leader attitudes and group creativity. *In* F. E. Fiedler, *A theory of leadership effectiveness.* New York: McGraw-Hill, 1967.

FIEDLER, F. E., MEUWESE, W. A. T., & OONK, S. Performance of laboratory tasks requiring group creativity. *Acta psychologica,* 1961, **18,** 100–119.

FIEDLER, F. E., O'BRIEN, G. E., & ILGEN, D. R. The effects of leadership style upon the performance and adjustment of volunteer teams operating in stressful foreign environment. *Human Relations,* 1969, **22,** 503–514.

FLANAGAN, J. C. The critical incident technique. *Psychological Bulletin,* 1954, **51,** 327–358.

GHISELLI, E. E. Moderating effects and differential reliability and validity. *Journal of Applied Psychology,* 1963, **47,** 81–86.

GODFREY, E. P., FIEDLER, F. E., & HALL, D. M. *Boards, managers, and company success.* Danville, IL: Interstate Press, 1959.

GRAEN, G., ALVARES, K. M., ORRIS, J. B., & MARTELLA, J. A. Contingency model of leadership effectiveness: Antecedent and evidential results. *Psychological Bulletin,* 1970, **74,** 285–296.

GRAEN, G., ORRIS, J. B., & ALVARES, K. M. Contingency model of leadership effectiveness: Some experimental results. *Journal of Applied Psychology,* 1971a, **55,** 196–201.

GRAEN, G., ORRIS, J. B., & ALVARES, K. M. Contingency model of leadership effectiveness: Some methodological issues. *Journal of Applied Psychology,* 1971b, **55,** 205–210.

HACKMAN, J. R., JONES, L. E., & MCGRATH, J. E. A set of dimensions for describing the general properties of group-generated written passages. *Psychological Bulletin,* 1967, **70,** 379–390.

HAVRON, M. D., LYBRAND, W. A., COHEN, E., KASSEBAUM, R. G., & MCGRATH, J. E. *The assessment and prediction of rifle squad effectiveness.* Washington, DC: The Adjutant General's Office, Personnel Research Branch, 1954.

HAWKINS, C. A. A study of factors mediating a relationship between leader rating behavior and group productivity. Unpublished doctoral dissertation, University of Minnesota, 1962.

HILL, W. The validation and extension of Fiedler's theory of leadership effectiveness. *Academy of Management Journal,* March, 1969, 33–47.

HUNT, J. G. A test of the leadership contingency model in three organizations. *Organizational Behavior and Human Performance,* 1967, **2,** 290–308.

HUTCHINS, E. B., & FIEDLER, F. E. Task-oriented and quasitherapeutic role functions of the leader in small military groups. *Sociometry,* 1960, **23,** 393–406.

KIRCHNER, W. K., & DUNNETTE, M. D. Identifying the critical factors in successful salesmanship. *Personnel,* 1957, **34,** 54–59.

KORMAN, A. On the development of contingency theories of leadership: Some

methodological considerations and a possible alternative. Paper presented at University of Chicago Graduate School of Business, December, 1971.

McGrath, J. E., & Julian, J. W. Interaction process and task outcome in experimentally created negotiation groups. *Journal of Psychological Studies,* 1963, **14,** 117–138.

Meuwese, W. A. T. The effect of the leader's ability and interpersonal attitudes on group creativity under varying conditions of stress. Unpublished doctoral dissertation, University of Amsterdam, 1964.

Meuwese, W. & Oonk, S. Enkele determinanten von creativiteit, structuur en proces in kleine experimentele groepen. Unpublished werkstuk, University of Amsterdam, 1960.

Mitchell, T. R. Leader complexity, leadership style, and group performance. Unpublished doctoral dissertation, University of Illinois, 1969.

Mitchell, T. R. The construct validity of three dimensions of leadership research. *Journal of Social Psychology,* 1970, **80,** 89–94.

Mitchell, T. R., Biglan, A., Oncken, G., & Fiedler, F. E. The contingency model: Criticism and suggestions. *Academy of Management Journal,* 1970, **13,** 253–267.

Myers, A. E. Team competition, success, and the adjustment of group members. *Journal of Abnormal and Social Psychology,* 1962, **65,** 325–332.

O'Brien, G. E. Group structure and the measurement of potential leader influence. *Australian Journal of Psychology,* 1969, **21,** 277–289.

Shaw, M. E. Scaling group tasks: A method for dimensional analysis. Gainesville, FL: University of Florida, 1963. Mimeograph.

Shima, H. The relationship between the leader's modes of interpersonal cognition and the performance of the group. *Japanese Psychological Research,* 1968, **10,** 13–30.

Steiner, I. D. Interpersonal orientation and assumed similarity between opposites. Urbana, IL: Group Effectiveness Research Laboratory, University of Illinois, 1959. Mimeograph.

Received: February 9, 1972