AD-750 583

THE CONTINGENCY MODEL: SOME IMPLICATIONS OF ITS STATISTICAL AND METHODOLOGICAL PROPERTIES

Samuel C. Shiflett

Army Medical Research Laboratory
Fort Knox, Kentucky

17 August 1972

US ARMY
MEDICAL RESEARCH LABORATORY
FORT KNOX, KENTUCKY 40121

REPORT NO. 996

THE CONTINGENCY MODEL: SOME IMPLICATIONS OF ITS STATISTICAL AND METHODOLOGICAL PROPERTIES

(Interim Report)

by

CPT Samuel C. Shiflett, MSC (Ph.D.)

17 August 1972

Approved for public release; distribution unlimited

UNITED STATES ARMY
MEDICAL RESEARCH AND DEVELOPMENT COMMAND

DISPOSITION

Destroy this report when no longer needed.
Do not return it to the originator.

ACKNOWLEDGMENT

The author wishes to thank James H. Davis, Fred E. Fiedler, and George Graen for their comments on an earlier version of this report. Gratitude is also expressed to Martin Chemers, Robert Hardy, Jerry Hunt, and George Skrzypek who so kindly provided their data for reanalysis. This report would have been impossible without their generosity.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) US Army Medical Research Laboratory Fort Knox, Kentucky 40121	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

THE CONTINGENCY MODEL: SOME IMPLICATIONS OF ITS STATISTICAL AND METHODOLOGICAL PROPERTIES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Interim Report

5. AUTHOR(S) (First name, middle initial, last name)

CPT Samuel C. Shiflett, MSC (Ph.D.)

6. REPORT DATE 17 August 1972	7a. TOTAL NO. OF PAGES 26	7b. NO. OF REFS 34
8a. CONTRACT OR GRANT NO. b. PROJECT NO. 3A061102B71R c. Task No. 03 d. Work Unit No. 128	9a. ORIGINATOR'S REPORT NUMBER(S) 996 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY US Army Medical Research and Development Command, Washington, D. C. 20314

13. ABSTRACT

The statistical and methodological adequacy of several aspects of Fiedler's (1967) contingency model of leadership effectiveness is considered. It is suggested that the use of rank order correlations, instead of product moment correlations, produces distortions related to sample size. LPC (least preferred coworker) is shown to be a reasonable predictor of group performance. But the model implies that LPC is unrelated to group performance under some conditions of situational favorableness. Problems with a multidimensional interpretation of LPC and a unidimensional interpretation of situational favorableness are discussed. Problems of research strategies used to test and validate the model are considered in terms of sample size, inclusion of non-significant results, the difficulty involved in adequately testing higher order interactions, and the inability to make between-octant comparisons. It was concluded that the model is capable of directing meaningful research as long as traditional research procedures designed to safeguard internal and external validity are carefully exercised.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

UNCLASSIFIED
Security Classification

I-A

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Leadership Effectiveness						
Contingency Model						
Least Preferred Coworker Scale						
Group Performance						

*U.S.Gov nment Printing Office: 1972 — 759-816/3059 Reg n 5-11

UNCLASSIFIED

Security Classification

I-B

AD ____________________

REPORT NO. 996

THE CONTINGENCY MODEL: SOME IMPLICATIONS OF ITS STATISTICAL AND METHODOLOGICAL PROPERTIES

(Interim Report)

by

CPT Samuel C. Shiflett, MSC (Ph.D.)*

US ARMY MEDICAL RESEARCH LABORATORY
Fort Knox, Kentucky 40121

17 August 1972

Military Performance:
Physical Decrement and Enhancement
Work Unit No. 128
Psychiatry
Task No. 03
Research in Biomedical Sciences
DA Project No. 3A061102B71R

*Present address:

US Army Motivation and Training Laboratory,
Room 239, Commonwealth Building, 1300 Wilson Boulevard,
Arlington, Virginia 22209

I-a.

USAMRL REPORT NO. 996
DA PROJECT NO. 3A061102B71R

ABSTRACT

THE CONTINGENCY MODEL: SOME IMPLICATIONS OF ITS STATISTICAL AND METHODOLOGICAL PROPERTIES

OBJECTIVE

To consider the statistical and methodological adequacy of several aspects of the contingency model of leadership effectiveness.

SUMMARY

It was suggested that the use of rank order correlations, instead of product moment correlations, produces distortions related to sample size. Least preferred coworker (LPC) scores were shown to be a reasonable predictor of group performance, but the model implies that LPC is unrelated to group performance under some conditions of situational favorableness. Problems with a multidimensional interpretation of LPC and a unidimensional interpretation of situational favorableness were discussed. Problems of research strategies used to test and validate the model were considered in terms of sample size, inclusion of non-significant results, the difficulty involved in adequately testing higher order interactions, and the inability to make between-octant comparisons.

CONCLUSION

In spite of various methodological and statistical problems, the model is capable of directing meaningful research, as long as traditional research procedures designed to safeguard internal and external validity are carefully and adequately exercised.

TABLE OF CONTENTS

THE CONTINGENCY MODEL: SOME IMPLICATIONS OF ITS STATISTICAL AND METHODOLOGICAL PROPERTIES

INTRODUCTION

The "contingency model of leadership effectiveness," proposed by Fiedler (1,2), has recently generated a substantial amount of interest, particularly of an evaluative nature. The first comments on the shortcomings and problems of the model were those presented by Fiedler himself (2). Soon thereafter, O'Brien (3) expanded upon the methodological problems involved in the analysis and definition of "situational favorableness." In 1970, Graen, Alvares, Orris, and Martella (4) published an extensive evaluation of the antecedent and evidential results which were used to develop and test the model; they concluded that non-significant results had cast "grave doubt on the plausibility of the contingency model of leadership effectiveness" (p. 295). An article by Mitchell, Biglan, Oncken, and Fiedler (5) presented a broad review of the model, again commenting on some methodological problems and concluded that "the total pattern of correlations are significantly in support of the predictions of the model" (p. 260). Subsequently, Fiedler (6) provided a more detailed review of research related to the model. He basically claimed strong support for the general hypothesis that the correlation between leadership style and group performance is moderated by the situational favorableness dimension, even though the dimension is operationalized in a wide variety of ways. Fiedler also concluded that the model, as originally defined by eight cells, is valid for prediction of leadership performance under field conditions. In his review, Fiedler (6) proposed several situational distinctions which he did not directly tie to the concept of situational favorableness. Specifically, he made distinctions between field and laboratory settings, interacting and coacting groups, and learning versus task performance.

A major concern of all of these reviews has been to determine whether the model can serve as a valid and useful basis for generating new hypotheses and good research. Except for the discussion of Fiedler's use of non-significant results by Graen *et al* (4), very little has been said about the implications of the statistical methodology of the model. The purpose of this report is to approach the model by discussing several explicit characteristics of the model which have not been adequately treated previously. In particular, some theoretical implications which result from the statistical treatment of the supporting data, even though the implications may not have been intended by Fiedler, will be discussed. It is hoped that this approach will lead to a psychologically reasonable, as well as a statistically sound, utilization of the model in guiding future research.

RANK ORDER VS. PRODUCT MOMENT CORRELATIONS

Fiedler has traditionally used the Spearman rank order correlation (ρ) in his model rather than the more common Pearson product moment correlation coefficient (r). His interpretation of the obtained coefficients closely follows traditional interpretations of r and he acts on the assumption that ρ is a valid estimate of the true population correlation ($\hat{\rho}$). The validity of his assumption is questionable on several grounds. Although ρ is usually assumed to be a reasonably good estimate of r, and therefore, of $\hat{\rho}$, this relation assumes relatively large samples and no tied ranks. Sample sizes used in calculating ρ's reported by Fiedler range from 6 to 22, with a substantial majority of the N's equal to 6, 7, or 8. Hays (7) notes that the distribution of ρ has a "curious jagged or serrated appearance due to particular constraints on the possible values of $\sum_i D_i^2$ for a given N" (p. 646). This is particularly true when N is small. Thus, a single reversal in a rank with a sample size of only 6 can cause a substantial shift in the size of the correlation. This effect has been demonstrated by Graen, Orris, and Alvares (8) in their exercise using successively smaller samples in testing the model. Nevertheless, the exact distribution of ρ is known when no tied ranks exist and significance levels can be determined. Under these circumstances, the best use of ρ would probably be to test the null hypothesis (H_0: $\rho = 0$), rather than to use it as an estimate of variance accounted for, or as an indicator of, the strength of a relationship. Unfortunately, most of the ρ's reported by Fiedler are not statistically significant and so his use of the correlations has been virtually opposite to that suggested here. He has certainly not presented evidence adequate enough to reject H_0 on the basis of the correlations alone. However, Fiedler (6) has demonstrated that the pattern of correlations within the various cells of the model is in the direction of support for the model.

The method of selecting leaders in several experiments dealing with the model requires some comment. Least preferred coworker (LPC) scores are typically obtained in some sort of pre-testing session and leaders are subsequently assigned to experimental conditions as a function of these scores, with subjects scoring at the extremes of the LPC range being designated as leaders. This procedure prevents the sampled leader LPC scores from approaching anything like a normal distribution. While this effect may be appropriately reflected by r, it will be obscured when using ρ 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. It is not clear whether potential bias exists as a result of this procedure, but it seems reasonable that a biserial correlation would be more appropriate, or even better perhaps, the analysis of variance in which leader LPC is treated as a dichotomous factor.

While the discussion above is, in general, critical of Fiedler's use of the rank order correlation, the criticisms are based primarily on theoretical grounds upon which statisticians are not in particularly close agreement. A more direct approach would be to simply make an empirical comparison between the two indices, based on data from several different studies purporting to test the model. This was done with the expectation that ρ would tend to slightly underestimate r. Guilford (9) suggests that the difference will be less than .02 when r is around .50. Raw data from several published studies were obtained and both ρ and r were calculated.*

Shiflett and Nealey (10) experimentally manipulated situational favorableness within laboratory situations defined by octants 3 and 4. Groups were assembled on the basis of LPC and intelligence scores. Correlations derived from this study are presented in Table 1. As one can see, the rank order correlations overestimated the product moment correlations in three of the four cases. In one case, ρ reached statistical significance although the corresponding r and the originally reported analysis of variance did not reach significance. In the fourth case, ρ underestimated r by a substantial margin. In no case was there close correspondence between ρ and r.

TABLE 1

Correlations Between Leader LPC and Group Performance on an Unstructured Task [Shiflett and Nealey (10)]

		High Ability Groups (N=8)	Low Ability Groups (N=8)
Strong Leader Power (Octant 3)			
	r	.644	-.491
	ρ	.405	.714*
Weak Leader Power (Octant 4)			
	r	-.160	.341
	ρ	-.321	.458

*p < .05

*Only studies which appeared to at least partially support the model were considered for reanalysis since results from non-supportive studies would have produced equivocal results within the context of the present approach. Thus, the Graen *et al* (8,29) studies were not considered for additional analysis in this study. In addition, these studies have been criticized by Fiedler (6,30) on methodological grounds, as was the Shiflett and Nealey (10) study, which is included in the present set of reanalyzed studies.

Hunt (11) reported a field study, conducted in three different organizations, in which he looked at the effectiveness of managers and foremen in a variety of situations. Correlations from his study are presented in Table 2. As can be seen, eight of the 11 ρ's overestimated the corresponding r while the remaining three underestimated r. In the latter case, one of the Pearson r's reached statistical significance thus providing even stronger support for Hunt's conclusion that the contingency model was supported.

TABLE 2

Correlations Between Leader LPC and Group Performance in Three Organizations [Hunt (11)]

	Good Group Atmosphere			Poor Group Atmosphere		
Octants 1 and 5	N	ρ	r	N	ρ	r
Research Chemists	7	-.64	-.40			
Shop Craftsmen	6	-.48	-.23	5	.80	.92*
Meat Departments	10	-.51	-.40	11	.21	.36
Grocery Departments	13	-.06	-.10	11	.49	.44
Octants 3 and 7						
Research Chemists	6	.60	.43	5	.30	.24
General Foremen	5	-.80	-.74	5	-.30	-.12

*$p < .05$

Hardy (12) tested half of the contingency model (octants 1 through 4) using classroom groups with good leader-member relations. Data from his study are presented in Table 3. In every case, ρ underestimates r, in contrast to most of the data reported above which suggest that the opposite relationship is true. In the present case, however, the Ns are relatively large, with 28 observations per coefficient. In the Hunt (11) study, two of the three cases where ρ underestimates r again involved relatively large Ns [11 and 13], while the overestimating ρ's involved much smaller Ns [5, 6, and 7]. The correlations from Shiflett and Nealey (10) were also based on rather small Ns [8]. Together, these data suggest that ρ tends to overestimate r when N is very small, but underestimates r when N is relatively large. These results indicate further that, while the estimated strength of the relationships may be distorted somewhat as a function of the sample size and the statistic used, the overall pattern of results remains essentially unchanged.

TABLE 3

Correlations Between Leader LPC and Group Performance
in Classroom Groups [Hardy (12)]

	Octant 1 High Power Structured Task	Octant 2 Low Power Structured Task	Octant 3 High Power Unstructured Task	Octant 4 Low Power Unstructured Task
ρ	-.387*	.209	-.211	.500**
r	-.455*	.281	-.280	.523**
N	28	28	28	28

*$p < .05$

**$p < .01$

THE POWER OF LPC

Statistical significance, by itself, does not really tell much about how "important" a relationship is "psychologically." That is, with a large enough sample and a powerful enough statistical technique, even the most trivial relationship can be shown to be statistically significant. As a result, there is now an increasing frequency of attempts to characterize psychological importance with various measures of the amount of variance actually accounted for. The "coefficient of determination" (r^2), represents the appropriate statistic for use in looking at the power of LPC within the model. It has already been pointed out that the ρ's used by Fiedler may or may not be very good estimates of r. Nevertheless, the median ρ's reported by Fiedler (2) will be treated as r's and ρ^2 will be estimated accordingly. This liberty is taken since the use of median correlations within the model probably tends to minimize any bias and so is a more accurate reflection of the median r than is any individual ρ likely to be an accurate estimate of the corresponding r. Furthermore, this approach is in keeping with Fiedler's traditional interpretation of the ρ's within the model, which basically assumes that ρ is an accurate estimate of r. The values of ρ^2 are presented in Table 4. As can be seen, only in octants 1, 2, and 4 is more than 20% of the variance in group performance accounted for by leader LPC. All three of these correlations are in the favorable half of the model, where group atmosphere is relatively good. In the less favorable half of the model, where group atmosphere is relatively poor, the relationship is not nearly so clear. One cell shows no relationship, one is undetermined, and two cells show 18% of the variance accounted for. By averaging the ρ^2 for each half of the model, it becomes apparent that LPC is about twice as effective a predictor when group atmosphere is good than when it is poor. To the extent that these are valid estimates of true variance, then for favorable situations LPC is as respectable a predictor as most other variables which have been studied in small groups [see, for example, Mann (13)]. It is still necessary to remember that most of the correlations used to derive these values were statistically non-significant and based on rather small Ns.

TABLE 4

Median Correlations Between Leader LPC and Group Performance [Fiedler (2)]

	Octant	Median ρ	N	ρ^2	$\overline{X}_\rho{}^2$
Favorable	1	-.52	8	.27	
	2	-.58	3	.34	
	3	-.33	12	.11	.24*
	4	.47	10	.22	(.20)**
Unfavorable	5	.42	6	.18	
	6	-	0	-	
	7	.05	12	.00	.12*
	8	-.43	12	.18	(.11)**

*Unweighted

**Weighted

THE MEANING OF LPC

What Fiedler refers to as "leadership style" has traditionally been measured by assessing the leader's evaluation of his LPC on a series of bipolar semantic differential scales. Each of the scales is summed to yield a single LPC score. The interpretation of this score has evolved and changed as Fiedler's research program has progressed over the years. In general, leaders who give low esteem to their LPC, referred to as low LPC leaders, are described by Fiedler as primarily oriented toward, or motivated by, task-related concerns. On the other hand, high LPC leaders, those who give high esteem to their LPC, are believed to be primarily interested in, and motivated by, interpersonal relations. Jacobs (14) has recently suggested that high LPC leaders be described as "social specialists" and low LPC leaders as "task specialists."

Regardless of the precise terminology used, or which interpretation of the score is advanced, one implication of these definitions has remained constant: a single score has been used to assess in which of these two directions a leader leans. This procedure implies that motives or behaviors characterized as "interpersonally oriented" and "task oriented" are opposite poles of a single dimension. This unidimensional conceptualization of leadership style is in direct contrast to the general proposition that there are at least two dimensions of leadership behavior, as suggested by the Ohio State studies (15), and the Michigan studies (16). This unidimensional concept also implies that if an individual is highly oriented toward the task, he must necessarily be completely unconcerned with or oriented away from social concerns, and vice versa. This interpretation is a necessary property of a unidimensional continuum. The status of "middle LPC" individuals in this state of affairs is unclear. Are they high, moderate, or low on both task and social concerns? Or, is the score meaningless for persons who don't give extreme ratings to their LPC? The more traditional approaches imply that the task orientation and social orientation dimensions are orthogonal, or at least, oblique. Thus, one's standing with respect to one dimension does not completely determine his standing on the other and he may be a leader who excels at <u>both</u> task and social concerns.

A unidimensional approach to leadership clearly renders impossible such approaches as suggested by Yukl (17) and by Graen, Dansereau, and Minami (18), wherein the interaction between consideration and structuring behaviors by the same leader are considered with respect to group performance. In support of a non-unidimensional interpretation of LPC, Yukl (19), Shiflett (20), and Fox, Hill, and Guertin (21) have presented evidence that the LPC scales may be measuring at least two distinguishable evaluative dimensions corresponding to task-related and socially-related characteristics. Current work by Fiedler and his associates (22) is addressed to the task of clarifying the meaning of LPC. Recent data suggest that LPC may be an index of goal hierarchies which change as a function of stress or situational favorableness. This research approach appears promising but its effectiveness and theoretical and statistical appropriateness will be partly contingent on the recognition that as long as LPC is expressed as a unidimensional score, its theoretical interpretation must also be unidimensional.

<u>CORRELATION VS. ANALYSIS OF VARIANCE</u>

One of the most telling criticisms of Fiedler's work is directed at his frequent use of statistically non-significant results [e.g., Graen *et al* (4)]. Basically, their criticism is well-founded since failure to abide by established criteria of statistical significance is to eschew one of the basic precautions against the inappropriate and inadvertent interpretation of effects due only to random variation. Nevertheless, it is suggested here that this criticism is too severe and that if Fiedler had utilized a more powerful statistical technique, such as the analysis of variance, many of his numerous research efforts might have yielded more significant findings than his reliance on rank order correlations has permitted.

Ironically, the model, as originally presented (1), is in the form of a factorial analysis of variance design with two levels each of leader position power, group atmosphere, task structure, and leader LPC. However, a major statistical problem presented itself in the process of compiling the many studies into a single theory. Group products from the various studies were usually not comparable and thus not directly amenable to the analysis of variance, hence Fiedler's extensive use of within-cell correlations, thereby using to his advantage the very nondimensional characteristic of correlations which will be described below as a major problem. The U-shaped curve, in a sense, reflects a four-way interaction among the variables. Fiedler was apparently willing to accept his intuition that the repeatedly appearing patterns of non-significant correlations did, in fact, reflect a real phenomenon and consequently was willing to ignore traditional standards of statistical reliability. He has consistently developed his theory of a higher order interaction between four variables by utilizing a series of usually independent tests of simple main effects without ever having recourse to testing the interaction, which is the crux of his whole model. It is certainly the hard way to prove a point.

A simple example of the additional power provided by the analysis of variance is demonstrated in a study reported by Shiflett and Nealey (10). The analysis of variance resulted in a significant interaction on group performance between LPC and leader position power for high ability groups. Neither of the two independent correlations between leader LPC and group performance, corresponding to the two simple main effects within that significant interaction, reached statistical significance (see Table 1). The analysis of variance clearly increased the power available to test the model. Another example is provided by Hardy's (12) data. Two analyses of variance, testing the model in octants 1 through 4, resulted in three of the four octants being supported by the data. Yet, independent correlations, presented earlier in Table 3, provide statistically significant support in only two of the octants, again demonstrating the increased power of the analysis of variance.

A study by Chemers and Skrzypek (23) provides an excellent example of the handicaps involved in using correlations instead of the analysis of variance. Their study has several valuable qualities: it tests the model in all eight cells; the methodology appears to have been carefully and adequately handled, and fairly closely reflects Fiedler's own definitions of the independent variables [see (6)]; the power manipulation includes a basis of legitimacy which is not usually available in laboratory studies, due to the nature of the subject population (military academy cadets); and the study clearly replicates the contingency model curve reported by Fiedler (2), using the Spearman ρ as the test statistic. Yet, not one of the reported correlations reached statistical significance. In order to permit some form of comparability between the structured and unstructured tasks in the present reanalysis, performance scores were transformed into T scores with $\bar{X} = 50$, and S.D. = 10. This procedure does not permit a test of the main effect of tasks within an analysis of variance design but does permit all other comparisons. These scores were then submitted to a 2 x 2 x 2 x 2 factorial analysis of variance with repeated measures on the task factor. A summary of this analysis is presented in Table 5. As can be seen, group atmosphere had a very strong effect on performance ($\underline{F}$ = 20.346, $\underline{p}$ < .001). Mean performance in groups with good leader-member relations was 55.15; while in groups with poor leader-member relations mean performance was 44.85. This one effect accounted for 24% of the variance in performance scores. In addition, the 3-way interaction of LPC, group atmosphere, and task was significant ($\underline{F}$ = 6.187, $\underline{p}$ < .025) and indicates that low LPC leaders did substantially better than high LPC leaders on a structured task when group atmosphere was good and on an unstructured task when group atmosphere was poor. This interaction will be discussed again in the next section, but it is worthy of note that the set of eight statistically non-significant correlations, reported by Chemers and Skrzypek, actually are reflecting statistically significant effects accounting for about 28% of the performance variability.

TABLE 5

Summary of Analysis of Variance of Transformed Performance Scores From Small Groups of Cadets [Chemers and Skrzypek (23)]

Source	df	MS	F
Between Subjects			
LPC (L)	1	182.96	2.192
Power (P)	1	16.55	< 1
Group Atmosphere (G)	1	1,697.96	20.346*
L x P	1	.21	< 1
L x G	1	42.20	< 1
P x G	1	5.45	< 1
L x P x G	1	92.57	1.109
Error (Ss within groups)	24	83.45	
Within Subjects			
Task (T)	1	0.00	0.0
L x T	1	79.54	1.269
P x T	1	2.38	< 1
G X T	1	88.29	1.408
L x P x T	1	10.75	< 1
L x G x T	1	387.84	6.187**
P x G x T	1	26.54	< 1
L x P x G x T	1	60.00	< 1
Error (T x Ss within)	24	62.69	

*p < .001; ω^2 = .24
**p < .025; ω^2 = .04

The problem of statistical significance in hypothesis testing brings up an interesting problem regarding research strategies involved in testing of the model. Tversky and Kahneman (24) have noted that, in general, a replication sample should be larger than the original sample in order to assure that the replication be statistically significant when using an independent test of significance. They also note that as a replication sample shrinks in size relative to the original sample, the probability of obtaining a comparable significant result drops alarmingly fast. They present the example of a case where the original sample contains 20 observations, and the replication sample contains only 10 observations. The probability of replication of a significant finding in the original sample under these conditions is less than 0.50. The extremely high frequency of non-significant correlations used to establish the contingency model [see Fiedler (2), pp. 134-141] argues very strongly that replication and validation samples sizes must be substantially increased. In spite of this fact, validation studies have usually involved similarly and frequently distressingly small samples. A case in point is the study by Chemers and Skrzypek (23) in which the contingency model is clearly supported in all eight cells, yet not a single correlation is statistically significant, using samples of eight. Under these circumstances, failure to replicate findings at acceptable levels of significance should not even be surprising--and does not necessarily indicate the absence of an effect--since there is always the possibility of making a Type II error (failure to reject the null hypothesis).

The blanket indictment of Fiedler's use of non-significant results which has been handed down by Graen *et al* (4) is probably too severe. Nevertheless, the lack of statistical significance is a definite problem in Fiedler's research program and efforts are clearly needed to establish and abide by definitive statistical procedures. The present data reanalyses clearly support the recommendation by Graen *et al* (4) that the analysis of variance would represent a substantial improvement in statistical technique. Recent tests of a post hoc nature reported by Fiedler (6), which show that the patterns of correlations are significantly different from chance expectations, are probably a reflection of this otherwise untested, but clearly implied, higher order interaction within the model. However, post hoc tests of this type are subject to criticism, as pointed out by Graen *et al* (4). The acceptance of non-significant results may be permissible for the organization of extant results or for developing a theoretical model, but firm criteria of statistical reliability cannot be ignored when attempting to validate or modify that theory or hypothesis. Without these criteria, there is no acceptable or predetermined criterion for either accepting or rejecting the null hypothesis. It is for this reason that an adequate method must be found to test the model's implied higher order interaction. The analysis of variance seems imminently qualified.

BETWEEN-CELL COMPARISONS

An implication that has been derived from the model, and done so rather explicitly by Fiedler (25), is that mean performance is fairly stable across the dimension of favorableness. However, this implication is unwarranted, based on a strict interpretation of the correlation coefficients as used within the model. A correlation coefficient, by mathematical definition, does not reflect any unit of measure. Thus, the model is silent regarding comparisons of relative performance across cells, or for that matter, between any one point of the favorableness dimension and any other point. Substantial differences in mean performance could exist between any two points along the dimension which would be completely unreflected in the corresponding correlations, as long as the relationship between leader LPC and group performance did not change. In fact, data reported by Fiedler [(2), p. 259] show a steady decline in performance from favorable to unfavorable situations for an unstructured task and rather erratic variation for two structured tasks. Yet, Fiedler (25) has suggested "situational engineering" as an alternative to leadership training. In other words, change the situation (shift a group up or down on the favorableness dimension) to fit the leadership style. But, if mean productivity is not constant across situations, one might find that a leader who has the "wrong" leadership style in one situation may become even less effective when his group is changed to a second situation where, according to the model, his leadership style is "right." The use of performance means, within an analysis of variance design, would help substantially to eliminate possible confusions of this type, as well as provide additional useful information not now available from statistical tests of the model which involve only correlations.

In order to more clearly illustrate this problem, the mean transformed performance data from the Chemers and Skrzypek (23) study were plotted across the eight octants of situational favorableness, as shown in Figure 1. As can be seen, the means within each cell generally reflect the correlations reported by Chemers and Skrzypek, shown at the bottom of the figure. An exception occurs in octant 3 where the means imply a negative correlation [as predicted by Fiedler (2)], but the reported correlation is positive. Also, the rather small difference between the means in octant 5 does not seem to bear out the substantially larger reported ρ. Beyond these peculiarities, it should be obvious that a great deal more information is available. Performance generally declines as a function of situational favorableness (reflecting the significant main effect of group atmosphere reported in Table 5). It can also be seen that only in octant 4 are high LPC leaders noticeably superior to low LPC leaders. In the other seven octants, high LPC leaders are either substantially inferior to, or are approximately equal to, low LPC leaders. Rather surprisingly, high LPC leaders do just as poorly in very favorable circumstances (octants 1 and 2) as they do in the unfavorable situations.

These data demonstrate the difficulties inherent in attempting to "engineer a situation" without first considering mean performance. In fact, to the extent that the Chemers and Skrzypek data are replicable and generalizable to many situations, the curves shown in Figure 1 have several implications for the use of the model in "situational engineering"

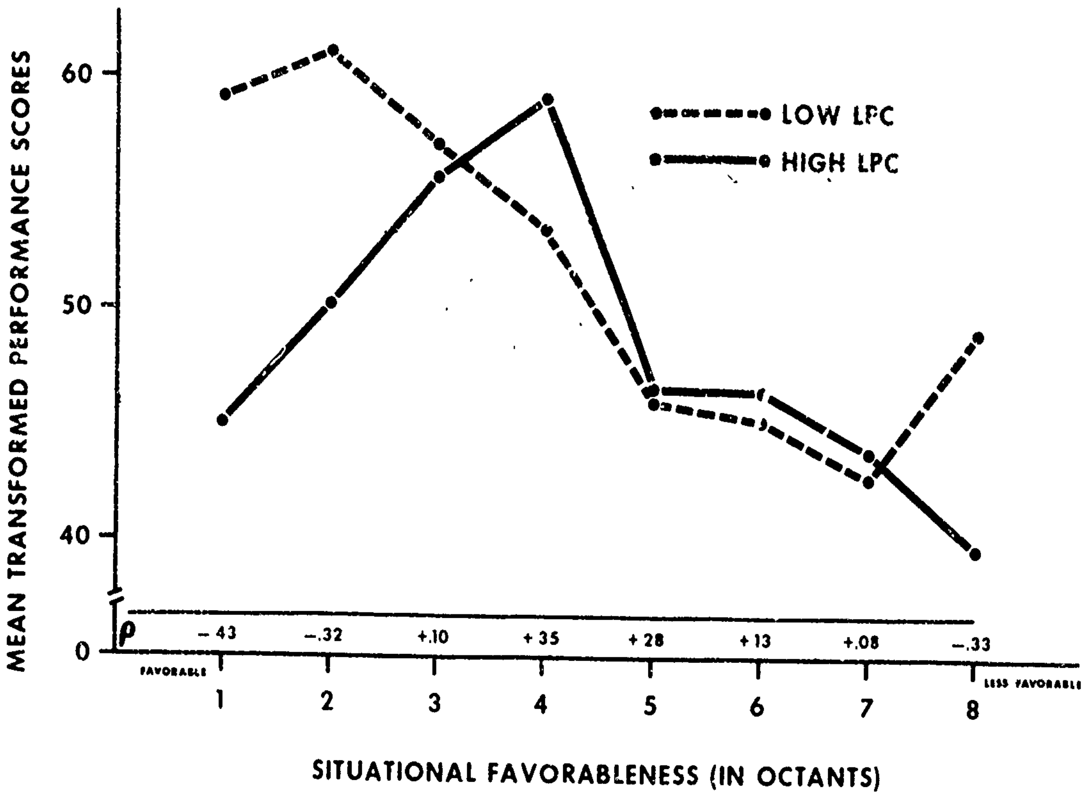

Fig. 1. Mean transformed performance data from small groups of cadets plotted along the situational favorableness dimension (23).

(25) or in assessing the effectiveness of training and experience (26). With respect to the latter approach, Fiedler has stated that "according to the Contingency Model, training should increase the favorableness of the situation and it should, therefore, increase the performance of one type of leader but decrease the performance of the other type" [(26), p. 115]. That is, Fiedler suggests that a high LPC leader who moves from an unfavorable situation to one of intermediate favorableness and on to a very favorable situation will see his performance go from poor to good and then become poor again. The low LPC leader would have the opposite effect occur: his performance would be good in unfavorable situations, poor in moderately favorable situations, and again good in very favorable situations. Notice that the high LPC curve in Figure 1 clearly supports Fiedler's hypothesis but the low LPC curve does not. In the case of the low LPC leaders there is a fairly linear relationship, where in unfavorable situations, performance is poor (although better than high LPC leaders), but in moderately favorable situations, performance has improved to an intermediate level of performance (although not as good as high LPC leaders), and in very favorable situations, performance continues to improve to a relatively high level, and much higher than that of high LPC leaders. Thus, the low LPC performance "function"

is not good-to-poor-to-good, but poor-to-intermediate-to-good. But both of these hypothesized functions, along with the high LPC function of poor-to-good-to-poor, would result in the same pattern of correlations, i.e., negative-to-positive-to-negative. It should be apparent at this point that any number of performance functions could occur and still reflect the same correlational function. Thus, knowledge of the corresponding correlations is not enough to make between-cell comparisons.

An "engineering" recommendation, based on the Chemers and Skrzypek data, might be to "keep the troops happy" and avoid hiring high LPC managers since they are generally a poor bet in terms of obtaining high levels of group performance. As for optimizing mean performance, situations should be designed to correspond to octants 3 and 4 (good group atmosphere, with an unstructured task), since mean performance here is high and differences in leader effectiveness as a function of LPC appear to be small. But Shiflett and Nealey (10) obtained large performance differences in these two octants as a function of leader LPC in interaction with group ability. Thus, the generality of these findings must be ascertained before clear-cut conclusions can be drawn. As a further example of possible restrictions on the generality of these data, a recent study by Csoka and Fiedler (27) indicated that there was little appreciable decline in performance across the favorableness dimension in several military settings.

A final comment on the Chemers and Skrzypek data is in order. The analysis of variance did not produce any significant effect involving leader position power. Thus, a more parsimonious explanation of the data would involve only three variables, not the four necessitated by a strict interpretation of the contingency model. Nevertheless, the present reanalyses do not in any way contradict the conclusions drawn by Chemers and Skrzypek nor do they contradict the basic assumptions or predictions of the contingency model. Rather, they do suggest new insights into the dynamics of the model, as well as indicating new methodological and statistical approaches to extensions of the model.

THE CONCEPT OF A SINGLE DIMENSION OF FAVORABLENESS

Fiedler's (1) initial presentation of the contingency model had three dichotomized dimensions collapsed onto a single dimension which is usually referred to as the favorableness of the group situation for the leader. Subsequent work [e.g., (2,6,28)] has involved an attempt to widen the definition of the dimension by assessing the "favorableness" of a number of variables such as group ability and cultural heterogeneity, then assigning appropriate favorableness weights and summing them to get a single index of favorableness. Most of this weight assessment, beyond the original three dimensions, has been done intuitively rather than having been guided by some formal set of postulates or rules. This technique has presented some problems, both methodological and theoretical, which have been discussed elsewhere (2,4,5,6). The purpose here is not to reiterate other critiques but to suggest that regardless of whether a single dimension of favorableness does or does not exist, evidence from the model itself implies that the unidimensional concept may have only limited usefulness, at least in terms of the use of LPC as an index of leadership style.

Earlier it was noted that the predictive ability, or the strength of the relationship between LPC and group performance, appeared to decline as favorableness decreased. This point will be expanded upon by asserting that the connecting of correlation median-points into the now-familiar inverted U-shaped curve directly implies that there are at least two areas along the favorableness dimension where LPC is completely unrelated to group performance. These areas exist around the point where the sign of the correlations changes, i.e., where the connecting line crosses the horizontal zero-correlation line between cells 3 and 4, and again near the center of cell 7. This assertion is based on the fact that r is a reflection of the strength of association between two variables. When the correlation is zero, the association and predictability are likewise zero. This effect is demonstrated in Figure 2. The broken line represents the absolute value of the median correlations reported by Fiedler [(2), p. 142]. The only difference between this figure and Fiedler's traditional inverted U is that negative correlations have been reflected onto the positive side of the graph. The solid line, enclosing

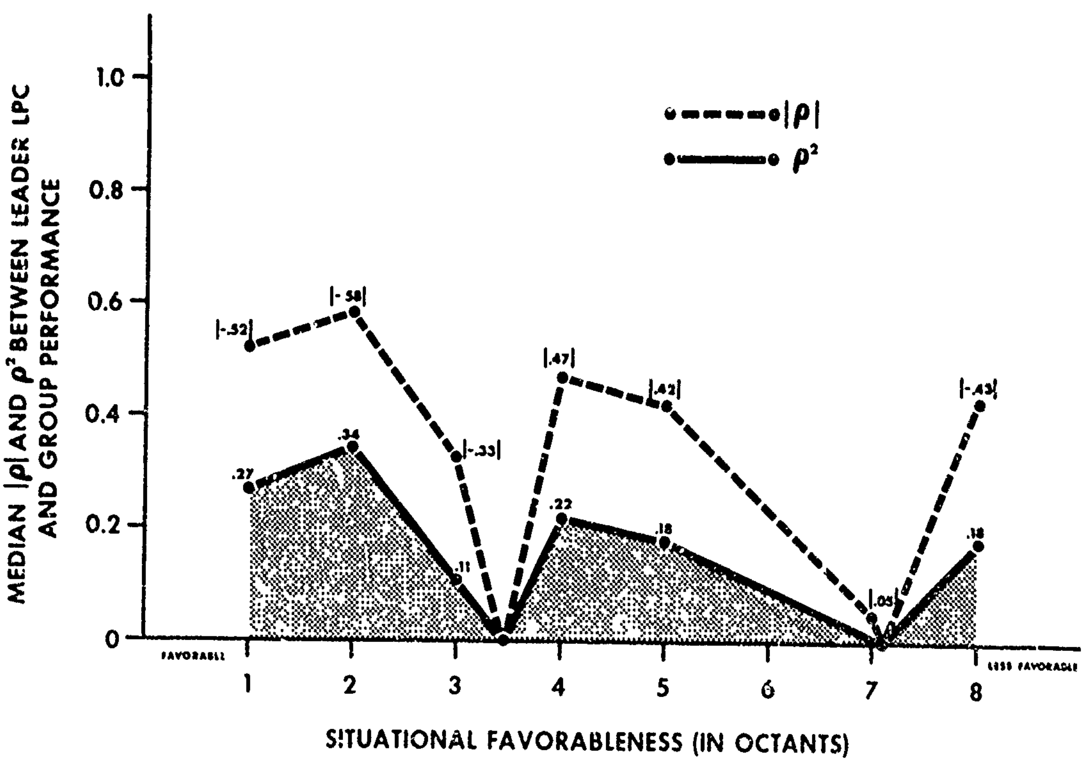

Fig. 2. The contingency model curve plotted across situational favorableness in terms of $|\rho|$ and ρ^2.

the shaded area, represents the degree of association, indexed by ρ^2, which is on an equal interval scale in contrast to $|\rho|$, which is not. Inspection of the figure shows the earlier noted decline in association between LPC and group performance as situational favorableness declines. The figure also shows the two areas of zero association wherein LPC presumably is completely unrelated to group performance. Since LPC is a key variable in the model, these areas represent "blank spots" in the theory, since nothing is said about them in any of Fiedler's work. Whether these are, in fact, blank spots depends on the validity of the unidimensional concept of favorableness and on the appropriateness of presenting data in this fashion. In other words, the appearance of blank spots may be nothing more than an artifact of the manner in which Fiedler has chosen to present his data. It seems rather unlikely to this author that a pattern of the type shown in Figure 2 exists in "real life." And to the extent that this is the case, the model itself argues against an attempt to organize the world along a single dimension. Whatever the case, the degree of association between LPC and group performance, as indicated in Figure 2, is an integral part of the model as it now exists, and includes two areas where the association is zero.

CONCLUSION

Recent attempts to extend and validate the contingency model [e.g., (2,6,28)] have aroused some concern due to the frequent use of statistically non-significant results and the increasing number of untested assumptions involved in attempting to incorporate a wide variety of variables into a single dimension of favorableness. While many of these problems have already been alluded to [e.g., (2-6,8,29,30)], few truly constructive suggestions have been made.

It has been shown here that LPC is more effective as a predictor of group effectiveness under some circumstance than under others. Furthermore, it has been shown that LPC may not be related to group effectiveness at all under some circumstances. In other words, LPC is not the panacea of small group research. What could be less startling! In this respect, then, LPC seems to be like most other variables. Recognition of this fact not only permits research to be directed at when (and why) LPC is related to group performance but also opens the door to research on when (and why) LPC is not related to group performance. Again, to be able to ask both questions requires a predetermined criterion for rejecting H_0. To suggest that LPC will be related to performance in some situations rather than others also implies a higher order interaction between LPC and other independent variables. The use of the analysis of variance has been strongly recommended as a means of increasing statistical power and for providing tests of both main effects and interactions. This, in turn, permits a determination of when LPC is or is not related to group performance. Further, the analysis of variance would permit not only within-cell comparisons of means but also across-cell comparisons. Measures of association such as ω^2 (7) are now readily accessible for use with the analysis of variance and should easily replace the use of r^2 and ρ^2 for the same purpose. In short, it is suggested that correlations have substantially outlived their usefulness within the framework of testing the contingency model. More powerful tools are available which provide even more information and with less ambiguity. Recent examples of attempts to test portions of the model using the analysis of variance include studies by Hardy (12), which substantially support the model, Shiflett and Nealey (10) which only partially support the model, and Graen *et al* (8,29) which generally do not support the model.

The attempt to squeeze all independent variables onto a single dimension of favorableness seems to be leading in a direction which, in the long run, will be nothing more than a dead end. Explaining the world by means of "favorableness to the leader," without clearly specifying in advance the procedures for defining "favorableness," reduces the capacity to generate meaningful research. Even if the relationships shown in the inverted U curve are valid across that favorableness dimension, favorableness may not even be the relevant dimension along which the LPC-performance relationship occurs. The situational favorableness dimension may only be correlated with the true dimension or dimensions along which that relationship occurs. The unusual curve shown in Figure 2 suggests that more than one dimension may be involved. Yet, these implications cannot be tested, and potentially correlated hypotheses cannot be partialled out as long as no well-defined set of rules exists for determining situational favorableness. In spite of an elaborate description of the definitional procedures used in the past (2), the application of these procedures is fairly difficult, especially when attempting to handle variables which have not previously been dealt with [see, for example, (25)]. O'Brien (3), recognizing this problem, suggested replacing Fiedler's "situational favorableness to the leader" with the concept of "leader influence" since leader influence can be defined mathematically within structural role theory (31,32,33). O'Brien and his associates (34) have recently demonstrated the feasibility of this experimental approach. This transition seems appropriate since Fiedler's (2) definition of situational favorableness incorporates the idea of the ease with which the leader can influence the group. Actually, Fiedler has already implicitly recognized the necessity of taking into consideration additional situational factors besides favorableness. He has distinguished between interacting, coacting, and counteracting groups (2) and, more recently (6), he has distinguished between laboratory and field studies and between training groups and those requiring previously learned performance. The statistical properties and problems of the model discussed above argue that it may be useful to carry this approach further by reconsidering the whole concept of a single dimension of favorableness.

Graen *et al* (4,8,29) have argued that the contingency model is not plausible because of lack of significant supportive evidence and that the model has lost its ability to direct research. While most of the criticisms of Graen and his associates are generally well-founded, it seems premature to simply reject either LPC or the model as not plausible. This seems akin to throwing out the proverbial baby with the proverbial bath water. Further, their assertion that the model has lost the capability of directing meaningful research can be true only to the extent that traditional research procedures designed to safeguard internal and external validity are ignored. A more reasonable treatment of the model would seem to be that of recognizing and avoiding its methodological and statistical weaknesses, and then to proceed to test, modify, and expand (or to eventually reject) the model with carefully controlled experiments, not with vitriolic diatribes.

LITERATURE CITED

1. Fiedler, F. E. A contingency model of leadership effectiveness. In: L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Vol. I, New York: Academic Press, 1964, pp. 149-190.

2. Fiedler, F. E. A Theory of Leadership Effectiveness. New York: McGraw-Hill, 1967.

3. O'Brien, G. E. Group structure and the measurement of potential leader influence. Aust. J. Psychol. 21: 277-289, 1969.

4. Graen, G., K. Alvares, J. B. Orris, and J. A. Martella. Contingency model of leadership effectiveness: Antecedent and evidential results. Psychol. Bull. 74: 285-296, 1970.

5. Mitchell, T. R., A. Biglan, G. R. Oncken, and F. E. Fiedler. The contingency model: Criticism and suggestions. Acad. Manag. J. 13: 253-267, 1970.

6. Fiedler, F. E. Validation and extension of the contingency model of leadership effectiveness: A review of empirical findings. Psychol. Bull. 76: 128-148, 1971.

7. Hays, W. L. Statistics for Psychologists. New York: Holt, Rinehart and Winston, Inc., 1963.

8. Graen, G., J. B. Orris, and K. M. Alvares. Contingency model of leadership effectiveness: Some methodological issues. J. Appl. Psychol. 55: 205-210, 1971.

9. Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1956.

10. Shiflett, S. C. and S. M. Nealey. The effects of changing leader power: A test of "Situational Engineering." Org. Behav. Human Perform. 7: 317-382, 1972.

11. Hunt, J. G. Fiedler's leadership contingency model: An empirical test in three organizations. Org. Behav. Human Perform. 2: 290-308, 1967.

12. Hardy, R. C. Effects of leadership style on the performance of small classroom groups: A test of the contingency model. J. Person. Social Psychol. 19: 367-374, 1971.

13. Mann, R. D. A review of the relationships between personality and performance in small groups. Psychol. Bull. 56: 241-270, 1959.

14. Jacobs, T. O. Leadership and Exchange in Formal Organizations. Alexandria, Va.: Human Resources Research Organization, 1971.

15. Stogdill, R. M. and A. E. Coons (Eds.). Leader Behavior: Its Description and Measurement. Research Monograph No. 88, Ohio State University, Columbus, Ohio, 1957.

16. Katz, D. Survey Research Center: An overview of the human relations program. In: H. Guetzkow (Ed.), Groups, Leadership and Men: Research in Human Relations. New York: Russell and Russell, Inc., 1951, pp. 68-85.

17. Yukl, G. Toward a behavioral theory of leadership. Org. Behav. Human Perform. 6: 414-440, 1971.

18. Graen, G., F. Dansereau, and T. Minami. Dysfunctional leadership styles. Org. Behav. Human Perform. 7: 216-236, 1972.

19. Yukl, G. Leader LPC scores: Attitude dimensions and behavioral correlates. J. Soc. Psychol. 80: 207-212, 1970.

20. Shiflett, S. C. Stereotyping and esteem for one's least preferred coworker. USAMRL Report (in preparation).

21. Fox, W. M., W. A. Hill, and W. H. Guertin. Dimensional analysis of the Least Preferred Co-Worker scales. J. Appl. Psychol. (in press).

22. Fiedler, F. E. Personality, motivational systems, and behavior of high and low LPC persons. Technical Report No. 70-12. Seattle, Wash.: Organizational Research Laboratory, University of Washington, 1970.

23. Chemers, M. M. and G. J. Skrzypek. An experimental test of the contingency model of leadership effectiveness. J. Person. Soc. Psychol. (in press).

24. Tversky, A. and D. Kahneman. Belief in the law of small numbers. Psychol. Bull. 76: 105-110, 1971.

25. Fiedler, F. E. Engineer the job to fit the manager. Harvard Business Review, September 1965, pp. 115-122.

26. Fiedler, F. E. Predicting the effects of leadership training and experience from the contingency model. J. Appl. Psychol. 56· 114-119, 1972.

27. Csoka, L. S. and F. E. Fiedler. The effect of leadership experience and training in structured military tasks—a test of the contingency model. Org. Behav. Human Perform. (in press).

28. Fiedler, F. E. The effect of leadership and cultured heterogeneity on group performance: A test of the contingency model. J. Exp. Soc. Psychol. 2: 237-264, 1966.

29. Graen, G., J. B. Orris, and K. M. Alvares. Contingency model of leadership effectiveness: Some experimental results. J. Appl. Psychol. 55: 196-201, 1971.

30. Fiedler, F. E. Note on the methodology of the Graen, Orris, and Alvares studies testing the contingency model. J. Appl. Psychol. 55: 202-204, 1971.

31. Oeser, O. A and F. Harary. A mathematical model for structural role theory, I. Human Relations, 15: 89-109, 1962.

32. Oeser, O. A. and F. Harary. A mathematical model for structural role theory, II. Human Relations, 17: 3-17, 1964.

33. Oeser, O. A. and G. E. O'Brien. A mathematical model for structural role theory, III. The analysis of group tasks. Human Relations, 20: 83-97, 1967.

34. O'Brien, G. E., A. Biglan, and J. Penna. Measurement of the distribution of potential influence and participation in groups and organizations. J. Appl. Psychol. 56: 11-18, 1972.