DISCREPANCIES BETWEEN FACTOR ANALYSIS AND MULTIVARIATE DISCRIMINATION AMONG GROUPS AS APPLIED TO PERSONALITY THEORY

By

LOWELL H. STORMS, B.A., M.S., Ph.D.

Institute of Psychiatry, Maudsley Hospital, University of London

IT is clearly advantageous to the personality theorist to know which of his measures covary and to what extent. The results of factor analyses can be useful in summarizing some of the patterns of covariation and classifying response measures. An indefinitely large number of sets of factors may equally well describe the same set of data, however, and, before any choice of one particular set of factors, dimensions, principal components, etc., can be considered as revealing basic intervening variables or constructs in a personality theory, independent evidence must be provided. It is the purpose of this paper to reveal some of the pitfalls in making such a choice by showing how concentration on one set of factors obtained by analysing a multivariate scatter can lead to a serious loss of information when applied to the analysis of group differences in the same set of data and can suggest interpretations which are at odds with the full evidence.

Professor H. J. Eysenck is the most notable example of a personality theorist who makes use of a variety of techniques of multivariate analysis to investigate a set of hypothetical dimensions, so we will compare our results with some of his applications of the same, or similar techniques (2, 3, 4, 5).

The techniques which were employed for the present purposes were factor analysis, discriminant function analysis, and canonical variate analysis. The factor analysis of the same date we used for our calculations is reported by H. J. Eysenck in two articles (2, 4). The data were collected for a Ph.D. thesis by Dr. H. P. Hildebrand (6), who did the original factor analysis, but the analyses to be presented here are additional to those he performed.

METHOD

Subjects

The psychoneurotic Ss, 145 in all, were male voluntary in-patients at Belmont Hospital, Sutton, Surrey. Patients were not included in the sample if they were under 18 or over 50 years of age, showed any evidence of psychotic or neurological disorder, had I.Q.s below 90, or were undergoing E.C.T. or insulin therapy. The patients were diagnosed by psychiatrists with varying points of view, and the diagnosis made by the psychiatrist in charge of a case was not altered by consultation or the pooling of judgments. The sample breakdown by diagnosis is as follows:

Anxiety states	25
Obsessives	10
Depressives	10
Hysterics	25
Psychopaths	20
Mixed neuroses	55

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

Further details regarding the neurotic sample can be obtained from Hildebrand (6, 7).

The normal group consisted of 22 adult males living in the London area selected by Eysenck and used by him in preference to the normal group used by Hildebrand (5, p. 29). Hildebrand's normals, 25 soldiers from an army re-allocation centre, were not used because they were apparently quite unrepresentative of normal adult males (6, p. 120).

Tests

Hildebrand gave twenty-two tests to all his subjects. The test battery included questionnaires, intelligence tests, and what Eysenck calls "objective behaviour tests". These are described in Hildebrand's thesis (6). For present purposes, a set of six tests and one of ten were selected from the twenty-two. These are listed below with the abbreviations which we shall use throughout.

Set 1	Set 2
± Maudsley Medical Questionnaire (MMQ)	? Maudsley Medical Questionnaire
+ Guilford R Scale (R)	Guilford R Scale
± MMQ Lie Scale (L.S.)	MMQ Lie Scale
± Annoyance Test (Ann.)	Annoyance Test
± Guilford S Scale (S)	Suggestibility Reversals (Rev.)
± Guilford T Scale (T)	Leg Persistence (L.P.)
± Summed Guilford D and C Scales (D±C)	
+ Porteus Mazes Weighted Wrong Directions (WWD)	
+ Porteus Mazes Crossed Lines (X line)	
Mill Hill Vocabulary (MHV)	

The six tests marked ± were chosen for Set 1 because they had the highest loadings on Hildebrand's neuroticism dimension; the three marked + were chosen because they had the highest loadings on his introversion-extraversion factor; the remaining test was chosen in order to provide a check on the effects of intelligence. Five of the tests in Set 2 were chosen because they were the only tests among the 22 which yielded F tests significant at the 5 per cent. level (or better) for discrimination among the six neurotic groups after transformations had corrected for heterogeneity of variance. The sixth, marked with a ?, was chosen as a possible contributor to the discriminant functions and because it was desired to have as many tests as groups.

Thus, Set 1 was selected to represent the main factors of the Hildebrand analysis, while Set 2 was picked according to the empirically observed discriminations yielded by the single tests. As a result, we have two groups of tests which can be compared for combined power of discrimination between pairs of neurotic groups.

The analyses for Set 1 were performed using all the Ss (N=167); the analyses for Set 2 were done using only the six neurotic groups (N=145), since these were the only groups to whom all six tests of this set were administered.

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

The Analyses

Three kinds of multivariate analysis were applied to the data for each set of tests. They are as follows:

1. Discriminant functions comparing each group with every other group. A convenient account of the method is given in Rao (10, p. 246–252). Twenty-one such comparisons were performed for Set 1 (seven groups of Ss) and fifteen for Set 2 (six groups). All discriminations were tested for significance using the analysis of variance technique described by K. S. Mather (8, p. 156). The percentage of total variance accounted for by the difference between means for each pair of groups was also computed.

2. The Hotelling principal components method, which can be found in Thompson (11, p. 108–126), was used to analyse the matrix of within groups variances and covariances for both sets of tests. This method is a form of factor analysis and uses units rather than estimated communalities in the diagonal of the correlation matrix.

3. A canonical variate analysis (10, p. 365–370), essentially a principal component analysis of the multivariate dispersion between groups, was performed on both matrices.

Our interest will centre primarily upon the comparative power of the two sets of tests as reflected in the discriminant functions and upon the relationship between the principal components and the canonical variates obtained for each set of data. This relationship constitutes a comparison of dimensions resulting from analysing within group and between group matrices (indeed, the principal components method was employed because of its comparability with canonical variate analysis).

These comparisons, and other aspects of the analysis, will be discussed and related to Eysenck's use of these methods.

TABLE I

Discriminant Function Significance Tests and Percentages of Variance Accounted for by Group Differences

Pairs of Groups	Set 1		Set 2	
	F Tests	Per cent. Variance	F Tests	Per cent. Variance
Anxiety—Obsessive	1·160	7·18	1·588	6·69
Anxiety—Depressive	1·024	6·39	0·708	3·09
Anxiety—Hysteric	2·275*	13·17	5·153†	18·86
Anxiety—Psychopath ..	2·883†	16·12	4·218†	15·99
Anxiety—Mixed	1·145	7·09	1·570	6·33
Anxiety—Normal	7·731†	34·01		
Obsessive—Depressive ..	1·226	7·56	1·089	4·68
Obsessive—Hysteric ..	1·806	10·75	2·501*	10·14
Obsessive—Psychopath ..	1·154	7·14	2·076	8·56
Obsessive—Mixed	0·784	4·97	1·077	4·63
Obsessive—Normal	3·824†	20·32		
Depressive—Hysteric ..	0·819	5·18	2·922*	11·65
Depressive—Psychopath ..	2·904†	16·22	3·610†	14·01
Depressive—Mixed	1·249	7·69	0·892	3·87
Depressive—Normal	3·834†	20·36		
Hysteric—Psychopath ..	2·259†	13·09	5·665†	20·35
Hysteric—Mixed	1·179	7·28	4·479†	16·81
Hysteric—Normal	5·947†	28·39		
Psychopath—Mixed	2·451†	14·05	2·932*	11·68
Psychopath—Normal ..	7·316†	32·79		
Mixed—Normal	0·103	0·68		

* Significant at 5 per cent. level. † Significant at 1 per cent. level.

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

TABLE II

Loadings of the Tests on the Canonical Variates (L) and Latent Roots (l)*

	Set 1						Set 2				
	L_1	L_2	L_3	l_1	l_2	l_3	L_1	L_2	l_1	l_2	l_3
1. MHV ..	·2658	·7398	−·1916	−·0780	−·5727	1·0000					
2. MMQ ..	1·0000	−·4204	·8087	·9124	·0027	−·4127	·0848	−·2664	1·0000	+·0017	−·1065
3. Lie scales	·1080	−·9173	·5732	−·4988	−·1657	−·4904	·6220	−·5970	−·6921	−·3970	−·7406
4. Ann. ..	−·3145	·9261	·4815	·7179	·0398	−·4041	·2727	+·8454	·8713	−·0100	−·3826
5. S	−·3845	−·4425	−·4801	·8675	−·3773	·0461					
6. T	·0625	·4506	·0005	·8127	−·1044	·4885					
7. D+C ..	·3222	·1273	−·8747	1·0000	·1495	−·1718					
8. R	·0321	1·0000	·3892	−·3954	·8218	−·1691	·5270	+1·0000	1·4827	·1404	+·5627
9. WWD ..	·1074	·5510	−·2264	·2332	1·0000	·1391					
10. X lines ..	−·1608	−·4441	1·0000	·2101	·9334	·6456					
11. L.P. ..							−·8543	+·1886	·2150	−·6102	+1·0000
12. Rev. ..							1·0000	−·1153	−·0688	1·0000	+·2336

* For maximal comparability, all latent vectors have been scaled so that the highest loading equals unity.

RESULTS

The F ratios showing how well pairs of groups are differentiated using discriminant function weights are presented in Table I on page 715, along with percentages of variance accounted for by differences between means of weighted scores.

Although the two sets of tests have four tests in common, it can be seen that among the neurotic groups (the only basis for comparison since the normal group was not used in the analysis for Set 2) the set with only two additional tests selected empirically yielded eight out of a possible fifteen significant discriminations, while the set with six additional tests selected to represent factors, yielded five. The percentages of variance accounted for by group differences are very slightly higher for Set 1, but where significant differences were found, Set 2, the shorter test battery, has a higher percentage in half the cases. (As in the case of multiple correlations, the addition of useless predictors almost always increases the percentage of variance accounted for and may at the same time reduce the statistical significance.)

From the Set 1 data, 4 significant principal components and 2 significant canonical variates were extracted. Set 2 matrices produced 4 significant principal components and 2 significant canonical variates. The significance test for principal component residuals is given by Bartlett (1), and the test for the canonical variates by Rao and Slater (9, p. 25).

It can be seen in Table II that the loadings on the common tests for the two sets of principal components and the two sets of canonical variates do not closely resemble one another, except that the first principal components are quite similar. This illustrates the differences that can arise from the use of different groups and different test batteries, but in the present case the portions of the effect due to tests and to samples cannot be separated.

A better comparison can be made between the principal components and the canonical variates obtained using the same Ss and the same tests. The patterns of loadings in Table II are seen to be different for the principal components as compared with the canonical variates for both sets of tests. For example, the L_1 loadings rank correlate ·19 with 1_1's for Set 1. The most legitimate comparisons that can be made, however, are the correlations between the subject's factor scores and canonical variate scores obtained by weighting test scores by the proper loadings. These correlations are presented in Table III.

TABLE III

Intercorrelations Between Principal Component Factor Scores and Canonical Variate Scores

Canonical Variates	Principal Components			
	11	12	13	14
L_1	·544	−·005	−·153	·016
L_2	·193	·374	·280	−·680
L_3	−·271	·553	−·157	−·050
L_4	−·073	−·620	·187	−·167

It is apparent that no correlation is high enough for the two dimensions involved to be considered identical. It is also evident that the dimensions which correlate most strongly are not necessarily the ones which came out in the same stages

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

of the analysis. Thus, for example, the first canonical variate and the first principal component correlate less highly than the second canonical variate and the fourth principal component. This indicates that in the present data the major dimensions describing the within groups scatter are not the same as those describing differences between groups. Even for the dimensions which correlate most strongly, the patterns of test loadings are widely divergent, as can be seen in Table II.

Discussion

The major conclusions of the present analysis are those drawn from the discriminant functions and from the comparison of principal components with canonical variates.

1. Selection of tests according to their loadings on supposedly relevant factors for the purpose of discriminating groups may yield results inferior to those which can be obtained from tests which are selected so as to provide as much information as possible. A shorter test battery selected on practical grounds may be more efficient than a long one with supposed dimensional relevance. Other influences than those represented by the factors may be entirely overlooked using the latter approach. This seems to have been what happened with the Eysenckian dimensions used in this analysis.

Because discriminant function analysis and multiple regression are analogous, it is apposite to mention that Watts and Slater found similar discrepancies between the results of factor analysis and multiple regression (12, p. 48–59).

2. The primary dimensions of covariation within groups may not be the same as the primary dimensions of between groups covariation. Indeed, there is no mathematical reason to expect that they will be, though theoretical considerations regarding the phenomena involved might lead to an expectation of such a correspondence. Eysenck seems to have propounded some such hypotheses about his dimensions. He found it necessary to rotate his canonical variates to make them consistent with his dimensions of neuroticism and psychoticism (3). He writes, "The original variates, Y_1 and Y_2, cannot be interpreted as corresponding to these factors as their position is not invariant". There is nothing to indicate that his new variates, Z_1 and Z_2, are any more invariant; at least Y_1 and Y_2, are invariant in the particular sample under consideration. Rotated canonical variates are no longer canonical variates because they no longer describe the major dimensions of the between groups multivariate scatter (strongly implying that they were orthogonal, he also did not point out that his rotated variates were correlated in his sample). Thus, Eysenck's findings seem to support our conclusion that within groups and between groups dispersions may yield different major dimensions.

From a practical standpoint, it is advantageous if the within groups and between groups dispersions lie along different dimensions. The two dimensional illustration in Figure 1 shows that, other things being equal, different major dimensions of within groups and between groups dispersion may allow more accurate discrimination of the groups than is the case when they are the same.

Although Eysenck will not find anything inconsistent in the finding of several principal components in each analysis, since he does not claim that his psychoticism, neuroticism, and extraversion-introversion dimensions cover the whole range of human variation, the emergence of three significant canonical variates from the analysis for Set 1 and two from the analysis of the Set 2 data poses a problem for him. He has tried to justify the use of the hysteric-

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

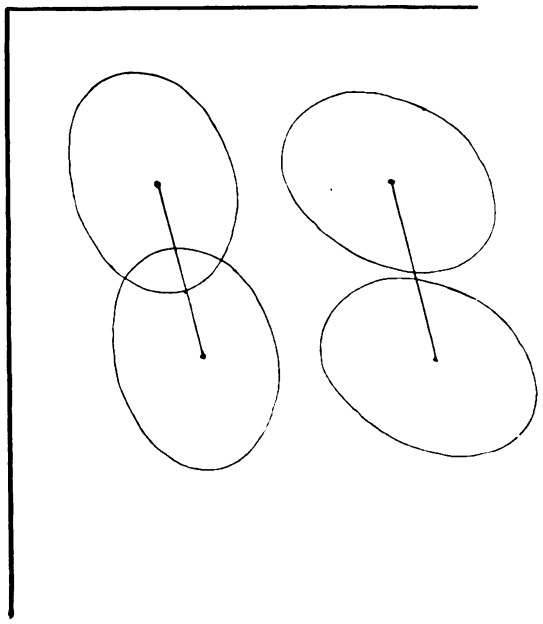


FIG. 1.—Comparison of between and within groups dispersions according to whether they have same or different major dimensions.

psychopath and dysthymic groups as criteria for extraversion and introversion, respectively (4, p. 22). Also, in his new book (5, p. 26) he makes this statement: "There is considerable evidence, which has been reviewed in great detail in *The Structure of Human Personality* (Eysenck, 1953), that the traditional psychiatric syndromes in the field of neurosis are generated by two factors, continua, or dimensions; namely, those of neuroticism and of extraversion-introversion." Discussing Hildebrand's results, he goes on to say that neurotic groups should be differentiated from normals on neuroticism and from each other in terms of extraversion-introversion (I–E). Although he further states: "It is not suggested that the three dimensions of psychoticism, neuroticism and extraversion-introversion . . . are the only ones which play a part in abnormal behaviour; it is highly probable that further research will succeed in unearthing other major dimensions", he implies that the important differences among neurotic groups are accounted for by different degrees of extraversion (5, p. 30). Our analyses make it clear that "other major dimensions" were implicit in the Hildebrand data all the time. The present results suggest not only that the major differences between neurotic groups lie on a "dimension" different from I-E, but also that there are more sources of between-group differences than Eysenck has indicated. These results were found using data from tests the selection of which Hildebrand describes thus: "Two groups of tests were chosen, one of which had been shown by previous work to differentiate between normal and neurotic groups and the other which previous work suggested having some loading on the I-E factors" (6, p. 18).

Hildebrand computed factor scores for his groups and found that hysterics and anxiety states were significantly differentiated (5 per cent. level). Eysenck takes this as support for his theory (4, 5), which seems justified as far as it goes. However, the fuller analysis reported here shows that the two groups are more efficiently discriminated by the six tests in Set 2, and the dimension defined by

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

the discriminant function weights bears little resemblance to introversion-extraversion (also, two tests, suggestibility reversals, which Hildebrand found to have a negligible loading on I-E, and Guilford's R scale, which had the highest I-E loading, were each found to differentiate the same two groups at the 1 per cent. level—reasonably enough, these two tests had the highest discriminant function weights). It appears that the data contain more relevant information than that summarized by the factor loadings.

A finding that Eysenck may find difficult to reconcile with his theory or with his practice of combining the hysteric and psychopathic groups (4, 5) to form a criterion group for extraversion is that psychopaths and hysterics were the two groups most significantly discriminated using both sets of tests. This is not surprising in view of the previously unreported fact that three of the tests used significantly discriminated hysterics and psychopaths ($p < \cdot 01$ in each case). These tests are suggestibility reversals, MPI lie scale, and leg persistence. All three are included in Set 2, only the lie scale in Set 1, so it was to be expected that Set 2 would discriminate the groups more significantly than Set 1. It should be noted that this discrimination also occurred using the same tests (those of Set 1) which are supposed to define the neuroticism and I-E dimensions. The groups in question differ either on these dimensions alone or on some other combination of characteristics. Eysenck has reported studies using these two groups pooled in varying proportions and calling them neurotic extraverts. This would be defensible if the groups differed only on extraversion, which is not the case. If they differ on neuroticism or other characteristics, such differences will enter into the differences between the pooled group and other groups in varying degrees from one experiment to another. Furthermore, the differences between the hysteric and psychopath groups may reflect dimensions or phenomena which have been overlooked and are worthy of study.

Summary

To illustrate some of the problems arising from attempts to relate results of factor analyses to differences among groups, discriminant functions, principal component analyses, and canonical variate analyses were performed using the data from two sets of tests which were administered to six neurotic groups and a normal group (the tests had been included in a larger test battery). One set consisted of ten tests selected primarily on the basis of high factor loadings on neuroticism and introversion:extraversion factors computed with the same subjects and the larger battery of tests. The other set consisted of six tests selected primarily because they gave the largest F ratios for differences among neurotic groups. Four of the tests were common to the two sets.

It was found that the smaller set of tests yielded more significant discriminations than the larger in the discriminant function analysis. Principal component analyses of the within groups variances and covariances resulted in a similar first "dimension" in each set of tests, but the other dimensions were not alike. An important discrepancy found in this analysis was that for neither set of tests did the major dimensions of the within groups scatter (principal components) correspond with the major dimensions of the between groups dispersion (canonical variates).

The results seem to indicate that adherence to dimensions derived from factor analyses when analysing differences among groups can lead to a serious loss of information and may lead to inefficiency in practical applications or over-simplification in theoretical interpretations. It was shown that dimensions describing between groups differences need not be the same as those appearing within groups and that it is better for purposes of discriminating groups if they are not.

On the basis of the results of our analysis, Eysenck's interpretation of a canonical variate analysis which he performed and his pooling of hysterics and psychopaths to form criterion groups for the study of extraversion were called into question.

Acknowledgments

The author expresses his gratitude to Mr. Patrick Slater for helpful suggestions and a critical reading of the manuscript and to Professor Eysenck for generously providing access to his data on the normal group and to the facilities of the Psychology Department at Maudsley Hospital.

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press

References

1. Bartlett, M. S., "Tests of significance in factor analysis", *Brit. J. Psychol. Stats. Sec.*, 1950, **3**, 77–85.
2. Eysenck, H. J., "Cortical inhibition, figural after-effect, and theory of personality", *J. abnorm. soc. Psychol.*, 1955, **51**, 94–106.
3. *Idem*, "Psychiatric diagnosis as a psychological and statistical problem", *Psychol. Rep.*, 1955, **1**, 3–17.
4. *Idem*, "A dynamic theory of anxiety and hysteria", *J. Ment. Sci.*, 1955, **101**, 28–51.
5. *Idem*, *The Dynamics of Anxiety and Hysteria*, 1957. London: Routledge and Kegan Paul.
6. Hildebrand, H. P., "A factorial study of introversion-extraversion by means of objective tests", 1953. Ph.D. Thesis, Univ. of London.
7. *Idem*, "A factorial study of introversion-extraversion", *Brit. J. Psychol.* In press.
8. Mather, K. S., *Statistical Analysis in Biology*, 1946. London: Methuen and Co.
9. Rao, C. R., and Slater, P., "Multivariate analysis applied to differences among neurotic groups", *Brit. J. Psychol., Stat. Sec.*, 1949, **2**, 17–29.
10. *Idem*, *Advanced Statistical Methods in Biometric Research*, 1952. New York: John Wiley and Sons.
11. Thompson, G. H., *The Factorial Analysis of Human Ability*, 1951. London: Univ. of London Press.
12. Watts, A. F., and Slater, P., *The Allocation of Primary School Leavers to Courses of Secondary Education*, 1950. London: Newnes Educational Publishing Co.

https://doi.org/10.1192/bjp.104.436.713 Published online by Cambridge University Press