*Brit. J. Psychiat.* (1976), 128, 404–7

# Multivariate Analysis: Data Reduction and Treatment Evaluation

By ALLAN WILSON

**Summary.** In a recent article, Everitt (1975) discussed several problems with multivariate techniques. However, two useful applications of multivariate techniques were not covered. The present paper describes the use of factor analysis to reduce a large array of outcome variables to a statistically manageable number, and multivariate analysis of variance to determine the relative effectiveness of several treatment regimes where a single outcome variable cannot be specified. It is concluded that the advantages of a multivariate approach outweigh the disadvantages, provided the researcher is careful in interpreting and reporting his results.

## Introduction

In a recent paper, Everitt (1975) outlined some of the problems associated with multivariate techniques. The ratio of number of variables to number of observations is certainly a much neglected parameter in the design of research, and the availability of a large number of factor rotations makes rotated factor analytic solutions excessively arbitrary. However, two very important aspects of multivariate analysis were not discussed in the Everitt article; the use of factor-analytic techniques for data reduction and the multivariate analysis of variance. These techniques can be of considerable value in assessing treatment effectiveness where the outcome criteria are not unequivocally indicated and where the criteria may include a number of psychological variables. Discussion with several colleagues who are responsible for evaluating psychiatric programmes revealed that the Everitt article had left them with an essentially negative impression of multivariate procedures—an impression I am sure the author did not intend beyond the use of factor-analytic techniques.

## Critera (Dependent Variables) for Treatment Evaluation

Unlike industrial organizations, where relevant outcome criteria can usually be specified in such concrete terms as production units, cost-benefit ratios, or simple revenues, mental health care organizations are usually hard-pressed to specify relevant criteria. For example, an alcoholism treatment programme might measure outcome by the number of days a patient ingests any alcohol during a given post-treatment period. Although this variable is probably valid, at least consensually, there are many other variables which might also be relevant to determining the overall effectiveness of the programme. Social adjustment, family adjustment, work history, attitudes toward self and others are just a few of the broad areas from which outcome variables might logically be drawn.

The question facing the researcher is which of the many possible variables to select. In the present illustration, selection of the single variable 'days drinking in the last six months' would almost certainly result in the criticism that there are many other variables of equal or greater importance in assessing the effectiveness of treatment of alcoholics. One solution is to include a number of dependent variables for which there is some consensual validity, and use multivariate techniques to organize the resulting complex of data.



https://doi.org/10.1192/bjp.128.4.404 Published online by Cambridge University Press

### Comparison of Treatment Regimes

In many treatment evaluation efforts, the effectiveness of one or more experimental procedures, treatments, or programmes is compared to a control or comparison programme. No-treatment controls are preferable, but for practical and ethical reasons the previously most effective treatment is usually employed as a quasi-control. If such experimental methods are used, between-group comparisons must be made, and the most powerful tool for such comparisons is the analysis of variance (or *t* test if only two treatments are involved). This, of course, assumes that the data are interval- or ratio-scaled and normally distributed. Adequate approximations of these criteria can usually be assured by careful selection and measurement of the outcome variables.

The relative effectiveness of the various treatments can be determined by analysis of variance followed by suitable multiple comparisons, simple effects, etc., to specify the nature of any significant between-treatment effects. However, univariate analysis of variance cannot be used if more than one dependent variable is being considered, as calculation of more than one *F* statistic reduces the confidence level. Thus, if multiple outcome criteria are to be considered, a multivariate analogue of the *F* or *t* test must be employed.

### Multivariate Analysis of Variance

Univariate analysis of variance provides a test of the significance of mean differences for two or more groups. In cases where a number of criteria (dependent variables) are relevant to testing the hypotheses of interest, a vector of two or more means is obtained for each group. Multivariate analysis of variance is the extension of analysis of variance to the multiple mean situation. An overall test of the differences in mean vectors for the various treatments is made, and any significant multivariate effects can be followed by univariate tests of the individual dependent variables. In this way multivariate analysis of variance provides a tool for determining the significance of differences between treatments when more than one outcome criterion is used.

### Sample Size for Multivariate Analysis of Variance

As advocated by Everitt, ten observations per dependent variable is a good rule of thumb for multivariate techniques. In addition, for multivariate analysis of variance eight to ten observations are desirable for each treatment under consideration. Thus, in a three-treatment comparison with ten dependent variables 240 to 300 observations would be required. As pointed out by Everitt, large samples are necessary for multivariate techniques. However, the potential gain in ability to detect significant differences between experimental samples is very great and certainly worth the extra effort required to obtain adequate samples.

### Multivariate Paradigm for Treatment Evaluation

The following example of the utility of combining factor-analytic and multivariate analysis of variance procedures is drawn from an ongoing programme evaluation of a hospital-based alcoholism treatment facility.* The hospital in which the unit is located was asked to develop more effective rehabilitation methods for skid-row alcoholics. Thus, the evaluation paradigm consists of a comparison of an experimental, hospital-based programme to the aggregate of treatment programmes already established in the community. Of course, the paradigm could be adapted to other areas of treatment evaluation.

Patients were randomly assigned to either the experimental programme or to another programme in the community. Prior to assignment to a programme all patients were required to complete a 24-item questionnaire containing questions about drinking history, family adjustment, social adjustment, work adjustment, and attitudes toward self and others. A parallel form of the questionnaire was administered six months after discharge from treatment to measure changes over time on the 24 items, and other parallel forms will be administered at 12- and 18-month intervals.

Before constructing the questionnaire, the literature pertaining to outcome evaluation of

* A pre-publication draft describing the facility, the programme, and the evaluation methodology is available from the author.

https://doi.org/10.1192/bjp.128.4.404 Published online by Cambridge University Press

treatment effectiveness was reviewed and it was found that a number of theoretical orientations have guided the selection of outcome criteria. Most studies have employed some measure of post-treatment drinking behaviour, but many have also argued for changes in working patterns, family adjustment, social adjustment and self-concept as important indices of the effectiveness of treatment. On the basis of these precedents, 24 items from these areas were incorporated into a questionnaire to be administered by an experienced counsellor.

The data collected before assignment to treatment and six months after discharge from treatment consisted of 24 dependent variables (responses to items) obtained from 100 patients randomly assigned to one of two treatment conditions. The problem was to determine the significance of differential changes over time on these 24 dependent variables and thus infer the relative effectiveness of the two treatment programmes. The obvious statistical technique (since the variables were designed to yield interval data and were approximately normally distributed) was multivariate analysis of variance. However, we have already noted that approximately ten observations are required per dependent variable per treatment condition—a total of 480 observations—and only 100 observations were available.

*Data reduction by factor analysis*

The 24 dependent variables were reduced to a statistically manageable number by principal components analysis of the pre-treatment data. The questionnaire had been administered to 300 sequential admissions to the unit, with the 100 experimental patients embedded as a block in the middle of this sample. Thus the 24 dependent variables were analysed by recourse to 300 observations on the assumption that the characteristics of the experimental sample would be accurately reflected in the larger sample. An orthogonal rotation was used to give a mathematically unique solution which is free of much of the criticism of excessive arbitrariness rightly attributed by Everitt to most rotated factor-analysis solutions. The rotation gave six factors with eigenvalues greater than 1·0 (an arbitrary but generally accepted retention criterion) accounting for about 65 per cent of the variance.

With an acceptable amount of the information for the 24 items reduced to six factors, it was then necessary to summarize each of the six factors for subsequent analysis. The technique employed was to take the item with the highest factor loading on each factor as being representative of the factor, compute a pre-treatment/post-treatment difference score for that item and submit the difference scores to subsequent analysis. It should be noted that some information is lost by such a reduction procedure, but careful examination of the percentages of variance accounted for by the factors and items will indicate whether the procedure is viable for a given case.

Alternative methods of summarizing the factors might involve combining the $n$ items with the highest factor loadings on a given factor, retaining the $n$ items with the highest loadings on a given factor without combining them (here impossible because of the sample size), or various methods of combining factor scores. Such combinative methods are, unfortunately, likely to result in an optimistic bias to subsequent tests of significance, and for this reason were rejected in the present example.

*Multivariate analysis of variance of pre-treatment data.*

Subsequent analyses of variance were to be carried out on the pre-treatment/post-treatment difference scores. Difference scores were used in preference to covariance techniques because they are conceptually simpler. However, if difference scores are to be used it is essential to demonstrate that the treatment groups are homogeneous with respect to the variables of interest.* Although random assignment and

* There are statistical arguments for using covariance techniques in preference to analysis of difference scores under certain conditions. Provided the distribution of low pre-treatment scores is uniform for the treatment groups, any bias introduced by using difference scores is likely to be insignificant (see Lord, 1963). This is particularly true where the relative overall changes in treatment groups are of primary interest. The data for the present example were also analysed by multivariate analysis of covariance, and the results were identical to those obtained by multivariance analysis of variance of the difference scores.

https://doi.org/10.1192/bjp.128.4.404 Published online by Cambridge University Press

adequate sample size should ensure this, a multivariate analysis of variance was computed to compare the two groups on the six-variable pre-treatment data. No significant between-group differences were obtained, indicating that the six variables were uniformly distributed in the two groups before treatment.

*Multivariate analysis of variance of difference scores*

The difference scores for the six variables were then submitted to multivariate analysis of variance to determine the significance of differences between the two treatment groups. In this study, a significant multivariate *F* indicated that the treatment groups did undergo differential changes as measured by the six-item vector. With a significant multivariate effect (and only with a significant overall test), univariate *F* tests can be computed separately for each of the variables. In the present case, it was found that the patients in the two treatment groups differed at a highly significant level on the pre-treatment/post-treatment change in number of days drinking, and that the difference favoured the established programmes. The other five univariate tests showed no significant between-treatment differences.

## Interpretation of Results

As any competent statistician will point out, questions can be raised about most applications of statistical procedures to data obtained in applied research. The onus falls on the researcher to ensure that his results are not optimistically interpreted. In the present example, one might question the method of representing a factor by a single item, but the researcher can minimize the problem by carefully assessing the contributions of the representative items to their respective factors. The reliability of the factor analysis as a data reduction technique might also be queried, as for logistic reasons a reliability check was not included in the example. However, data obtained from subsequent samples from the same population are currently being analysed in the same way, and the reliability will be determined by comparison of the factor structures obtained from the two samples.

The potential gain from such a multivariate approach to treatment evaluation far exceeds any disadvantages. If consistent methodology is used in different treatment settings or in the same settings at different times, comparability of results will be achieved. The multivariate approach to treatment evaluation facilitates such a systematic approach, and frees the researcher of the highly-restrictive necessity of specifying *a priori* a single outcome criterion. However, with multivariate techniques, results must be interpreted cautiously, and, as Everitt points out, replication assumes even greater importance.

## References

Everitt, B. S. (1975) Multivariate analysis: the need for data, and other problems. *British Journal of Psychiatry*, **126**, 237–40.

Lord, F. M. (1963) Elementary models for measuring change. In *Problems in Measuring Change* (ed. C. W. Harris). Madison: University of Wisconsin Press.

A. Wilson, Ph.D., *Assistant Professor of Psychiatry, University of Manitoba, 770 Bannatyne Avenue, Winnipeg, Canada, R3E 0W3*

(*Received 6 May 1975*)

https://doi.org/10.1192/bjp.128.4.404 Published online by Cambridge University Press