Brit. J. Psychiat. (1979), **134**, 516–21

Problems of Repertory Grid Analysis and a Cluster Analysis Solution

By PETER GOODGE

SUMMARY Statistical and requirement difficulties encountered with existing methods of Repertory Grid Analysis are considered. A simple purpose-built method of Cluster Analysis is outlined, and its extension to the difficult problem of comparative analysis described.

Kelly's Repertory Grid technique (Kelly, 1955) is essentially a questionnaire method of empirically determining the structure of an individual's perceptions. The technique has been extensively used in psychiatric settings (e.g. Ryle, 1975; Fransella and Adams, 1966). Market researchers have used Repertory Grids to determine how differing products are perceived, and increasing interest is shown by organizational researchers (e.g. Pirani and Reynolds, 1976).

The Repertory Grid technique begins with a list of objects, termed 'elements'. Elements can be persons, roles, products, organizational problems, etc. The number of elements is usually between 8 and 25. The respondent's task is to choose an adjective (or a pair of adjectives of opposite meaning) that describes one (or more) of the elements. Each adjective (or adjective pair) is termed a 'construct'. An effective limit is put upon the number of constructs by the respondent's ability to think up genuinely different ones. Each element is now rated or ranked according to how much of each construct quality it has, and the respondent's judgements placed in the cells of a matrix. Table I shows a conventional form of Repertory Grid.

Kelly's original form of Repertory Grid differed somewhat from that described here. Kelly used his 'triples' method to elicit constructs, and employed a binary system of scoring. Kelly's variety of grid seems unpopular to-day, probably because of the relative complexity of the triples method and because of the advantages of greater freedom of response with rating schemes. Kelly's binary scoring method also suffers from methodological problems (Bannister and Mair, 1968; Benjafield and Adams-Webber, 1975).

Repertory Grids do present formidable problems of data analysis. Completed Repertory Grids contain a large amount of complex data, and the difficulties of somehow statistically managing such data do seem to have deterred the wider utilisation of grid techniques. There are accepted methods of data analysis, such as Principal Components Analysis, but, as I hope to show later, these are not very useful. It is also true that problems of analysis are made significantly more difficult when we wish to compare two or more Repertory Grids.

In this paper I should like to review the problems of Repertory Grid analysis and to outline a cluster technique which seems to overcome many of the difficulties.

Problems of data

The difficulties we encounter when analyzing Repertory Grids are really of two types: (i) limitations imposed by the nature of the data itself and (ii) the adequacy and utility of the results provided by particular types of analysis. I shall begin by considering the question of data limitations.

Let us suppose we have ratings in the cells of the Repertory Grid, as in the Grid in Table I. Ratings constitute an 'ordinal' form of data; meaning that such data correspond to the rough, subjective estimates of respondents and

TABLE I

Example of Repertory Grid focusing on marital relations

Constructs	Elements							
	Mother	Father	Wife	Self	Ideal self	Ideal wife	Ex-flame	Male friend
Nervous	4	3	3	3	1	1	4	2
Determined	3	4	2	4	4	3	3	2
Shy	4	3	5	3	2	2	3	2
Thoughtful	3	3	2	4	3	3	2	2
Happy	4	3	4	3	5	4	3	4
Sexy	2	2	2	4	4	4	3	3
Unfaithful	2	1	3	2	2	1	4	2
Free	3	2	3	3	4	4	3	4

are in no way a 'measure' of anything. A rating of '4' is not somehow worth twice as much as a rating of '2'. With ratings we may not perform any of the usual arithmetic operations of addition or multiplication. For example, an 'average rating' is meaningless and it would be erroneous to determine one. The important ordinal character of ratings data seems to have been overlooked by a number of researchers who have erroneously used orthodox statistical techniques on ratings data. For example, a two-factor Principal Components Analysis performed on our example Repertory Grid is diagrammatically given in Diagram 1. There are a number of oddities in the factor analysis results; FATHER is rated in the Grid as less nervous than at least two other persons, but the factor analysis would have us believe he is the most nervous. Similarly, MALE FRIEND is placed as the most THOUGHTFUL, yet rated in the Grid as one of the least THOUGHTFUL. Anomalies such as these are relatively frequent in the factor analysis of Repertory Grids.

It is difficult to know what to do with ratings-type data. It is possible to perform some elementary non-hierarchical cluster analysis, but this would not give very satisfactory results. Perhaps the easiest way to put ratings into a manageable form is to rank order them. Ranked data is more convenient, but the price of transforming from ratings to ranks is often a considerable loss of information.

With rankings in the cells of the Repertory Grid analysis is a little easier. Rankings constitute 'interval'-type data, and we may add and subtract ranks. I would suggest that we might also find an average of ranks as this is a meaningful statistic. The multiplication of ranks is an invalid operation, however; hence we may not determine product-moment correlations with ranked data. We may legitimately find rank correlation coefficients. In summary, then, it appears that we need our Repertory Grid data in rank order form and we may only use rank order correlations as measures of association.

The above considerations also severely limit the extent to which we may compare two or more Repertory Grids. Some researchers have added, or subtracted, values from the same cells of different Grids and analyzed the resulting matrix. In this way it was hoped that compound Grids made up of several individuals' Grids could provide data on a collection of individuals. Similarly, it was thought that a

https://doi.org/10.1192/bjp.134.5.516 Published online by Cambridge University Press

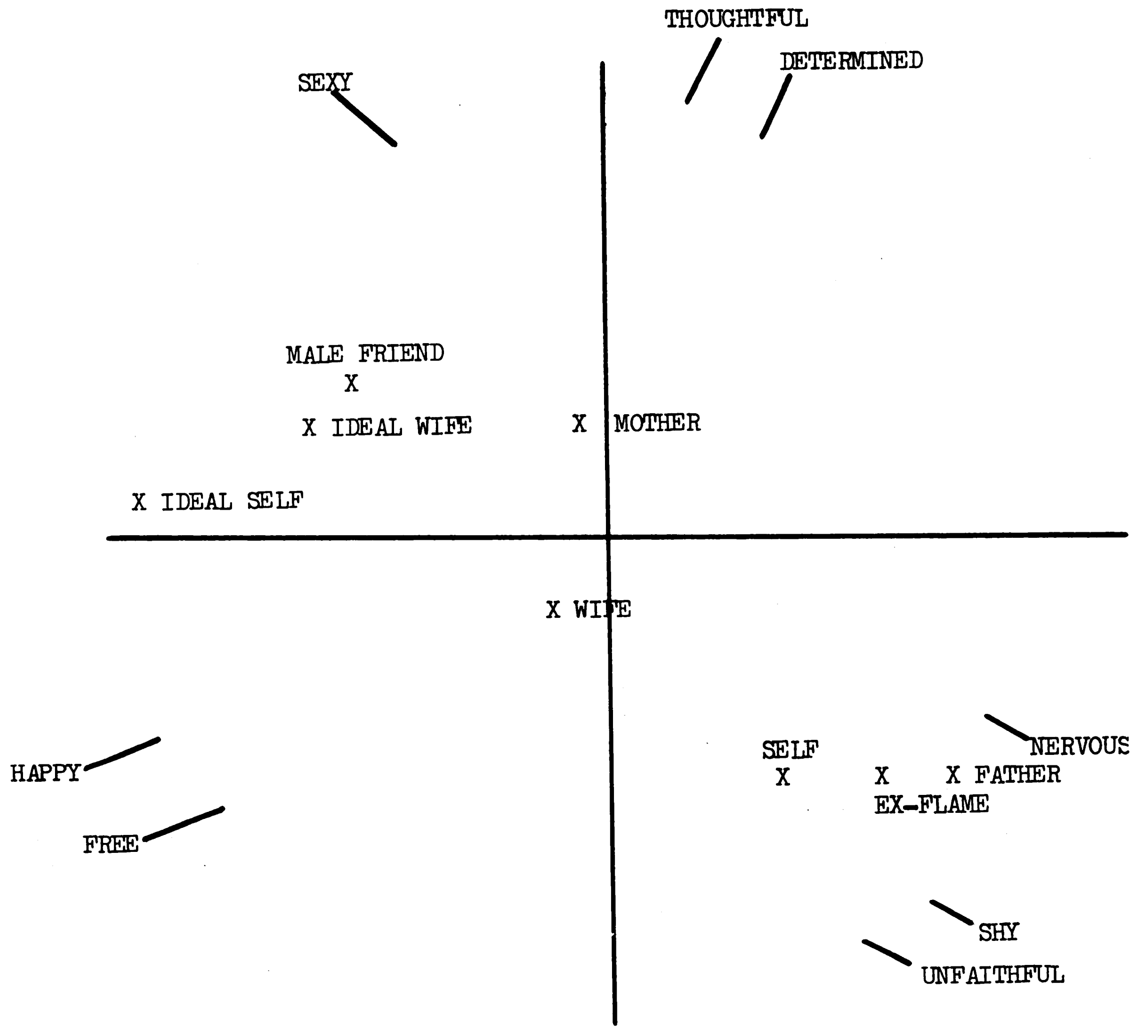

DIAGRAM 1.—Factor analysis of the example grid.

'differences Grid' formed by subtracting only two Grids could be used for comparative analysis. Clearly, the adding and subtracting of Grids with ratings-type data is invalid—such figures are meaningless. Such operations may be permissable with ranked data, but this does constitute a rather crude form of analysis. In addition, this approach to comparative analysis seems methodologically doubtful, since it is inappropriate to have to use the same set of provided constructs for several persons (see Caine and Smail, 1967). What is needed would seem to be a method of comparatively analyzing Repertory Grids on which respondents would be free to choose their own constructs.

Additional minor problems arise with missing data. It seems wrong to pressure respondents to 'fill in' a cell if they find it inappropriate to apply a construct to a certain element. It is also erroneous for the researcher to 'guess' the values of missing data. The absence of data should, I think, itself, be treated as data especially as the work of Landfield (1976) suggests that a relative abundance of missing entries is associated with personal problems. An adequate method of Repertory Grid analysis, therefore, needs to take some account of missing data.

Adequacy of analysis

The user of Repertory Grids is faced with a choice of employing a computationally simple method of analysis or using a more elaborate technique, which certainly necessitates the use of a computer. Both simplistic and elaborate methods raise questions about the adequacy of the results they provide.

Less sophisticated methods, such as those described by Bannister and Mair (1968), tend to make simplifying assumptions which can result in misleading results. In addition, not all the results that we might require can be provided with such a method. For example, Bannister's method (Bannister and Mair, 1968) presumes the orthogonality of the two constructs that explain the most variance, and this might imply erroneous representation of the data. Bannister's method does not provide figures to plot the elements in relation to the structure of the constructs.

Different problems are encountered with factor analysis. Here the major problem is how to describe verbally the derived factors. For example, Diagram 1 shows a rotated two-factor solution of our example Grid, but the factors are very hard to label. It seems to me that it is wrong for the researcher to apply his verbal descriptions to some other person's construct systems, as this presumes that the constructs have the same meaning for researcher and respondent. It is also probably difficult for the respondent to find high-level constructs with which to label the results of factor analysis. One of the advantages of cluster analysis is that the clusters are effectively self-labelling.

An adequate form of analysis must then avoid over-simplifying assumptions and present the results in a comprehensible form.

A method of analysis

Based upon the above considerations, I have adapted a method of cluster analysis to be suitable for Repertory Grid data. It is hoped that several of the problems encountered with other methods will be less prevalent with this technique.

I shall assume that all the Grid data have been rank-ordered row-wise (i.e. construct-wise), and shall employ rank-order correlation coefficients as a measure of relationship between constructs. Rank-order correlations are very quick to calculate.

The essence of the analytic method is Forgy's method of cluster analysis (Anderburg, 1973). Forgy's procedure defines each cluster by a 'seed point'; in the case of Repertory Grid data, a seed point will consist of a row of numbers similar to the rows of the Grid. I shall define the construct as the adjective(s) associated with each row of the Grid and the rows of the Grid as construct row vectors, in order to distinguish between the semantic label and the numeric row vector. The procedure is an iterative one involving clustering (i.e. associating) each construct row vector of the Grid data with the seed point with which its absolute correlation is greatest, then finding new seed points as the arithmetic means of the row vectors in each cluster, and then clustering again. The procedure terminates when an iteration produces no change in the clustering (i.e. each construct row vector remains in the same cluster). More explicitly, the method involves:

Step 1: Choose k initial seed points. There are several ways of choosing initial seeds (see Anderburg, 1973), but simply selecting k construct row vectors from the Grid data works quite well.

The number of seed points, k, is a matter of personal choice. For normal Grid data, two or three seeds (and, hence, the same number of clusters) are usually sufficient.

Step 2: Rank-order correlate each construct row vector with each seed point, and cluster each construct with that seed with which the magnitude of its correlation is greatest. Ignore the sign of the correlation coefficient at this stage.

Step 3: If one or more of the construct row vectors has now changed its cluster membership, then go on to Step 4, else terminate the clustering.

The number of iterations required is usually very small, say less than four, although this figure will increase with the number of constructs in the data and the number of seeds used.

Step 4: We now have k clusters. Find a new seed point for each by taking the mean row vector of the constructs in the cluster. Where a construct row vector correlates negatively with its seed, we need to 'reflect' its rankings, i.e. the highest rank becomes the lowest, the lowest becomes the highest, and so on. Failure to do this means that the new seed will not be found in a logically consistent manner. 'Reflected' rankings are only used in the calculation of new seeds. The iterative procedure now continues by going back to Step 2.

This is a relatively simple and quick procedure, usually involving much less arithmetic calculation than the computation of a correlation matrix. It provides a clustering of constructs with correlations with each seed rather like 'factor loadings'. It also gives element scores on the cluster seeds in terms of rank scores.

This form of cluster analysis applied to the example Grid required only two iterations when using two arbitrarily chosen construct row vectors of the Grid as the initial seeds. Using other seeds gave very similar results. The results are given in diagramatic form below.

The axes are not perpendicular because a small correlation was found between the cluster's centroids (i.e. final seeds). It is questionable if clusters ought to be orthogonal, as we cannot really expect construct subsystems to be totally independent. Indeed, correlations between cluster centroids is a useful indicant of

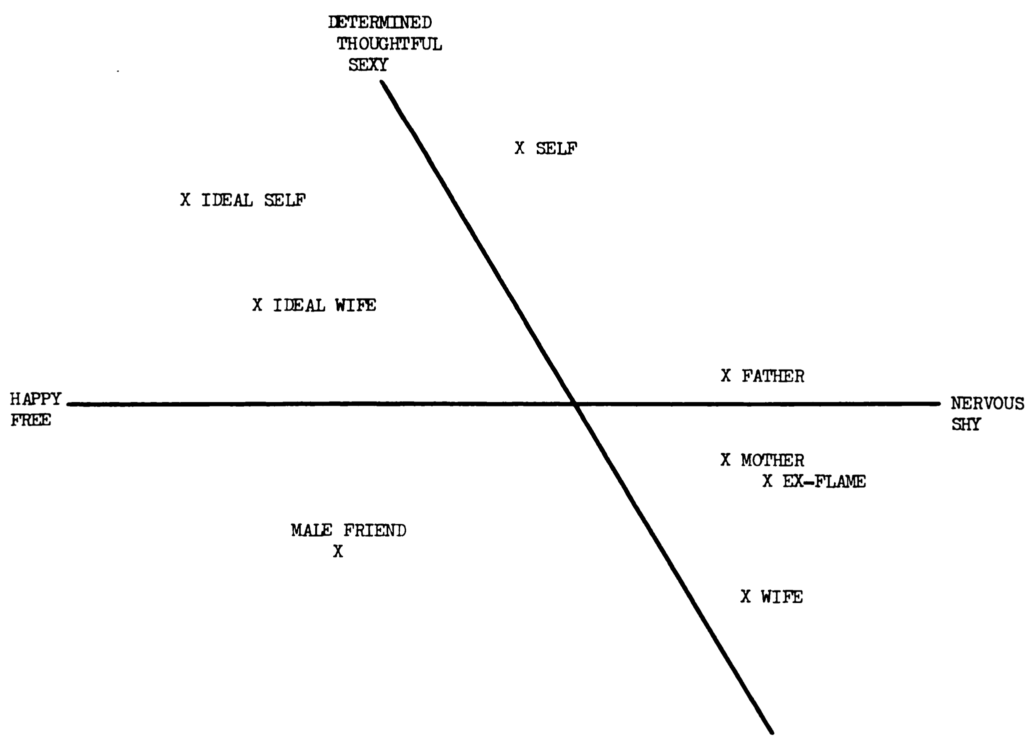

DIAGRAM 2.—Cluster analysis of the example grid.

how dependent construct subsystems (as represented by clusters) are.

If we have a construct row vector with missing data units (i.e. a cell is empty), then we may still find rank order correlations and cluster the construct. When we correlate the construct row vector with another we imagine that the corresponding cells in the other seed row vector are also missing and reduce the number of data units (elements) in our calculation. Ranking and the calculation proceeds normal. We cannot use a construct row vector with missing entries in finding cluster centroids. For this reason, too much missing data will produce anomalous results.

With this method it is perhaps best to determine a number of solutions using different number of clusters (say between one and three), and then to choose that solution which appears best. The quality of a clustering may be judged by examining the absolute values of the correlations between construct and seed row vectors. Ideally every construct should correlate highly with its cluster seed. If high construct-seed correlations are possible with fewer clusters, then presumably this would be a more preferred solution. Should a set of constructs correlate poorly with their seeds, this would suggest more clusters (i.e. seed points) are needed. The simplicity of this method makes it feasible to obtain two or three cluster solutions with the use of a pocket calculator.

Comparative analysis

As we have seen, accepted methods of Repertory Grid analysis do not extend very well to comparative analysis. The new method described above has much less of a problem in performing comparisons.

Assuming we have two Repertory Grids with the same elements but different constructs, then we may parcel all the constructs together and cluster analyze the compounded Grid. The results of such an analysis would permit us to determine the similarity and differences between the Grids, as judged by the extent to which they share the same clusters. Relations between the constructs of the Grids are also examinable.

With more than two Repertory Grids it becomes impractical to pool all the constructs together into the same cluster analysis. This does not produce much simplification of the data. Quite a useful method of comparative analysis may, however, be performed by completing a second cluster analysis on the final seed points derived from separate cluster analyses of all the Repertory Grids. In detail this would involve determining one or two clusters for each Grid by the method described above, and then pooling all the final cluster seed points into a second cluster analysis. If we take the seed points of the cluster analysis of each Grid as some numeric summary of the Grid itself, then our second cluster analysis gives us a representation of how clusters from different Repertory Grids are related. We might, for example, use this method to determine how the construct subsystems belonging to the members of a family or group were tied together.

References

Anderburg, M. R. (1973) *Cluster Analysis for Applications*. Academic Press.

Bannister, D. & Mair, J. M. (1968) *The Evaluation of Personal Constructs*. Academic Press.

Benjafield, J. & Adams-Webber, J. (1975) Assimalative projection and construct balance in the Repertory Grid. *British Journal of Psychology*, **66**, 169-73.

Caine, T. M. & Smail, D. J. (1967) Personal relevance and the choice of constructs for the Repertory Grid technique. *British Journal of Psychiatry*, **113**, 517-20.

Fransella, F. & Adams, B. (1966) An illustration of the use of the Repertory Grid technique in a clinical setting. *British Journal of Social and Clinical Psychology*, **5**, 51-62.

Kelly, G. A. (1955) *The Psychology of Personal Constructs*. New York: W. W. Norton.

Landfield, A. (1976) A personal construct approach to suicidal behaviour. In Slater, P. (Ed.) *Explorations of Intrapersonal Space*. Volume I. Wiley.

Pirani, M. & Reynolds, J. (1976) Gearing up for safety. *Personnel Management*, February, 25-9.

Ryle, A. (1975) *Frames and Cages*. Sussex University Press.

Peter Goodge, B.A., M.SC., *Computer Evaluation Team, City General Hospital, Newcastle Road, Stoke-on-Trent, Staffordshire*

(*Received 29 December 1977*)