TECHNICAL NOTE

SELECTING A "TYPICAL" CURVE

When presenting the results, in any study involving the experimental analysis of behavior, it may be desirable to select the performance of a single organism to represent or be typical of a class of behavior. This can be justified on the grounds that the averaged curve does not in fact show how any real organism behaved. The counter argument is that there is room for subjective bias to operate in the selection of a particular record. However, important information could be lost, or never found, if only averaged curves were studied (see Sidman, 1952, and Estes, 1956, for a theoretical discussion of some of the problems and Hayes, 1953, for some practical illustrations). The need to present an individual record is agreed. It is the process by which it is selected that is open to argument.

The meaning of the word typical in the present context is, "serving as a representative of a class", and the point at issue is, representative in what way? The selection of something typical may be carried out in two very different ways, *i.e.*, using different criteria. Firstly, for didactic use a typical curve may be selected because it shows in extreme or simplified form something which is not always so clear-cut. This implies of course that the underlying processes are well known and understood (at least by the teacher) and that this "typical" case is seen by all to be an exaggeration or simplification of what is usually found. This is a common use of the word typical. Secondly and what is more usually understood by a typical case is where, for purposes of exposition or argument, a typical case is displayed because (a) it is typical in the sense that it most resembles all the other cases studied and (b) useful information would be lost by presenting all the data as an averaged curve. Briefly, argument can arise because a curve presented as satisfying the second criterion but on subjective grounds only, is always open to the charge that consciously or unconsciously it has been selected according to the first criterion when there are no grounds (*i.e.*, full understanding of the process) for doing so. To put it another way, where a typical curve is selected on some unstated subjective basis, the suspicion in the eye of the beholder will always be that it has been selected because it shows strongly certain tendencies which the experimenter feels to be important but which were not present to the same degree in the rest of the data. Such curves could thus be typical of the preconceptions of the experimenter, rather than of the class of behavior studied by him.

It seems clear that there is a need for an objective method of selecting the typical record from a family of records, in that of all the records, it most nearly represents a consensus in terms of configuration. If a consensus in terms of mean values but still retaining an individual curve was wanted, some sort of least squares procedure could be used to find which curve had the least sum of squared deviations from the mean curve. However, this curve could have all the disadvantages of the mean curve and none of the advantages of the individual curve, *i.e.*, such a procedure could select the curve which was actually least typical in shape in some situations. It seems that what is wanted is a method of selecting a typical curve which satisfies the second meaning of "typical" given above and is objective, *i.e.*, not in any way subject to bias on the part of the experimenter.

One way of doing this is to adapt the method quoted by Thomson (1951, p. 253) and previously used by Burt (1936, quoted in Thomson) for weighting the marks given by different examiners. Burt intercorrelates examiners over candidates and finds the typical examiner to be the one with the highest sum of intercorrelations, *i.e.*, the one who most resembles all the others. Curves over occasions could be intercorrelated to arrive at the same result. This procedure produces a typical curve which preserves all the virtues of the individual curve and is not open to the objection of biased selection.

Given a large number of organisms over a great many trials and no access to a computer, this would be a very time consuming procedure. However, if one does have access to a computer then, apart from punching time, only a few minutes work is involved. Once the computer has arrived at the table of correlations, the organism with the highest sum of correlations could be found (1) by hand if the matrix is not very large, (2) by the computer in a matter of seconds by writing a small program especially for this purpose, (3) by using any of the standard Hotelling-type factor analysis programs (*e.g.*, as in Horst, 1965) to extract the first factor from the matrix. The organism with the highest factor loading on the first factor will then be the one required. The second method described above is recommended if one makes frequent use of this technique, the third if one wishes to try it once to see how it works.

Given below is a miniature example using real data (from Pollard and Lysons, 1967) showing how this technique can be used. The problem is to pick the

typical record from the performance of five cats on a closed field maze problem. The error scores are as in Table 1.

Table 1
Error Scores of Five Cats on Problem 6

Runs	Cats				
	1	2	3	4	5
1	194	136	12	138	40
2	16	0	0	148	18
3	24	52	0	46	0
4	22	0	0	44	0
5	6	0	0	38	0
6	18	0	0	54	0
7	0	0	0	0	0
8	18	0	0	24	0

The intercorrelations of these five performances are shown as Table 2.

Table 2
Intercorrelations of Five Performances

	1	2	3	4	Sum
1					3.44
2	94				3.21
3	99	93			3.40
4	61	53	58		1.78
5	90	81	90	06	2.64

As can be seen from Table 2, the performance of cat No. 1 would be selected as the typical performance. There will be instances where the mean curve, if included as another organism, would be selected as

the typical record. This can happen when the mean curve is like most of the individual curves in shape. If one wishes to present a mean curve this could be used as a test of whether it is appropriate to do so. This use of correlation coefficients may be unusual but it is not invalid. Correlation here is taken to be simply a convenient index of concomitant variation and nothing more. It is not always realized that as such it carries with it no assumptions concerning distributions. However, it is a necessary requirement of correlational procedures that values must be paired on some logical basis. In some cases such a basis might prove difficult to find. In other cases the labor involved in reducing the data to a convenient form for correlations might be too great. Hence, this technique is not put forward as a general solution for all cases.

REFERENCES

Estes, W. K. The problem of inference from curves based on group data. *Psychological Bulletin*, 1956, 53, 134-140.
Hayes, K. J. The backward curve: a method for the study of learning. *Psychological Review*, 1953, 60, 269-275.
Horst, P. *Factor analysis of data matrices*. New York: Holt, Rhinehart & Winston, 1965.
Pollard, J. S. and Lysons, A. M. Possums in the closed field test. *Animal Behavior*, 1967, 15, 129-133.
Sidman, M. A note on functional relations obtained from group data. *Psychological Bulletin*, 1952, 49, 263-269.
Thomson, G. H. *The factorial analysis of human ability*. London: University of London Press, 1951.

H. F. PRIEST
J. S. POLLARD
University of Canterbury
Christchurch, N.Z.