PATTERN ANALYSIS[1]

James R. Swanson
Florida State University

This paper is a discussion of another method of predicting single criteria from multiple predictors: the method of pattern analysis. Pattern analysis is a sort of family name applied to a group of techniques which seeks to find relationships not clearly evident when traditionally-used scoring or correlational methods are used. Within this general classification, two major subclasses can be identified. One is termed "profile analysis," which deals with ordered sets of data such as test score profiles or some other pattern of standing on a linear continuum. The other is called "configural scoring," which is directed toward finding sets or groups or configurations of test items, or similar elements of unordered data, in the hope that these groups will have a predictive validity which cannot be found in the complete units which they form.

The major emphasis of this paper is profile analysis. The use of configural scoring is limited in ordinary educational prediction at the present time, for reasons which will be discussed. However, it is a method of extreme value in clinical work, and may prove to be even more worthwhile in the field of educational prediction as the methods and means of handling and analyzing the data come into more general use. It will be of some value, then, to discuss configural scoring in order that its possibilities as well as its limitations can be seen.

One of the earliest references to pattern analysis was made by Zubin in 1936, when he stated "pattern analysis leads to fractionating a group into several types, characterized by specific patterns or syndromes." This seems to express the idea of an empirical device which provides a means of classifying individuals on the basis of a certain pattern of responses, in comparison with other individuals who obtained the same pattern.

The worth of this idea has been borne out by several studies made of abnormal persons and the responses they gave to the items of the Minnesota Multiphasic Personality Inventory (MMPI). Those

---

[1]Paper presented to the Eighth Annual FERA Testing Conferences, Tallahassee, Florida, January 25, 1964.

readers who are familiar with this instrument know that it contains a large number of items, which are "scored" on any or all of sixteen different scales. By isolating patterns of responses given by persons with diagnosed mental disorders, several researchers have been able to identify the same disorders in other patients simply by comparing response patterns. This is an example of one of the techniques of configural scoring. Notice that there is no particular reference to order here, but only the relationship detected by empirical observation of certain response patterns to certain personality characteristics.

A similar concept, which can be illustrated with even more convincing results, was described by Meehl (1950). "Meehl's Paradox", as the situation has become known, is a hypothetical description of responses made to two true-false items by a group of 200 persons. The analysis of the responses is illustrated in Table I. It indicates that in the hypothetical group, which consisted of 100 normals and 100 schizophrenics, 50 from each category responded positively to each item, while the other 50 in each category did not. Considered separately, the responses show no power of discrimination between normals and schizophrenics whatever. But when the pattern of responses to both items is taken into account, the two items discriminate perfectly. The normal subjects responded either true to each item or false to each item, while the schizophrenic subjects responded true to one item and false to the other.

Table 1

An Example of Meehl's Paradox

Normals		Item 1		
		True	False	Total
	True	50	0	50
Item 2	False	0	50	50
	Total	50	50	100

Schizophrenics		Item 1		
		True	False	Total
	True	0	50	50
Item 2	False	50	0	50
	Total	50	50	100

Meehl's Paradox points up the values inherent in configural scoring. While both of the examples mentioned deal with personality traits, the concept of configural scoring can be easily generalized. L. L. McQuitty (1957) has suggested some configural scoring methods which hold promise in educational prediction. Although the mechanics of these methods are too lengthy to describe in this paper, I would like to present a brief sketch of them to you.

McQuitty suggests one method which he calls the "cumulative" method, in which the first step is to find the single item which correlates best with the criterion. This item is then paired with the remaining items until the best-predictive pair is found; this pair is matched with the remaining items until the best-predictive triad is found; and so forth. Eventually a point is reached where the addition of further items will not increase the predictive validity, and so the process is stopped. Another method suggested is called the "reductive" method, in which a single response pattern is taken, and reduced to one or more patterns of less than all the items. The application of this method results in the isolation of those major response patterns on which all the members of a criterion group agree.

In general, configural scoring is practical only when there is a large number of subjects and a relatively small number of items. As the number of items increases, or as the number of choices per item increases, the number of patterns increases exponentially. In Meehl's Paradox, there were two items with two choices each, and 100 subjects. This involved the classification of 100 subjects into four patterns. The number of possible patterns can be generalized to N^t, where N is the number of choices and t is the number of items; if Meehl had chosen to work with 10 items, he would have had to deal with 2^{10} patterns (1024 patterns). Placing 100 subjects into 1024 patterns will obviously leave many empty patterns, upon which no prediction can be based. Since most available predictor tests utilize at least four choices, and contain many more than ten items, the impracticality of the method in a limited situation can be seen. The methods proposed by McQuitty encounter similar difficulty; he cites a report of the use of the cumulative method on a sample of 1474 subjects, which noted that answer patterns with no subject in them began to appear after the fifth item was selected. This may or may not be a real disadvantage, depending upon whether empirical or ideal best-predictor patterns are wanted. Also, the items selected serially by this method might not be the items which would predict best if the items were selected in groups. That is, the first three serially chosen items might not predict as well as the best of all possible triads selected by threes.

A further difficulty presented by configural scoring is the method by which the patterns obtained are represented. The problem is to develop a score which will at once represent the magnitude of the scores as well as their overlap. Lubin and Osborn (1957) have proposed a mathematical method which involves representing the configural scale as a polynomial function of the item scores. This method is a very sophisticated approach, and Lubin and Osborn point out that unless the assumption of linearity cannot be made, the usual total-score correlation with the criterion will be as predictive, if not more so.

These remarks on configural scoring can be concluded by observing that unless the items are built for a particular purpose, or unless there is some reason to believe that certain patterns of responses will be related to the criterion in a certain way, a large scale configural scoring project is not justified, and will probably prove unprofitable.

As previously mentioned, profile analysis takes into account ordered sets of data. It differs from configural scoring in several important ways, although the idea of increasing predictive validity through utilizing pattern similarity is the basis of both techniques. Since profile analysis has some practical applications in school situations, I am going to point my comments toward the description of a method, and stay away from a mathematically oriented theoretical approach. Suffice it to say that the methods do have sound mathematical bases, and the literature contains references to several different methods for several different purposes.

The most important aspect in a practical application of profile analysis is the way in which the relationships between two or more of a student's scores are displayed. The development of a meaningful way to show these relationships requires the use of three characteristics of the score profile: elevation, shape, and scatter.

Elevation is a measure of the amount of each trait present in the profile. That is, the higher a raw score, the higher the elevation of that score. The elevation of all scores in the profile, when taken together, gives the elevation of the profile.

Scatter is the measure of the absolute differences between the scores in the score profile. If all scores are at the same elevation, then there is no scatter; but if the scores differ widely from one another, the scatter is great.

A simply applied method of describing the relationships between scores is called "profile coding". This method uses the position of each of the student's scores within the distribution of scores for a particular test. When profiles are coded, they will fall into several "code groups," upon which prediction of some criterion can be made. These are analogous to the patterns mentioned in the remarks on configural scoring.

The measure of elevation is of greatest consideration in this method. Cutting scores are established in such a way that the distribution of scores for each test in the profile is divided into three nearly equal parts. This can be done most accurately by first transforming raw scores into standard scores, determining the standard scores which correspond to points on the baseline of the distribution which would divide the distribution exactly into thirds, and coding the obtained standard scores according to their position relative to the dividing points. It will be sufficient, however, to use percentile ranks, cutting at the 33rd and the 67th percentiles.

The profile code is determined for each student simply by writing a series of digits which indicates the portion of each distribution into which the student's scores fall. By naming the three portions of the distribution "1", "2", and "3", a three-digit numeric code is developed, which indicates at once the elevation and scatter (and shape for that matter) of each profile.

For example, if John Jones made percentile ranks of 65, 80, and 28 on three tests, his coded profile would be "231". For Mary Smith, whose percentile ranks were 95, 99, and 98, the profile code would be "333". It is obvious that if the profiles are to maintain a uniform meaning, the ordering of the score codes must be the same within every profile.

When profile code groups have been established, the mean criterion score for each code group is obtained. This mean score becomes the predicted score for the profile group.

The method I have described is not the most accurate predictor available. As long as the predictor score distributions are not badly skewed, multiple regression is the most accurate predictive tool at our disposal, with the possible exception of the multivariate techniques for non-linear regression. However, there are some advantages to profile analysis, for which the sacrifice of a little accuracy might be considered:

(1) When percentile scores are used as cutting scores, the coding of profiles becomes a clerical task.

(2) After the mean criterion scores are obtained for the code groups, prediction becomes a clerical task That is, the rather frightening job of plugging numbers into a regression equation with fractional coefficients is eliminated.

(3) Information may be coded into the profile which cannot be utilized in the usual methods of correlation. For example, a size-of-school code, a code representing the result of some ipsative measure, or a grade point average code can be legitimately included in the profile. It is preferable that each predictor have the same number of code variables to keep the measure of elevation constant for each predictor. However, for straight predictive use, this is not absolutely essential.

(4) The shape can be used to good advantage in differential prediction. That is, if enough cases are available, an investigation can be made as to the value of the shape of the profile in predicting success in certain vocational fields or courses of study.

(5) Relationships between variables need not be linear.

The method has the disadvantage attendant to all pattern analytic techniques--that of establishing N^t code groups, some of which will be empty or nearly empty, and upon which prediction cannot be based.

Table 2 presents the results of an example problem.

Table 2

An Example of Pattern Analysis

Coded Scores from 9th Grade Test Data	N	Mean Scores from 12th Grade Test Data	Coded Scores from 9th Grade Test Data	N	Mean Scores from 12th Grade Test Data
111	56	080.5	223	9	302.1
112	13	158.3	231	2	-
113	1	-	232	19	273.9
121	15	122.9	233	16	313.2
122	12	153.7	311	3	-
123	1	-	312	6	294.5
131	1	-	313	7	358.9
132	3	-	321	0	-
133	0	-	322	5	390.4
211	15	166.1	323	34	379.8
212	18	219.8	331	0	-
213	1	-	332	12	363.9
221	7	191.3	333	84	435.8
222	41	232.1			

On the basis of this table, it might be generalized that a student whose profile contained at least two 3's could be expected to attain a twelfth grade sum-of-scores which would gain him admission to one of the state universities; that if the profile contained at least two 2's, he might expect to be in the "twilight zone" between 200 and 300, particularly if the third code was 3; and that if the profile showed two 1's, his score probably would not exceed 200.

Although dividing a distribution of scores into only three parts may seem to be a crude partitioning of available information, further subdivision will greatly decrease the number of cases in each code group, and will increase the number of empty or nearly empty cells. This will decrease the predictive power of the technique even more than a less refined subdivision of information.

Particularly in cases such as the one illustrated, the need for cross-validation is great. The presence of only a few spuriously high or low scores in one of the code groups can adversely affect the predictive value of the method.

References

1. Cronbach, Lee J. and Gleser, Goldine C. "Assessing Similarity between Profiles," Psychological Bulletin, 50 (November, 1953), 456-473.

2. Fricke, Benno G. "A Coded Profile Method for Predicting Achievement," Educational and Psychological Measurement, 17 (Spring, 1957), 98-108.

3. Gaier, Eugene L. and Lee, Marilyn C. "Pattern Analysis: the Configural Approach to Predictive Measurement," Psychological Bulletin, 50 (March, 1953), 140-148.

4. Hartford, Donald L. "An Investigation of a Profile Method for Predicting College Grades from Scores on the College Qualification Tests." Unpublished Ed. D. Dissertation, Graduate School, University of Kentucky, 1959.

5. Horst, Paul. "Pattern Analysis and Configural Scoring," Journal of Clinical Psychology, 11 (January, 1954), 1-11.

6. Kelley, E. L. and Lingoes, James C. "Data Processing in Psychological Research." In Borko, Harold (ed.), Computer Applications in the Behavioral Sciences. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1962.

7. McQuitty, Louis L. "Isolating Predictor Patterns Associated with Major Criterion Patterns," Educational and Psychological Measurement, 17 (Spring, 1957), 3-42.