Rcpts.
35691
v.I21
A-1

AD 227732

(1)

UNIVERSITY OF ILLINOIS

Contract No. AF 33(038)-25726

Memorandum Report A-1
Task A

1 August 1952

DDC
RECEIVED
AUG 11 1971
DCS

TEST RESPONSE PATTERNING

Eugene L. Gaier and Marilyn C. Lee

One of the more promising trends in current psychometric research is the increasing concern with methods of evaluating patterns of test scores and test responses. In the areas of clinical, vocational, social and educational psychology, there is a growing appreciation that test scores do not represent separate and isolated variables but may be more adequately and effectively analyzed with references to their interrelationships. Evidence is being offered that predictive variables will yield higher validities if treated as patterns rather than as mere "and-summations" of separate and independent factors.

The purpose of the present paper is to review recent theoretical and technical developments in pattern analysis with particular emphasis on its most recent application: the study of individual test-item responses. For the present paper, we are making the initial assumption (after Burt, 1), that taking account of the manner in which individual test responses are associated or combined may lead to more precise and valid results than the usual method of merely reporting an average for the separate test items. As long ago as 1937, Zubin (30) pointed out that a total test score may conceal as much as it reveals, since the psychological equivalence of test items can seldom be established nor their additive character proved. A total score may thus carry considerably less diagnostic significance than a detailed analysis of item response patterns.

The present review will be devoted to an evaluation of results based on the thesis that consideration of the manner in which responses are associated by the individual case is an important factor in psychological prediction. Attention will be devoted first to considerations emerging from the analysis of total score patterns. Subsequently, the extension of these principles to

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

the analysis of response patterns within a single test will be examined. Our basic hypothesis is that test data may be so treated as to yield a higher degree of predictive utility than that obtainable by current "additive" methods.

Aspects of Pattern Analysis

The problem of pattern analysis has three aspects which must be differentiated for purposes of conceptual clarity. The first may be designated as "pattern representation," the second as "pattern matching," and the third as "pattern prediction."

The adequate representation of a pattern in numerical terms is in itself a problem which necessitates the prior resolution of several statistical difficulties. The main problem is that of developing statistical indices which can simultaneously depict both the strength and interrelationships of the several variables composing the pattern. Another is that of correcting for the errors of measurement characteristic of all psychometric instruments, so as to accurately estimate the "true" profile. Can we prevent the effects of unreliability from contaminating the statistical representations of the profiles to be studied? These problems must be solved before one can proceed to the second aspect of pattern analysis -- that of pattern matching.

In order to determine the degree of congruence between two patterns it is again necessary to develop statistical coefficients which simultaneously consider both the level and configuration[1] of the variables employed. The subsequent discussion of attempts to solve this problem will serve to illuminate the inadequacies of existing methods as well as the theoretical and methodological considerations involved.

The third facet of pattern analysis involves the prediction of complex criteria by the simultaneous use of several predictive variables. This aspect of pattern analytic methodology shares certain elements in common with both pattern matching and multiple prediction problems. If the criterion to be predicted is itself formulated in terms of a profile, as is the case in most clinical and sociological studies, then the predictive problem becomes that of ascertaining the degree of similarity between the undiagnosed (predictive) profile and the reference (criterial) profile. If, however, the criterion is a single composite index, as is usual in vocational and educational research, then the investigator's problem involves ascertaining that shape

---

[1] Level refers to the strength or intensity of any given trait. Configuration takes into consideration the interrelatedness of the variables.

and level of the pattern of predictors which will yield the best prediction of the criterion involved. "Pattern prediction" may thus include "pattern matching," and both may depend heavily upon the prior development of adequate methods of "pattern representation."

These distinctions form the theoretical background from which follows the subsequent discussion of pattern analytic theory and methodology. Any statistical technique which employs several predictors and attempts to capitalize upon the predictive power contributed by a consideration of their interrelationships, will find a place in the above classification. The multiple regression formula, for example, may be conceived as an illustration of "pattern prediction"; Q-technique and the multiple discriminant function involve, respectively, "pattern matching" and "pattern representation."

Analyzing a Test

In contrast to factor analysis, which attempts to determine the general underlying components in terms of which test relationships may be understood, pattern analysis leads to the division of a group of individuals into types, characterized by specific test score patterns or syndromes. Cattell (3) has pointed out that since a score pattern consists of both traits and relations a duality exists in every pattern which must be considered in its definition. In any test profile, the relations can be considered either apart from or in combination with the particular traits involved. Thus attention may be concentrated only on the configuration or pattern, it may consider both configuration and level simultaneously, or it ma y consider merely the level or intensity of the profile components.

Exclusive consideration of either level or configuration alone ignores the fact that these two aspects of patterns may be interdependent both statistically and psychologically. Although some writers do not consider both intensity and configuration essential to maximal discrimination, Du Mas (8) asserts that intensity differences within the same pattern may have different operational meanings. Consequently, similar patterns located at different levels may indicate interpretations quite at variance with each other as far as behavior is concerned. Although the majority of present psychological tests depend upon level alone for discrimination, the consensus (3, 6, 8, 14, 19, 30) appears to be that: (a) configuration is important in its own right; (b) it is more than a mere summation of the elements involved; and (c) it cannot be reduced to linear combinations of scores.

In his Study of Error, published in 1932, C. C. Brigham (1) summed up the results of six years' work with the Scholastic Aptitude Test in the following statement:

> These studies might be summarized by the general conclusion that a test item, regarded as a specific problem situation to which a certain number of answers may be made, when given to two or more populations sampled for study in the same manner, will show approximately the same distribution of criterion scores for each possible answer in the several groups sampled. (1, p. 69)

Brigham regarded this as having great importance for educational research, since it demonstrated that similar groups of individuals tend to accept and reject certain alternative answers with approximately equal frequencies. For example, he found that responses to the five alternatives of logical inference test items tended characteristically to distribute themselves bimodally. This suggests that it is possible to specify patterns of response or meaning peculiar to and characteristic of particular groups of people. Brigham thus emphasized the potential fruitfulness for predictive purposes of breaking up the typical datum of educational measurements (i.e., total score in a series of tests or subtests) into smaller units consisting of particular test items. By studying individual test items as specific situations, and by analyzing the patterning of their internal and external relationships, he hoped to isolate specific response configurations which could be used to differentiate various classes or types of individuals from each other and from mixed samples.

This principle has received widespread recognition and application in recent vocational and educational prediction (11, 12, 26, 27). In 1935, an attempt to objectify intuitive methods of interpreting aptitude test profiles was made by the Minnesota Employment Stabilization Research Institute (11, 21). Batteries of tests were administered to persons in a variety of occupations in an effort to ascertain the nature of their test score patternings. These batteries were then administered to an individual and the resulting profile was compared with the reference profiles obtained from the various occupational groups to which the battery had originally been given. These studies indicated that individuals successful in a given occupation tend to manifest patterns of ability quite distinct from those in other occupations and in the general population. This rationale has directed subsequent use in vocational and educational guidance of such standard batteries as the Differential Aptitude Tests and the USES General Aptitude Test Battery. A significant aspect of these techniques is that emphasis is placed on the importance of the configuration as well as on amount of individual aptitudes.

In clinical psychology, the pre-quantitative analogues of pattern analysis were the concepts of type and syndrome. Although "types" have

fallen into disrepute because of their tendency to multiply and their lack of specificity regarding component variables, this concept might be revived if sufficiently powerful statistical techniques were available.[2]

The syndrome, no doubt because of its respectable medical parentage, has remained long and consistently enough in vogue to stimulate attempts at statistical refinement. The concept of the syndrome underlies the development of multi-score clinical tests which yield configural results based on several different tests or subtests.

Cronbach (6) has suggested that the results of clinical tests (Rorschach, MMPI, Wechsler-Bellvue) and factorial inventories (Guilford, Thurstone) may be conceived as a point in k-space where k equals the total number of traits measured. The problem of pattern analysis may thus be conceived as an attempt to describe how these multiple measurements are distributed in k-space, and how the distributions compare with one another. From this point of view, one way to study test profiles is to treat these regions of greater concentration in k-space and to determine the significance of the obtained differences in concentration. This conception involves the study of patterns in their own right without reference to an external criterion.

Cronbach further points out that while the statistical comparison of score profiles is logically justified by the k-space concept, the existing methodology of such analysis is quite inadequate. A major problem is that of accurately determining the extent to which two patterns agree when the component scores of either or both may be unequally reliable and unequally correlated. Under such conditions, confidence limits for the significance of obtained differences are difficult to establish and the assumption of linear relationships difficult to justify. For the same reasons, the problems of curvilinear distributions, multivariate regression and multimodal concentrations in k-space have necessarily remained untouched, as well as that of handling of the changes with time which may occur in any of the dimensions measured.

The inadequacies of the methods for dealing with the multiple discrimination problems involved in pattern matching generally inhere in their failure to correct for unfulfilled assumptions of linearity and equality of reliability and/or intercorrelations among items or tests. In addition, some of the methods fail to take account of the direction of the obtained differences or concentrate on configuration to the exclusion of level. However, the contributions of these methods to objectivity of interpretation justify brief mention of each.

---

[2] Stephenson's (23, 24) application of his Q-technique to Jungian typology has been an attempt in this direction.

Methods for Studying Test Patterns

1. Correlation coefficients among the scores, measures, or traits. The use of product moment r's ignores the intensity of the components and for that reason is unable to yield a measure of absolute agreement (congruity) between two patterns. Although its assumption of linearity of relationship may be circumvented by the use of rank-difference correlations, these tend to give undue weight to small and unreliable differences at the shorter end of skewed distributions (5). With both types of correlation it is possible in extreme cases to obtain maximally good prediction and perfect comparison when the two profiles are very dissimilar in both intensity and pattern (8).

2. Q-technique. This method tends to restrict study to those subdivisions of k-space where many cases fall (6), and groups patterns according to gross similarities without adequate reference to level (3). Moreover, it is subject to the limitations attendant upon the use of human judges, and is therefore probably not adequate for a precise comparison of individual profiles.

3. Ratios among the determinants. The reliability of the variates is of particular importance in the use of this method. Minor score changes may drastically affect the size of the obtained ratio, especially when the absolute size of the numerals involved is small (5). Moreover, the technique ignores level (3) and contains the hidden and often unjustified assumption that patterns yielding equal ratios are psychologically equivalent (5).

4. Du Mas' (7, 9, 10) method of pattern comparison. This is a quick but rather gross method that describes the similarity of profiles in terms of the slope of corresponding segments. The method considers both level and configuration, but like the rank difference correlation method it can yield maximally good prediction when the two profiles are actually dissimilar in both shape and level.

5. Zubin's original expression for the agreement of patterns. This is also a relatively unprecise method dealing only with dichotomous scores. This similarity coefficient is based on the number of fundaments for which the scores are the same, divided by the total number of scores in the pattern. This ignores the expected frequency of agreements and tends to ignore intensity (3, 30).

6. Multiple regression equations. This technique considers the intercorrelations of scores and weights them appropriately, thus giving good predictions even when dealing with unreliable figures containing large random elements. However, it proves unstable on cross validation because random

variation is maximized to the experimenter's advantage. While taking into account both level and configuraion, it assumes that strength in one component compensates linearly for weakness in another, and because of this considers the degree but not the direction of deviations involved (5).

The least squares solution provides a good but spurious fit since all or random variation is maximized to the experimenter's advantage. As a result, multiple correlation and the principal-axes factor solution appear to be more exact and elegant than they really are.

Like the factorial specification equation, which may be used if the regression of each component on the criterion is known, multiple regression assumes linearity and takes no account of the possible optimum level of each of the variables (3). Moreover, the composite index it yields is based on a combination of several variables and thus throws away much information by reducing several dimensions to one (5). In extreme cases it may result in an r of one when the profiles are actually very dissimilar (8).

7. Pattern tabulation. Developed by Cronbach (4), this method permits study of the distribution of patterns within a group. The significance of a difference in pattern concentrations may be tested by Chi square. However, the technique can handle only two or three scores at a time, and functions best when the scores utilized are equally reliable and equally correlated. It assumes that profiles having the same shape are psychologically equivalent and thus ignores level by defining configuration in terms of deviation from the average for each score considered (5).

8. Discriminant function. This formula will discriminate two categories of subjects from a mixed sample and thus is of value in studies comparing different types of subjects. Like the multiple regression equation, however, it assumes compensation and reinforcement among the fundaments of the pattern, reduces several dimensions to one composite index, and thus does not give an analytic picture of the types of interprofile deviations which occur (5).

9. Festinger's (13) technique based on comparisons of mean rank. This method considers only the rank order of individuals and thus assumes nothing about normality of distributions or linearity of relationships. It is capable of indicating whether two profiles differ significantly, but is not useful when there are numerous ties in score. Though it has greater stability than the Chi square test for small samples, it is essentially a multiplication of single tests for the significance of differences. Hence, it tends to ignore pattern and the significance of interrelationships among the traits of the pattern (5).

10. Munroe's check list method. Developed in a study of Rorschach protocols (20) this technique consists of a list of "signs" characteristic of some special group which serve to differentiate it from other groups. A common method of developing checklists is to compare the means of two defined groups on one raw score after another and to list all significant differences obtained as "signs." The resulting checklist score is essentially a composite of mean differences in which the configuration is lost. It has the further disadvantage of capitalizing on differences due to accidents of sampling. Like other single score methods, the checklist and signs techniques lose too much information when the component variables are intercorrelated (5).

11. Coefficient of profile similarity. Developed by Cattell (2), this technique considers both shape and level and yields an overall index of the deviations of the two patterns with respect to each variable in the profile. Consequently, it can be used as a measure of absolute agreement between two patterns. However, it may not take sufficient account of configuration in some cases, for it is based on the assumption of uncorrelated and normally distributed components.

It can be seen that the methods of pattern analysis developed suffer either from lack of flexibility or from inability to meet the necessary psychological and statistical assumptions. As Cronbach (6) points out, techniques are needed for evaluating significance when samples are compared on many separate correlated variables, and for estimating the true profiles given unequally fallible, correlated scores. There are at present no practical procedures for studying very complex interrelations of scores or for analyzing the infinite variety of unique patterns which result from the use of multiscore instruments. In order to use statistical methods of analysis, the data must frequently be reduced to fewer dimensions by considering certain patterns to be equivalent in order to obtain a sample sufficiently large to insure adequate reliability. This often entails equating patterns not psychologically equivalent, as, for example, in the psychometric index methods that assume compensation among the pattern components. Nevertheless, the emphasis upon patterning represents a much-needed trend toward the development of objective techniques for representing the organization of ability and personality traits as they occur in the individual.

Analysis of Response Patterns

The emphasis upon the significance of associations between predictors was first extended to item responses within a single test by Zubin (29, 30). He was quick to point out that the traditional practice of assigning empirical weights to responses and summing these to yield a total score assumes unjustifiably that numerically equal scores are psychologically equivalent.

Zubin contended that a knowledge of the patterns of test responses may be more important than knowledge of the total number correctly passed. He hypothesized that response patterns, or clusters of responses to a given group of items having a frequency greater than that expected by chance, will yield results of greater diagnostic significance than a total score which represents merely the summation of discrete responses to qualitatively different items.

To test this hypothesis, Zubin selected a number of personality inventory items, each of which was believed capable of discriminating between normal and abnormal groups. Grouping these items into triads, he tested for the significance of pattern differences by Method 5 above. It was found that scoring by patterns rather than by individual responses is possible, and that normal subjects show significantly more pattern integration than psychiatric patients. The patients exhibited several unique patterns characteristic only of their own diagnostic group. There appeared to be differences among the psychiatric categories in the extent to which they displayed normal and unique patterns of response. Although each item that entered into any of the triads was itself diagnostic, a further significant finding was that not all of the triads of these individually diagnostic items proved to be diagnostic as a pattern.

The other side of this coin has recently been considered by Meehl (19), who points out that "configural" scoring of an item pair may give perfect discrimination between two groups even though each item, taken separately, has zero validity. Thus, dichotomous items having no discriminatory power by themselves might separate two groups completely when scored with respect to pattern. Moreover, the proportion of misclassifications was shown to decrease the more the correlation between the two items differed for each group.

Lubin[3] has appealed to Rao's (24) maximum likelihood scoring technique for a general solution to "Meehl's paradox." According to Lubin, Meehl's finding applies to continuous as well as to dichotomous variables and also to the case where both groups are scored on only one variable (differentiation being achieved by means of individual item scores). Rao's method applies to these and to all other problems which involve assigning subjects to one of several categories and computing the proportion of

---

[3] Lubin, A. Maximum likelihood scoring: a solution to Meehl's paradox. (Manuscript.)

misclassifications to be found. Essentially, it involves reducing the number of misclassifications to a minimum by calculating the likelihood of a particular response pattern for each group and then assigning all subjects with this pattern to the group where the likelihood is a maximum. The method will significantly reduce misclassification only if the frequency distributions are different for each group being studied.

The findings reviewed suggest that configural analysis, as applied to item responses within a single test, may possess greater discriminatory power than the additive techniques which ignore inter-item relationships. The rationale of pattern analysis assumes that the joint presence of several responses may carry a diagnostic significance not inherent in those responses taken by themselves. This concept of emergent phenomena is well summed up by Burt (2):

> "...In certain cases, it is desirable to work with compound characteristics. Thus, not infrequently the mere presence of one or other of two characteristics may have no diagnostic significance by themselves, but their joint presence may be highly significant. ...Similarly, in emotional and moral disturbances, it is the total syndrome that is of diagnostic importance, not the isolated symptom" (p. 171).

Less explicitly, this idea is also inherent in the practice of developing a number of different scoring keys for the same test or inventory on the basis of empirical data showing the relationships of these items to various criteria. To cite but two instances, this method has been used by Strong (25) in validating his interest inventory for a variety of occupations, and by the Human Resources Research Center which has developed navigator, bombadier and pilot scoring keys for the Airman Biographical Inventory. Such empirical keys indicate that analysis of differential patterns of response not only yields higher validity coefficients but also makes it possible to use the same psychometric instrument for a variety of predictive tasks.

Walker (28) has pointed out that no test has a unique answer pattern, for this depends not only on the difficulties of the items but also upon the character of the group of testees. As a result, score scatter is the product of the interaction of the test and population, and as such reflects the characteristics of both.

These principles have been extensively utilized by McQuitty (15, 16, 17) since his attempt in 1937 to differentiate between normal and psychiatric

groups by response patterns. McQuitty's investigations constitute perhaps the most extensive attack made to date on the diagnostic significance of response patterns in personality questionnaires. The most recent development of his technique (18) involves ascertaining the degree of association between every response alternative and every other response alternative for all items in the test being studied. The obtained matrix of association coefficients is factor analyzed using the principle axes solution. Factor measurements for each individual are then obtained from the table of known factor loadings in order to ascertain the degree of integration manifested in the given individual's responses.

Thus, if the subject has chosen many response alternatives characteristic of a large proportion of the original group, i.e., responses having high loadings on the first factor, his pattern may be said to represent majority opinion and his personality said to be integrated according to contemporary cultural standards. If, on the other hand, the subject's responses may be accounted for only by means of a large number of factors, he may be regarded as disintegrated, since his response pattern suggests the presence of attributes characteristic of several disparate groups. This finding is supported by other evidence indicating that group conformity is highly related to personality integration in the sense that an inventory of items weighted to measure this factor discriminates between community persons and mental hospital patients.

The basic hypothesis tested by McQuitty is that personality is integrated to the extent that the individual's ideas are consistent with each other [4]. Psychometrically, personality integration is proportionate to the extent that an individual's responses to a personality inventory are so interassociated that they may be accounted for in terms of a single factor. In assigning primary importance to the interrelationships among the predictors, this method has utilized the analysis of response patterns in order to obtain diagnostic individual factor structures. By evaluating patterns of responses with respect to their similarity-diversity, i.e., the probability that, given one response, another either will or will not tend to occur, McQuitty has been able to differentiate various selected groups with an extremely low percentage of overlap.

---

[4] McQuitty, L. L. Implications of certain measures of personality integration for theories of social psychology. (Manuscript.)

Attribute analysis, or factor analysis of discrete characteristics, has been applied by McQuitty [5] to a variety of test materials in an effort to isolate patterns of responses peculiar to the individual or to types of individuals. In attribute analysis, each response is assumed to reflect a discrete characteristic rather than a continuous variable, Concomitance indices or phi coefficients are used to measure the relationships between responses. Factorization of these indices, computed for all possible pairs of answers, yields patterns of attributes rather than the continuous homogeneous variables presumably isolated by ordinary factor analysis.

The correlation coefficients usually entered into the conventional factorial matrix are derived from test scores where the responses to ability test items are usually classified as either correct or incorrect. One of the most significant aspects of McQuitty's technique is that the relation of each test response, correct or incorrect, is computed with reference to every other answer. This approach therefore provides a new interpretation of wrong responses. It provides for a discovery that the individual does not respond by chance alone when he does not know the correct answer, and that there may be meaningfulness in the individual's pattern of wrong responses.

Summary

It is apparent that there is considerable predictive advantage to be gained by giving adequate statistical recognition to the response patterns which characterize psychometric data. Research to date has demonstrated clearly that the discriminatory power of a prognostic instrument is partially a function of how it is scored, and that an analysis which takes patterns of response into account as well as the individual responses may have more predictive power than the traditional total score approach. Such psychometric techniques are but a special case of one of psychology's general problems, i.e., that of describing the relative frequencies of response in each of the classes he has abstracted from the behavior complex. Since unordered classifications have so far been the rule in most experimental as well as psychometric research, the majority of statistical techniques are not designed to handle the relations existing among the classes. However, the recent research which has indicated that responses are not independent has forced attention to the analysis of patterns and dependent probabilities to the decided advantage of predictive experimentation.

---

[5] McQuitty, L. L. Factor analysis of an individual's discrete responses: An approach to the study of personality integration. (Manuscript

References

1. Brigham, C.C. A study of error. New York: College Entrance Examination Board, 1932.

2. Burt, Cyril. The factorial analysis of qualitative data. Brit. J. Psychol., Statist. Sect., 1950, 3, 166-185.

3. Cattell, R.B. r p and other coefficients of pattern similarity. Psychometrika, 1949, 14, 279-298.

4. Cronbach, L.J. "Pattern tabulation": A statistical method for analysis of limited patterns of scores, with particular reference to the Rorschach test. Educ. psychol. Measmt., 1949, 9, 149-171.

5. Cronbach, L.J. Statistical methods applied to Rorschach scores: A review. Psychol. Bull., 1949, 46, 393-429.

6. Cronbach, L.J. Statistical methods for multiscore tests. J. clin. Psychol., 1950, 6, 21-25.

7. Du Mas, F.M. A quick method of analyzing the similarity of profiles. J. clin. Psychol., 1946, 2, 80-83.

8. Du Mas, F.M. On the interpretation of personality profiles. J. clin. Psychol., 1947, 3, 57-65.

9. Du Mas, F.M. The coefficient of profile similarity. J. clin. Psychol., 1949, 5, 123-131.

10. Du Mas, F.M. A note on the coefficient of profile similarity. J. clin. Psychol., 1950, 3, 300-301.

11. Dvorak, Beatrice J. /Differential occupational ability patterns. Empl. Stab. Res. Inst. Bull., v. 3, No. 8. Minneapolis: Univer. of Minn. Press, 1935.

12. Dvorak, B.J. The new U.S.E.S. general aptitude test battery. J. appl. Psychol., 1947, 31, 373-376.

13. Festinger, L. The significance of difference between means without reference to the frequency distributing function. Psychometrika, 1946, 11, 97-105.

14. Lecznar, William B. Evaluation of a new technique for keying biographical inventories empirically. Research Bulletin 51-2, Human Resources Research Center, Lackland AFB, San Antonio, Texas, March 1951.

15. McQuitty, L.L. An approach to the measurement of individual differences in personality. Char. and Pers., 1938, 7, 81-95.

16. McQuitty, L.L. A measure of individual differences in personality integration. Canad. J. Psychol., 1950, 4, 171-178.

17. McQuitty, L.L. Clinical implications for a measure of mental health. J. abnorm. soc. Psychol., 1951, 46, 73-78.

18. McQuitty, L.L. A statistical method for studying personality integration. In O.H. Mowrer (Ed.), Psychotherapy: A symposium on theory and research. Ronald Press. (In press.)

19. Meehl, P.E. Configural scoring. J. consult. Psychol., 1950, 14, 165-171.

20. Munroe, Ruth L. Prediction of the adjustment and academic performance of college students by a modification of the Rorschach method. Appl. Psychol. Monogr., 1945, No. 7.

21. Paterson, D.G., and Darley, J.G. Men, women and jobs. Minneapolis: Univer. of Minnesota Press, 1936.

22. Rao, C.R. Utilization of multiple measurements in problems of biological classifications. J. roy. statist. Soc., Series B. 1948, 10, 159-93.

23. Stephenson, W. Methodological consideration of Jung's typology. J. ment. Sci., 1939, 85, 185-205.

24. Stephenson, W. A statistical approach to typology: The study of trait universes. J. clin. Psychol., 1950, 6, 26-38.

25. Strong, E.K., Jr. Vocational interests of men and women. Palo Alto: Stanford Univer. Press, 1943.

26. Toops, H., and Kuder, G.F. Test construction and statistical interpretation. Rev. educ. Res., 1935, 5, 229-441.

27. Trabue, M.R. Occupational ability patterns. Personnel J., 1933, 11, 344-351.

28. Walker, D.A. Answer-pattern and score-scatter in tests and examinations. Brit. J. Psychol., 1931, 22, 73-86.

29. Zubin, J. A technique for pattern analysis. Psychol. Bull., 1936, 33, 773. (Abstract.)

30. Zubin, J. The determination of response patterns in personality adjustment inventories. J. educ. Psychol., 1937, 28, 401-413.