A REVIEW OF THE APPLICABILITY OF DISCRIMINANT ANALYSIS TO SOCIAL SCIENCE RESEARCH

Beth Dawson, Southern Illinois University School of Medicine

Discriminant analysis is a statistical technique that allows the researcher to examine differences among groups. As such, it provides a unified approach to research situations where there are several groups and it is desirable to: 1) establish significant group differences; 2) study and describe the variables on which groups differ; 3) classify future individuals into the most appropriate group (Tatsouka and Tiedeman, 1954).

The primary purpose of the present paper is to describe the relevance of discriminant analysis in social science research. This will be attempted by reviewing variable selection methods, appropriate statistical tests of significance, and interpretation of discriminant functions. Finally, a brief review of some of the studies utilizing discriminant analysis in education is presented with a discussion of the relevance of the procedure for specific applications.

Both conceptually and resultwise, Tatsouka (1971) points out that discriminant analysis in the two group case is equivalent to regression analysis using a "dummy" criterion variable. More generally, with k groups and p predictor variables, results are equivalent to those obtained using canonical correlation when the set of criterion variables contains k-1 variables coded dichotomously such that:

$Y_{ij}=1$ if observation j is on a member of group i
$=0$ otherwise; $i=1,2\ldots,k-1$; $j=1,2,\ldots,n_k$.

While k-1 (or p, if smaller) discriminant functions are generated, it is almost always the case that two or three functions are sufficient to account for the group differences. Thus, the procedure results in a reduction of the space necessary to describe group differences. In addition to examining main effects, Tatsouka (1969) suggests that the method is equally effective to study the nature of interaction effects following multivariate analysis of variance. In this case, $\underset{\sim}{S}_{int}$, the sums of squares and cross products matrix (SSCP) of the appropriate interaction effect is substituted for the between groups matrix in the characteristic equation, and the eigenvalue-eigenvector pairs are found for $\underset{\sim}{W}^{-1}\underset{\sim}{S}_{int}$, where $\underset{\sim}{W}$ is the within groups SSCP matrix.

One of the practical problems confronting the researcher who uses discriminant analysis is similar to that of multiple regression: selection of variables that significantly differentiate between the groups. The problem is confounded in the case of discriminant analysis due to the additional consideration of number of groups. Several authors have addressed the variable selection problem, and, while there is no consensus on the "best" method, the following are some of the approaches that have been suggested.

In the two group case, Collier (1963) developed a test of whether the addition of p-q predictors add significantly to the discrimination obtained using q predictors alone. Assuming bivariate normality and equal variance-covariance matrices, Collier shows that

$$\frac{(n_1+n_2-p-1)}{p-q} \cdot \frac{R^2(p)\ R^2(q)}{1-R^2(p)} \sim F(p-q, n_1+n_2-p-1)$$

under the assumptions of regression where n_1 and n_2 are sample sizes of groups 1 and 2. The generalized distance function can be estimated by

$$D^2(p) = \Sigma_i^p \Sigma_j^p S^{ij}(\bar{X}_{1i}-\bar{X}_{2i})(\bar{X}_{1j}-\bar{X}_{2j})$$

where S^{ij} is the ijth element of $\underset{\sim}{S}^{-1}$, the sample estimate of $\underset{\sim}{\Sigma}^{-1}$. Using the functional relationship between R^2 and D^2, Collier suggests the following statistic for testing whether the generalized distance is the same when p variables are used as when only q variables are used ($q<p$):

$$\frac{(n_1+n_2-p-1)}{p-q} \cdot \frac{n_1n_2(D^2(p)-D^2(q))}{(n_1+n_2)(n_1+n_2-2) + n_1n_2D^2(q)}$$

distributed $F(p-q, n_1+n_2-p-1)$.

Related to tests for multivariate analysis of variance, Roy and Bargmann (1958) developed an a posteriori procedure for testing specific contributions by predictor variables. The step-down method is applied to an a priori ordering of the predictor variables: the most important variables are ordered first and the least important are ordered last. Testing occurs in reverse order and stops as soon as a significant variable is encountered.

Bargmann (1970) recommends a hierarachical method to select variables that may be used when more than one discriminant function is retained. He suggests the correlation between the first discriminant function and the predictor variables be used as an index of discrimination. Those variables with high correlations would than be deleted and the analysis redone on the remaining variables. This process is continued until most of the variables have been assigned to one of the functions in the hierarchy or until correlations become small enough to ignore.

Kraatz (1975) compares three different methods of variable selection. The first method is based upon transforming predictor variables into predictive components using the validities of these variables with a single criterion (Tucker, 1973). Kraatz extends the procedure to the special case where group membership is the criterion variable and compares this method of variable selection with a stepwise procedure and with the use of all predictor variables in a Monte Carlo study. The criterion used to judge the effectiveness of each method was the value of Wilks' lambda computed for the population. When sample values of lambda are computed for each of the three methods, the method of using all variables consistently overestimates the amount of variance accounted for between groups in the population. When population values are used for the within and total covariance matrices, all methods underestimate the amount of variance accounted for, although the method based on predictive composites of validities is the best estimator of lambda for middle ranges values of lambda. The stepwise method is never a best estimate in this situation. Finally, when a cross validation sample is considered, the predictive composite method again performs most satisfactorily.

Huberty (1971) presents a review of variable selection methods and then compares six methods. Two sets of actual data with three and five groups respectively were used for analysis and the index of discriminatory power was the proportion of correct classifications. The six selection methods were based upon beta weights, F-ratios resulting from a priori univariate analysis of variance, the stepwise procedure used in the BMD computer program, component loadings following a rotation of a principal component analysis, factor analytic-discriminatory correlations (selecting those variables that correlate highly with the first discriminant function for each of the clusters of variables defined by a factor analysis), and correlations between variables and the leading discriminant function. In general, Huberty found that the stepwise procedure results in a higher proportion of correct classifications the majority of the time.

The entire subject of variable selection needs more study. It would appear that Monte Carlo studies varying the number of variables, the number of groups and the intercorrelations between variables are in order. In addition, the effect of violation of the assumptions of multivariate normality and common variance-covariance matrices on various methods needs investigation. At the present time, there appears to be no clear-cut answer as to which variable selection method is best; it is quite possible that no one method will prove robust under all of the conditions mentioned above.

Several tests have been suggested as appropriate in determining the significance of the discrimination between groups. These tests follow naturally from the tests suggested in multivariate analysis of variance. The most commonly used statistics as given by Tatsouka (1971) are the likelihood ratio statistic, Wilks' lambda (Λ); Hotelling s trace criterion (τ); and Roy's largest root criterion (Θ).

In the case of discriminant analysis, the formulae are, letting

$r = \min(k-1,p)$
S_e = SSCP matrix for error effects
S_h = SSCP matrix for hypothesis effects
λ_i = the ith eigenvalue

$$\Lambda h = \frac{|S_e|}{|S_h + S_e|} = |S_e^{-1} S_h + I|^{-1} = \Pi_i^r (1+\lambda_i)^{-1}$$

$$\tau = \Sigma_i^r \lambda_i = \Sigma_i^r (S_e^{-1} S_h)_{ii} = tr(S_e^{-1} S_h)$$

$$\Theta = \lambda_1/(1+\lambda_1)$$

Pillai derived the distribution for Λ and tabled the null distribution for selected centile points. Heck and Pillai also developed tables for selected centile points of Θ following Roy's derivation of the null distribtuion of Θ (Bock and Haggard, 1968). The sampling distribution of Λ was obtained by Schatzoff but exact numerical computations are feasible only when p is even and k is odd. Consequently, two approximations to Λ commonly used are Bartlett's V, a function of the logarithm of Λ with an approximate chi-square distribution; and Rao's R, a radical function of Λ with an approximate F distribution. While R is a closer approximate test and is exact for p=1 or k=2, V an be expressed as the sum of several terms (Tatsuoka, 1971).

Schatzoff (1966) investigated these three statistics along with three others to determine their respective sensitivity to a wide variety of parameterizations. As the criterion, Schatzoff used the concept of expected significance level (ESL) that may be defined as one minus the power of a test, averaged over uniform values for α. Therefore, when comparing two test statistics, the one with the smaller ESL is preferred. Results from a Monte Carlo analysis indicate that while on one statistic is best in all situations, the largest root Θ is much worse than the other two except for the two group case. He concludes that a researcher would do equally well to use either Wilks' Λ or Hotelling's τ since both are fairly robust with respect to a wide range of alternate hypotheses.

Another approach to the question of significant discrimination is the use of a statistic such as Hays' $\hat{\omega}^2$, an estimate of the amount of variance that can be attributed to the relevant independent variable(s) in the population. Tatsuoka (1970) defined a multivariate analogue of $\hat{\omega}^2$ for use in one-way multivariate analysis of variance, denoted

$$\hat{\omega}^2_{mult} = \frac{|S_h| - |S_e| - ((k-1)/(N-k))|S_e|}{|S_h| + (1/(n-k))|S_e|}$$

where N is the total sample size. He performed a Monte Carlo simulation (1973a) to examine the statistical properties of $\hat{\omega}^2$mult. The statistic was found to be quite positvely biased unless the ratio of subjects to variables is extremely large. Therefore, an empirical approach to correct the bias was conducted with the result

$$\hat{\omega}^2_{corr} = \hat{\omega}^2_{mult} - \frac{p^2+(k-1)^2}{3N}(1-\hat{\omega}^2_{mult})$$

Upon testing, the corrected index was judged adequate for cases where $p(k-1) \leq 49$ and $75 \leq N \leq 2000$. To test the significance of individual discriminant functions, Tatsuoka (1970, 1971) utilized the fact that Bartlett's V function of Λ is approximately distributed chi-square. Since the discriminant functions are uncorrelated and the terms in V are statistically independent, a sequence of tests may be defined to assess the contribution of each subsequent root. The testing procedure continues until for some j the test is not significant, leading to the conclusion that only the first j+1 discriminant functions are significant. Although subsequently questioned by Harris (1974), a recent open letter to Harris from Tatsuoka (1976) has re-established the validity of this test.

Two methods have been suggested for interpreting the relative contribution to discrimination by a given predictor variable. The first method (Tatsuoka, 1970, 1971) consists of standardizing the coefficients of the predictor variables in the discriminant function. The other method (Cooley & Lohnes, 1971) is based upon the factor structure defined by the correlations between the original variables and each of the discriminant functions. This approach is more meaningful when it is desired to interpret the discriminant functions as opposed to measuring the relative contribution of each variable to the discriminant functions (Tatsuoka, 1973b). Bargmann (1970) recommends the use of correlations between variables and only the first discriminant function.

Huberty (1975) reports a Monte Carlo study designed to study the stability of the standardized coefficients and variable-discriminant function correlations under repeated sampling. Neither measure cross validated with great stability, and the study was limited to the first discriminant function.

Following is a brief review of several studies using discriminant analysis in education. Emphasis is placed on techniques used and thoroughness of reporting rather than results of the analyses. The articles are not intended to be either representative or comprehensive; the reader is referred to the article by Tatsuoka and Tiedeman (1954) for references to uses of discriminant analysis up to that date and to the publications by Huberty (1975) and Lachenbruch (1975) for more complete listings of recent applications.

Discriminant analysis has been used in several studies related to evaluation of student success in higher education. Selover (1942), in one of the earliest applications, investigated the utility of a sophomore testing program in determining success in major curriculum groups. Data were obtained from four years of testing, and 21 curriculum groupings were analyzed. Using four subtests, Selover computed a linear discriminant function to differentiate between various sets of two groups each.

More recently, Keenen and Holmes (1970) studied college withdrawal and failure. Groups were defined on students who graduated, withdrew, or failed at a liberal arts college. Four intellectual variables (not described by the authors) and thirty non-intellectual measures based on demographic variables and stated interests were considered. Two discriminant functions were extracted on each of three analyses: the first included both types of variables, the second only intellectual variables, and the third only non-intellectual measures. Cross validation of classification using non-intellectual measures was performed on a ten percent holdout sample.

Study techniques were examined to discriminate between students in four undergraduate disciplines who had above and below average grades (Goldman & Warren, 1973). A four by two multivariate analysis of variance was performed; interaction between discipline and grade average was not significant. Subsequent discriminant analysis of the discipline effect resulted in two significant functions; no classification or cross validation analyses were reported.

Discriminant analysis has been applied in educational evaluation to study the effect of differences in learning environments. Anderson <u>et al</u>. (1969) used student perceptions of learning climates to evaluate the impact of a new physics curriculum. Three groups were defined on the basis of the experimental curriculum, the traditional curriculum, and the experience of the teachers with the new curriculum. Results of a questionnaire administered to students were the predictor variables. Discriminant analysis was followed by rotation of the two significant functions to increase interpretability. This study is uncommon with respect to the clarity and detail of reporting the methods of analysis.

A fairly common application of discriminant analysis has been the prediction of vocational choice. Cohen (1971) reports results of a study to differentiate between male college seniors in business administration and teacher education. The single discriminant function was obtained using factors from an instrument measuring motivation for career choice. Results of classification on a replication sample were reported. Porebski (1966) used three tests to discriminate between four technical trade groups. Two discriminant functions were retained; no classification was performed. Using only three predictor variables enabled the author to present considerable computational detail.

In the field of medical education, Checker <u>et al</u>. (1972) attempted to predict specialty choice for four major specialty fields from the four Medical College Admission Test (MCAT) scores obtained from students prior to admission to medical school. Two functions were retained for analysis, and classification as correct in 36 percent of the cases. No replication studies were reported.

Paiva <u>et al</u>. (1974) examined career choice transition during medical school. Three major career categories were considered. Students were surveyed as to their career preference upon entering medical school and again at the end of the senior year; the consistencies and shifts in career choice were used to define several groups. Two discriminant functions accounting for 77 percent of the differentiation were extracted and retained for interpretation. No classification analysis was performed.

Applied research in the social sciences is often hampered due to small sample sizes available for analysis. A study not so impacted resulted in a very comprehensive and detailed report on the use of discriminant analysis on Project TALENT five-year follow-up data by Hall (1967). In one study, Hall used 30 predictor variables to differentiate between 17 private colleges subsequently attended by students for whom 1960 test scores were available. In another study, Hall utilized twelve predictor variables related to student perceptions of self, home, and future and how they changed between grades nine and twelve. These papers are presented in great detail and explanation of variables used and procedures performed.

The final study to be summarized is one performed by the present author (1975) which indicates that insights may be gained into the generalizability of performance from one class to another during the developmental years of a new institution. In an exploratory study, discriminant analysis was used to identify the pre-admission variables that predict success in the first year of medical school. For the purpose of this analysis, success was defined in terms of a composite criterion score derived from performance on various examinations during the first year. Students were divided into three groups on the basis of their composite score.

Although multiple regression analysis is also an appropriate statistical procedure to predict success, the medical school curriculum is competency based and there is little desire to use a procedure that results in a precise ranking of predicted success. It is more relevant to identify those students who may experience some difficulties in the first year and for whom remedial instruction should be anticipated. Likewise, it is helpful to identify students who are expected to perform consistently above criterion level so that educational enrichment activities may be developed.

Seventeen predictor variables were examined and included the four MCAT scores, various GPA measures, the number of course hours in given areas, and data on specific courses. Data on the classes admitted in 1973 and 1974 were analyzed; sample sizes were 48 and 63. For each class, discriminant functions were found and then used to classify both the analysis class and the other class for cross validation.

A stepwise method was used to select the significant variables in the analysis. Entry of variables was controlled by using an F value of 2.00, approximating an $\alpha \leq .05$ level. The selection criterion used in this analysis was the method in which the variable that minimizes Wilks' lambda is entered at each step. Prior probabilities equal to the sample percentages were assigned for use in subsequent classification analysis.

Analysis on the first class resulted in six significant variables loading on the discriminant functions. The eigenvalues associated with the two discriminant functions account for 71.7% and 28.3% of the total variance existing in the discriminating variables (See Table I). The canonical correlations are .63 and .46 for the first and second functions respectively. Application of the sequential chi-square test of individual functions results in $p \leq .001$ for one root, $p \leq .08$ for two roots.

Table II shows a plot of the group centroids and indicates that function one primarily separates the High and Low groups, and function two differentiates between Medium and Low groups. Due to the evidence of group separation and the exploratory nature of the analysis, it was decided to retain both discriminant functions.

Interpretation of the standardized weights in Table III indicates that the most important dimension separating the groups is an ability in quantitative and science areas along with a history of specific science courses. The groups are ordered from low to high on this dimension consistent with their performance on the composite criterion. Another important dimension is verbal ability, with relevant contributions from a course in biochemistry, that separates the Medium group from the Low group. One might tentatively suggest that students well based in science with attendant ability as measured by the Quantitative MCAT perform best in the first year of medical school; students with less science and quantitative ability may still perform satisfactorily if they are well qualified on verbally related skills and have specifically taken a course in biochemistry. These results are not particularly surprising, but they tend to reinforce our prior expectations.

TABLE I

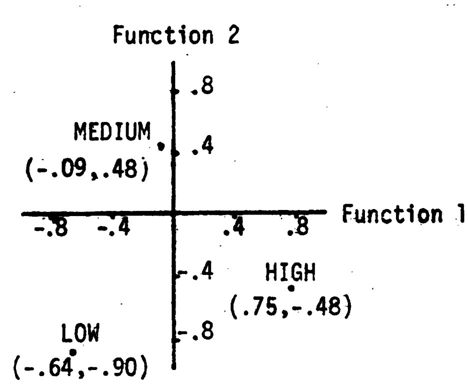
DETERMINATION OF NUMBER OF DISCRIMINANT FUNCTIONS TO BE RETAINED

Discrm Function	Eigenvalue	Relative Percentage	Canonical Correlation
1	0.670	71.69	0.633
2	0.265	28.31	0.457

Function Derived	Wilks' Lambda	Chi-Square	df	Significance
0	0.474	31.78	12	.001
1	0.791	9.98	5	.076

TABLE II

PLOT OF GROUP CENTROIDS



The final step in this analysis was to examine the accuracy of classification of students into correct groups. Comparison of actual group membership with predicted membership is shown in Table IV; 72.9% of the students are correctly classified using two functions. Computing these functions on the second class for cross validation results in correct classification of 56.6%.

Although complete details are not presented here, it is of interest to note some of the results obtained when the discriminant analysis was performed on the second class. Again, six variables loaded on the discriminant functions and both functions were retained. The functions correctly classified 84.9% of the class on whom they were computed and 54.2% of the cross validation class.

Only two variables entering the discriminant functions were the same both both analyses, an indication that the two classes may differ in abilities that tend to discriminate performance. Alternately, the composite score used as a criterion may not be an equivalent measure for the two classes. In both analyses, a considerable reduction in the percent of correct classifications occurs when the functions are applied to the cross validation sample. Although some shrinkage in correct classification is not unusual, differences in the classes or inequivalence of the criterion may be contributing to this shrinkage. Thus, the present analysis has identified two areas for further research and indicates that generalization with respect to performance of students from one class to another should be extremely cautious in the formative years of a curriculum.

TABLE III

STANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	Function 1	Function 2
Chemistry GPA	.322	-.277
Biochemistry Taken	.217	.651
Cell Biology Taken	.309	.101
Verbal MCAT	-.236	.626
Quantitative MCAT	.593	-.045
All Other GPA	-.072	1.016

GROUP CENTROIDS

	Function 1	Function 2
High Group	.75	-.48
Medium Group	-.09	.48
Low Group	-.64	-.90

TABLE IV

RESULTS OF CLASSIFICATION

Class on Whom Analysis Performed

Actual Group	No. of Cases	Predicted Group High	Med	Low
High	11	6	4	1
Med	28	2	25	1
Low	9	0	5	4

Percent Correctly Classified: 72.9%

Second (Validation) Class

Actual Group	No. of Cases	Predicted Group High	Med	Low
High	15	3	12	0
Med	28	2	22	4
Low	10	0	5	5

Percent Correctly Classified: 56.6%

Summary

The foregoing discussion has attempted to identify some of the methodological areas of research in discriminant analysis and examine the applicability of this method in a few selected studies. Social science research almost invariable attends to situations that are complex to study. Multiple independent and dependent measures are frequently assessed. Researchers in these fields have recently begun to utilize the greater analytic power made possible by multivariate methods. Discriminant analysis appears to be especially relevant to many research questions that involve a polychotomous, unranked criterion. The concurrent sophistication of statistical computer programs has rendered multivariate techniques accessible and easy to apply. Perhaps they are too easy; researchers should be aware of the assumptions of the models they utilize and remain cognizant of the measurement problems still existing in the study of human behaviors.

In reading reports of research that have used discriminant analysis, two particular aspects were noted that, if properly attended, could result in better research and more relevant utilization of research findings. First is the problem in discriminant analysis of the validity of the initial classification of subjects into groups. The model assumes that classification partitions the sample: each individual is a member of one and only one group. The validity and generalizability of all subsequent analyses are based upon the accuracy of the initial classification. This calls for more precise definition of groups being studied. In many situations, definition of group membership is simple and presents no problem; i.e., classification on the basis of sex, age, other biographic characteristics. Studies that examine career choice or college major must recognize that the classification is valid only for that point in time and does not necessarily extend to future preferences. The problem is even more apparent in research that utilizes diagnostic categories to describe groups, such as medical research and personality studies. The accuracy of the initial classification must be examined and findings interpreted accordingly.

Second, although it may seem elementary, greater attention should be given to complete and precise description of the research. The very nature of the groups analyzed, the basis for determining initial classifications, and the types of predictor variables used to derive the discriminant functions are frequently referred to in a very cursory manner. In addition, the lack of description of the exact procedures used is sometimes appalling. Method(s) of variable selection, criteria used to determine the number of functions retained for analysis, and the type of classification functions used should be explicitly noted in all reports of research. The space constraints of many professional journals make complete description difficult. But a brief paragraph outlining methods and criteria should be insisted upon by the author; such description need not be eloquent and can generally be condensed into several concise sentences. Certainly the primary purpose of publishing research results is to communicate methods and findings to other researchers. Many published studies do not achieve this objective.

References

Anderson, G.J et al. Curriculum effects on the social climate of learning; A new representation of discriminant functions. Amer. Educ. Research J., 1969, 6:315-328.

Bargmann, R.E. Interpretation and use of a generalized discriminant function. In R.E. Bose, et al. (Eds.), Essays in Probability and Statistics. Chapel Hill: Univ. of N.C. Press, 1970.

Bock, R.D. et al. The use of multivariate analysis of variance in behavior research. In D.K. Whitla (Ed.) Handbook of Measurement and Assessment in Behavioral Sciences. Reading, MS: Addison-Wesley, 1968.

Checker, A. et al. Can the choice of medical specialty be predicted by the Medical College Admission Test? Proceedings, 80th Annual Convention, APA, 1972.

Cohen, D. Differentiating motivations underlying vocational choice. J. of Educ. Research, 1971, 64:229-234.

Collier, R.O. A note on the multiple regression techniques for deleting variables in the discriminant function. J. of Exper. Educ., 1963, 31:351-353.

Dawson, B.K. Discriminant analysis. Paper presented at AAMC/RIME, Washington, D.C., 1975.

Goldman, R.D. et al. Discriminant analysis of study strategies connected with college grade success in different major fields. J. of Educ. Meas., 1973, 10:39-47.

Hall, C.E. Three papers in multivariate analysis. Project TALENT Five-Year Follow-up studies: Interim Report 2, AIR and Univ. of Pittsburg, 1967.

Harris, R.J. Multivariate analysis of variance. Paper presented at AERA, Chicago, 1974.

Huberty, C.J. On the variable selection problem in multiple group discriminant analysis. Paper presented at AERA, New York, 1971.

Huberty, C.J. Discriminant analysis. Review of Educ. Research, 1975, 45:543-598.

Keenen, C.B. et al. Predicting graduation withdrawal and failure in college by multiple discriminant analysis. J. of Educ. Meas., 1970, 7:91-95.

Kraatz, J. The use of predictive composites in multiple discriminant analysis. Unpubl. doct. dissertation, Urbana: Univ. of Illinois, 1975.

Lachenbruch, P.A. Discriminant Analysis. New York: Hafner, 1975.

Paiva, R.E.A. et al. Career choice transition during medical school. Paper presented at AAMC Colloq. on Career Dev't of Physicians, Washington, D.C., 1974.

Porebski, O.R. Discriminatory and canonical analysis of technical college data. Brit. J. of Math and Stat. Psych., 1966, 19:215-236.

Roy, S.N. et al. Tests of multiple independence and the associated confidence bounds. Annals of Math. Stat., 1958, 29:491-503.

Schatzoff, M. Sensitivity comparisons among tests of the general linear hypothesis. Amer. Stat. Assn. J., 1966, 415-435.

Selover, R.B. A study of the sophomore testing program at the University of Minnesota. J. of Applied Psych., 1942, 26:296-307, 456-467, 587-593.

Tatsuoka, M.M. Multivariate Analysis. Review of Educ. Research, 1969, 39:739-744.

Tatsuoka, M.M. Discriminant Analysis. Champaign, IL: Institute for Personality and Ability Testing, 1970.

Tatsuoka, M.M. Multivariate Analysis. New York: Wiley, 1971.

Tatsuoka, M.M. An examination of the statistical properties of a multivariate measure of strength of relationship. Final Report, Proj. No. 2-E-020, HEW, 1973a.

Tatsuoka, M.M. Multivariate analysis in educational research. In F.N. Kerlinger (Ed.), Review of Research in Education. Itaska, IL: Peacock, 1973b.

Tatsuoka, M.M. Letter of response to Richard Harris. Personal communication, 1976.

Tatsuoka, M.M. et al. Discriminant analysis. Review of Educ. Research, 1954, 24:402-420.

Tucker, L.R. Notes on a system to determine predictive composites based on validities. Unpubl. paper (mimeo), Urbana: Univ. of Illinois, 1973.