Some New Ways of Analyzing MANOVA Data in a Post Hoc Context

Tungshan Chou

ABSTRACT

This paper presents the multivariate statistical procedure, MANOVA, in a non-traditional way unfamiliar to most educational researchers. Whereas traditional way of using MANOVA has been viewed only as a preliminary test for safeguarding the experimentwise type I error rate for conducting the individual univariate tests, the non-traditional way as presented in this paper seeks to explain the grouping variable effect in a truely "multivariate" framework. A challenge was made to the logic of traditional use of MANOVA with respect to both experimentwise type I error rate and statistical power.

It was argued that all traditional MANOVA follow-up procedures fail to interpret the MANOVA effect in a multivariate context. For remedying the traditional procedures, some extensions were proposed through the use of discriminant analysis statistical procedure. These extensions were made in three perspectives: searching for underlying dimensions, determining relative variable importance, and conducting multivariate group contrasts.

Finally, the techniques discussed in the paper were used to illustrate a real life research scenario. The example study attempts to investigate the differences of opinion as expressed by community residents, teachers, and school principals in a recent mail survey in Hualien area. Our analyses indicated that two underlying attitudes were present which could be accountable for the differences of opinion among these three categories of people: an apathetic attitude toward school-community relationship and the attitude toward schools providing extra-curriculum services for the community residents. None of our findings could have been reached with the traditional MANOVA follow-up procedures.

The statistical procedure of multivariate analysis of variance, commonly known as MANOVA, has been widely adopted in the works of applied researchers in education and psychology as well as social sciences in general. This procedure has largely been used in research situations where the effect of the independent variable(s) (also called grouping variable(s)) is to be investigated in the presence of multiple dependent variables (also called outcome variables), as a natural generalization of its univariate counterpart, ANOVA. In fact, almost all of the currently existing statistical books which include MANOVA as one of its topics recommend the practice of using MANOVA as a preliminary test prior to conducting the multiple univariate ANOVAs. However, this practice has recently been challenged by some researchers (Bray & Maxwell, 1982; Huberty & Morris, 1989). The challenge was primarily addressed to the concerns that information obtained from multiple ANOVAs ignore the correlations among the multiple outcome variables,

國立台中師範學院教育測驗評量與統計方法研究發展中心
測驗統計年刊 第三輯 民84年12月 237-265頁

and consequently yields only fragmentary information which could not be meaningfully fit together for researchers to make sense out of their data. In this paper, the term "univariate analysis" refers to the analysis when a single outcome variable is involved, whereas "multivariate analysis" refers to the analysis that involves multiple outcome variables. Multiple univariate analyses refer to a group of univariate analyses which this current paper argues, should not be confused with a multivariate analysis.

One well-developed technique that could be used to follow up on significant MANOVA results is discriminant analysis. Tatsuoka (1969) had described the application of discriminant analysis as a follow-up procedure to MANOVA as "probably the most significant development during the 1960s" (p. 742). While the mathematical basis of this technique was largely developed in the thirties and forties, only in relatively recent years, and due to the availability of the computing facilities made to the general research community, we have seen calls for their applications in the educational research (Bray and Maxwell, 1982). However, despite the easy access to the softwares capable of performing discriminant analysis, a recent survey of six major research journals published by either the American Psychological Association or American Educational Research Association revealed the fact that out of a total of 222 MANOVA research situations, only two used discriminant analysis to follow up on their significant MANOVA results (Huberty and Morris, 1989). Not surprisingly, for the educational research community at large, it is almost an automatic response to follow up on a significant MANOVA result with multiple univariate ANOVAs.

Purpose

The purpose of this paper is to present MANOVA statistical procedure in a truely multivariate-conceived framework as opposed to the common conception of MANOVA simply as a summary of collective univariate results. Then, the technique of discriminant analysis in its versatile potential, as a MANOVA-compatible multivariate procedure will be illustrated as a natural follow-up procedure to the solutions of various forms of research interest associated with MANOVA.

The current paper is organized into three major parts. In the first part, an overview of the MANOVA procedure is presented, followed with a challenge to the current traditional practice in the interpretations of the MANOVA results. In the second part, discriminant analysis is introduced as a follow-up procedure, with extensions made to some other aspects of the MANOVA research questions. It is argued that the choice of a follow-up procedure to MANOVA should be guided by the specific research question the researcher wishes to answer, which in turn should be determined by the motivation of the researcher. Three conceivable motivations for following up on a significant MANOVA result are covered in this study: searching for underlyingdimensions, determining relative variable importance, and conducting multivariate group contrast. In the last part, a real data set will be used to illustrate the techniques discussed in this paper.

An Overview of MANOVA

Multivariate analysis of variance is typically used to examine the effect of one or more grouping variables on multiple outcome variables. We shall use the general notations of k to represent the number of levels (or groups) in the grouping variable with n_j observations in the jth group and p to represent the number of outcome variables. Let's begin with the univariate scenario of k groups measured on one outcome variable (p=1). The null hypothesis of interest is the equality of k population means. The grouping effect can be assessed under the usual linear model parameterized as

$$X_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

where X_{ij} denotes the ith individual from group j, μ is an overall effect common to all groups (or populations, to be precise), α_j is the specific effect associated with the population from which group j was sampled, and ε_{ij} is the random residual unaccounted for by the model. This setting has been referred to as the univariate one-way classification. For the case of extending the number of outcome variables from 1 to p, MANOVA is simply a generalization of the above linear layout by replacing each term with its matrix form,

$$\mathbf{X}_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

where $\mathbf{X}_{ij}$ represent independent p dimensional vectors of observations sampled from k multivariate normal populations with mean vector $\mu_j = \mu + \alpha_j$, j=1 to k and common variance-covariance matrix Σ. To test the null hypothesis of equal mean vectors, the likelihood ratio criterion is commonly known as the Wilk's lambda, computed as

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{T}|} = \frac{|\mathbf{E}|}{|\mathbf{B} + \mathbf{E}|}$$

where **E**, **B**, and **T** are the within-group, between-group, and total sums-of-squares matrices, respectively. Under H_o of no treatment effect, then $\mathbf{B} = \mathbf{0}$ and Λ will take on the value of 1, whereas if treatment effect is large relative to **E**, Λ will be small. Since Λ is defined in terms of the determinants of **E** and **T** and the fact that the determinant of a covariance matrix is the generalized variance for the p variables involved, the complement of the Wilk's lambda, $1-\Lambda$, can thus be used as a generalized R^2 (or η^2, to be technically correct) indicator which indicates the amount of variance in the p-variable system that is attributable to the effect of the grouping variable. Unfortunately, this information has often been ignored by the MANOVA users.

The sampling distribution of Λ is generally very complicated. Only under a few special cases can Λ be converted to exact F distributions through transformations. These special cases are summarized in Table1. Besides these cases, an approximation to more familiar distributions is generally adopted. Two approximations are currently available: (1) Bartlett's χ^2 and (2) Rao's F. Bartlett's χ^2 is given by

$$\chi^2 = -[(N - 1) - .5 (p + k)]\ln \Lambda$$

with p(k-1) degrees of freedom. Rao's F approximation is quite complicated to compute, and is available in the MANOVA procedure of most major statistical computing software packages such as SAS and SPSS. Interestedreaders may refer to their manuals for the detailed computational steps. Some empirical research had indicated that Bartlett's χ^2 is a good approximation for moderate to large sample size and Rao's F is a better approximation for smaller sample size (Lohnes, 1961), although both statistics generally lead to the same decision on H_o. It should be noted that the degrees of freedom for error with Rao's F can be non-integer, researchers should not be alarmed at such a finding.

Table 1. Special cases in which transformations of Λ are exactly distributed as F

p	k	Transformation	df's of Exact F
Any number	2	$(\frac{1-\Lambda}{\Lambda})(\frac{n-p-1}{p})$	p, n - p - 1
Any number	3	$(\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}})(\frac{n-p-2}{p})$	2p, 2(n - p -2)
1	Any number	$(\frac{1-\Lambda}{\Lambda})(\frac{n-k}{k-1})$	k - 1, n - k
2	Any number	$(\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}})(\frac{n-k-1}{k-1})$	2(k-1), 2(n-k-1)

Source: Dillon and Goldstein (1984), p. 422

In addition to Wilk's Λ, three other multivariate test statistics commonly seen in statistical software packages are (1) Roy's largest root, (2) the Hotelling-Lawley trace, and (3) the Pillai-Bartlett trace. Roy's largest root refers to the largest eigenvalue of $\mathbf{BE}^{-1}$ which is a natural extension of the univariate F = MS_b/MS_w, i.e., a measure of between- to within- association. The Hotelling-Lawley trace is the sum of the eigenvalues of $\mathbf{BE}^{-1}$, whereas the Pillai-Bartlett trace is the sum of the eigenvalues of $\mathbf{BT}^{-1}$. Since Wilk's Λ can be expressed as a product of eigenvalues of $\mathbf{ET}^{-1}$, it is noteworthy that all four of the multivariate test statistics are functions of some eigenvalues. Thus, understanding of the properties of these eigenvalues is a key to the interpretation of the multivariate problem, which will be the central thrust of this paper.

Turning to the problem of selecting an appropriate MANOVA test statistic, Olson (1976) suggested the use of Pillai-Bartlett trace over the other three because of its supposed robustness property in the presence of heterogeneous covariance matrices. However, Stevens (1979) found that there was virtually no practical differences among Wilk's Λ, Pillai-Bartlett trace, and Hotelling-Lawley trace concerning robustness; whereas in his rejoinder Olson (1979) reaffirmed his original recommendations on the robustness of the Pillai-Bartlett trace. Olson (1974) found that power differences among all four multivariate test statistics are generally small. Based on the summary information from empirical research on robustness and power, this paper recommends the use of Wilk's Λ over other three under ordinary conditions. In the case of serious violations of multinormality and covariance homogeneity, Pillai-Bartlett trace may be used to ensure the validity of the test. Moreover, this trace statistic can be obtained by submitting the data, converted to ranks or normalized scores, to a statistical computing package such as SAS, SPSS, or BMDP. Monte Carlo research has indicated that such a nonparametric approach may achieve reasonably valid results under violations of assumptions (Zwick, 1985).

Traditional MANOVA Follow-up Procedures

A significant MANOVA result only indicates that the groups are not exactly the same with respect to their outcome variable mean vectors. Just how the group mean vectors are different remains to be explored. Because multivariate tests of equality of mean vectors appear to be natural extensions of their univariate counterparts, most researchers are led to believe what's left to be done is to "locate" the variable(s) on which the groups show significant differences. The focus of the follow-up procedures inevitably centers on each individual outcome variable.

The most popular approach to follow up on a significant MANOVA result is to conduct p univariate F tests. One would assume that if the multivariate null hypothesis is rejected, then at least one of these univariate F's should be significant. Nonetheless, multivariate test and univariate test make use of different information. The univariate test concerns the relationship between only one outcome variable and the grouping variable whereas the multivariate test makes use of the covariance information among all p outcome variables. The covariance information translates into the correlations among p outcome variables. Consequently, there is no necessary relationship between multivariate significance and univariate significance. Timm (1975, p. 166) has already pointed out that the rejection of multivariate test does not guarantee that there exists at least one significant univariate F value. Pedhazur (1982) presented an empirical example of three groups and two outcome variables where the MANOVA test is significant at the .001 level, yet neither univariate test is significant, even at the .05 level. Stevens (1986, p. 131) explained that such a phenomenon is due to the high correlation among two outcome variables. When the correlation is strong, most error on the second variable is accounted for by the first variable, making multivariate error term much smaller, and consequently results in a more powerful test for MANOVA. If univariate F's are used to follow up on a significant MANOVA result under such a situation, researchers may find themselves in an embarassing dilemma, and are left at a loss over the interpretation of the MANOVA situation.

If one wishes to prevent the above mentioned problem from happening, multivariate test information must be used in the follow-up procedures. One common approach is to use Hotelling's T^2 to perform all pairwise comparisons for the grouping avriable on all outcome variables. In general, a significant MANOVA will generally produce some significant T^2's. Then researchers most often use univariate t-tests, each at a certain α level, to determine which of the individuals are contributing to the significant multivariate pairwise differences. To conroll experimentwise type I error rate at a specified level for all pairwise differences, Tukey's confidence intervals or Roy-Bose simultaneous confidence intervals are sometimes used.

Of course, multivariate test is not limited to pairwise differences only. Just as a univariate t-test can be used to test either a pairwise contrast or a complex contrast, so can Hotelling's T^2 be used to test their multivariate counterpart. A multivariate contrast defines the relations among the mean vectors associated with all levels of the grouping variable through a set of coefficients. Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ denote the population mean vectors, a multivariate contrast is given by $\underline{\psi}$ as follows:

$$\underline{\psi} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \cdots + c_k \mathbf{u}_k$$

To test the $H_0 : \underline{\psi} = \mathbf{0}$, the contrast is estimated by replacing the population mean vectors with the sample mean vectors:

$$\hat{\underline{\psi}} = c_1 \bar{\mathbf{x}}_1 + c_2 \bar{\mathbf{x}}_2 + \cdots + c_k \bar{\mathbf{x}}_k$$

Let $\mathbf{S}$ be the estimated common within-group covariance matrix, the Hotelling's T^2 for any contrast can be calculated as

$$T^2 = \left[\sum_{i=1}^{k} \frac{c_i^2}{n_i}\right]^{-1} \underline{\hat{\psi}}' \mathbf{S}^{-1} \underline{\hat{\psi}}$$

The computed T^2 can be transformed to an exact F by

$$F = \frac{N - k - p + 1}{(N - k)p} T^2$$

with p and (N - k - p + 1) degrees of freedom. As one might have guessed, if the transformed F value is found to be statistically significant, the typical practice would be to follow up on each outcome variable with a univariate t test.

A Challenge to the Traditional Practice

As mentioned in the beginning of this paper, the primary reason for conducting a MANOVA is to determine whether there is a grouping variable effect in the presence of a collection of outcome variables. Since the number of the outcome variables is more than one, these outcome variables may be to some extent interrelated to one another. In as much as one cannot necessarily judge outliers in the space of multiple predictors in a regression scenario based on a collection of univariate analyses, the same logic applies to the MANOVA situation: one may not be able to tell systematic differences among groups based on the information yielded by univariate tests alone. However, as discussed in the previous section, almost all traditional follow-up procedures to a significant MANOVA end up with interpretations based on univariate analyses only. Even most recent textbooks on multivariate methods (Morrison, 1990; Hair et. al., 1992) have their coverage of MANOVA follow-up procedures focus on individual outcome variables with multivariate test statistics (such as using a T^2 to test two-group differences) simply serving as some licenses for carrying out their later-stage univariate interpretations. What is missing from interpretations as obtained from all these follow-up procedures is the information about the correlations among the outcome variables. As a matter of fact, all follow-up procedures which yields interpretations that ignore the correlations among the outcome variable system follow the logic of pursuing univariate ANOVAs only after a significant MANOVA is detected. Therefore in this paper, we shall follow Huberty and Morris's track (1989) to collectively refer to these procedures as the MANOVA-ANOVAs approach.

The MANOVA-ANOVAs approach has been very common among educational researchers for a long time. A justification for such an approach has primarily been argued to control for the overall type I error rate. The rationale behind this justification goes that if we conduct many univariate ANOVAs with the α_c set at a prespecified level for each outcome variable, we could end up with an experimentwise α_E much larger than originally anticipated. Thus, in order to prevent spurious ANOVA test statistics from being detected as statistically significant, MANOVA should be brought in to provide safeguard against it. This is in fact very much like the idea of Fisher's protected F test. However, the ability of a protected F test to completely control for type I error probablity has been open to criticism (Bird & Hadzi-Pavlovic, 1983; Bray and Maxwell, 1982, p. 343), The main arguements against the protected approach is that it works only under the situation where H_o is 100% true, but does not provide protection for the individual F tests if the H_o is false. Specifically in the MANOVA-ANOVAs context, Wilkson (1975) found that such a notion also failed to gain support in empirical research.

Now suppose that the MANOVA-ANOVAs approach was indeed able to control the experimentwise error rate at the prespecified level, it would then seem that MANOVA is no more than conducting a set of univariate ANOVAs. No more information whatsoever would be available to researchers besides what can be known from univariate ANOVAs. This can be somewhat disturbing to those who wish to see things "multivariately". Huberty and Morris (1989) argued that there is no need to conduct a MANOVA as a preliminary step to conducting the univariate ANOVAs. Instead, the guiding force for choosing either a MANOVA or multiple univariate ANOVAs should be the nature of the reseach questions. Another words, research questions that the researcher has in mind should dictate, be it multivariate or univariate. For an excellent discussion on the difference between "multivariate" and "univariate" questions, readers may refer to his original paper appearing in Psychological Bulletin of 1989 issue. In the rest of the section, we will illustrate how simple ways of overcoming spurious type I error rate when performing multiple univariate ANOVAs not at the expense of sacrificing statistical power can be adopted, to further challenge the non-necessity of the traditional logic of MANOVA-ANOVAs pairing approach. Then, multivariate questions which require "multivariate" examination of the data are proposed to set tone for the next discussion on extended multivariate follow-up procedures.

Type I Error Rate and Power

Let's Denote the set of univariate null hypotheses for p outcome variables to be H_1, H_2, ..., H_p. To control for the experimentwise type I error rate at a prespecified level of α_E, researchers typically invoke the Bonferroni procedure which states that if we test each hypothesis at the α_c level, then the overall experimentwise type I error rate should not exceed the sum of total α_c's, namely, $\alpha_E \leqq \sum \alpha_{ci}$. Therefore, if we set $\alpha_c = \alpha_E/p$ for testing each univariate null hypothesis, the overall type I error rate will not exceed our nominal level of α_E. This procedure has not been used to a great extent by applied reseachers, perhaps due to its inevitably low statistical power in the presence of a large number of outcome variables. Taking the example of 10 outcome variables (p=10), to control for α_E=.05 each univariate null hypothesis would have to be tested at α_c=.005. If we have 20 outcome variables (p=20), α_c would have to be set at .0025. It is not difficult to see that the control for overall type I error rate is accomplished only at the expense of a tremendous loss in statistical power. There is certainly no surprise that it has not been received with much enthusiasm, even in the context of ANOVA multiple comparisons.

Holm (1979) proposed a modified Bonferroni procedure with which statistical power can greatly improve. First we order the empirically obtained p-values, say we obtained 10 p-values (suppose we have 10 outcome variables) from their univariate ANOVA analyses, we order them from smallest to the largest and denote the ordered $\{ p_i \}$ by $p_{(1)} \leqq p_{(2)} \leqq \ldots \leqq p_{(10)}$. Tied p-values can be ordered arbitrarily. Suppose i^* is the integer associated with a specific null hypothesis between H_1 and H_p (inclusive), null hypothesis $H_{(i^*)}$ may be rejected if

$$p_{(i^*)} < \alpha_E / (p - i^* + 1)$$

The value of α for Holme's procedure for testing the first null hypothesis H_1 is the same as that obtained by the Bonferroni procedure, namely, when $i^*=1$, $p_{(i^*)}$ would have to be less than α_E / p for the null hypothesis to be rejected. If the null hypothesis is not rejected at H_1, no further hypothesis testing needs to be carried out for the rest of the null hypotheses. If the hypothesis is rejected at H_{i^*}, then α for rejecting the next null hypothesis H_{i^*+1} will be adjusted upward. Holm has proved that this procedure guarantees that there is at most α_E chance of rejecting at least one of the true null hypotheses. The increased power comes from the fact that Holme's procedure may reject some hypotheses not rejected by the Bonferroni procedure.

Univariate Questions Versus Multivariate Questions

Although some applied researchers take the approach of formulating research questions after trying out fitting many feasible statistical designs to their data (this approach is often called "data snooping"), there is a consensus among conscientious researchers that research questions should guide the selection of the appropriate statistical methods and not otherwise. Huberty and Morris (1989) were among the earliest methodologists to call for attention to the issue of univariate versus multivariate research questions in the context of MANOVA.

Research questions which intend to discover the effect of the grouping variable on individual outcome variables are univariate questions. For instance, suppose we wish to examine the effect of different teaching strategies on three outcome variables: attitude, achievement, and the ability to generalize to other related tasks. If we intend to answer three separate questions like: Do the different strategies differ with respect to the resulting students' attitude? Do the different strategies differ in the resulting students' achievement? Do the different strategies differ in the students' ability to generalize the acquired knowledge to other related tasks? Three separate ANOVAs can certainly be employed to anwser such research questions and they do not require the consideration of the inter-correlations. Holme's procedure can then be applied under such a circumstance to both insure the overall type I error protection and sufficient power to detect significant effects. There is, however, no need whatsoever, to conduct a MANOVA for answering such questions.

On the other hand, any time researchers work with a number of outcome variables, it is almost universally true (unless empirically manipulated otherwise) that some correlations exist among these outcome variables. Looking only at the results from multiple ANOVAs cannot provide researchers with a complete picture. It is like putting pieces of information randomly together. The question being asked by MANOVA is, Are there any overall effects associated with the grouping variable(s) significant in the presence of a collection of outcome variables? These effects can be either main effects or interaction effects. As mentioned earlier, all multivariate test statistics involve some forms of covariance (or sums of squares and crossproducts) matrices for the outcome variables, namely, between, within, or total. Since the computation of the test statistic uses the information based on their inter-relationships, it would only seem reasonable that our explanations of the MANOVA results also make use of that inter-correlational information. Thus, our choice of MANOVA over multiple ANOVAs should be dependent on whether we have research questions which would require the investigation of the correlational patterns among outcome variables should the effect of the grouping variable be found significant.

Huberty (1986) suggested three purposes for using multivariate tests that involve the consideration of the inter-correlations among outcome variables in the context of MANOVA: (1) determining outcome variable subsets that account for group separation; (2) determining the relative contribution to group separation of the outcome variables in the final subset; and (3) identifying underlying constructs associated with the obatined MANOVA results. The first purpose is commonly referred to as the "variable selection problem", and has been studied extensively by MaCabe (1975). The second purpose has to do with the "variable ordering problem" and has been discussed by some methodologists (Bibb and Roncek, 1976; Blalock, 1961; Eisenbeis, Gilbert, & Avery 1973; Huberty and Wisenbaker, 1992). The third purpose has to do with the identification of the underlying construct(s) which is/are responsible for the group separation. It is probably considered as the most pertinent purpose of the three with respect to the decision of conducting a MANOVA analysis. None of the above purposes could be attained by performing multiple ANOVAs. Therefore, it is the view of this paper that unless the inter-correlations among outcome variables can be somehow accounted for under some reasonable framework, it cannot be viewed as a multivariate analysis.

Having defined MANOVA as a statistical procedure to analyze grouping variable effect in a strictly "multivariate" way, it would then only be appropriate to follow up on a significant MANOVA result with some procedures which may help us account for the grouping variable effect "multivariately". In the next section, we shall see how some of the well-established multivariate methods may be used in the context of MANOVA to help researchers extract information which would not otherwise be obtained had multiple univariate analyses been adopted.

Some Extended Follow-up Procedures

Consider a one-way layout of p response measures taken across k groups. Let **E** be the p×p error matrix of sums of sqaures and cross products (SSCP), and **B** be the p×p between-group SSCP matrix. As mentioned earlier, the test of MANOVA essentially compares the relative magnitudes of **B** and **E** in terms of their determinants. All of the information in the SSCP matrices are used to derive the test statistic. Therefore it appears reasonable we should also interpret the overall difference among k groups in the context of all p outcome variables. Another words, we should use the covariance (or correlation, equivalently) information to explain for the overall grouping effect. In this section we shall see how this overall difference can be attributed to some underlying dimension(s) which we may hold responsible for the variations displayed among k groups and how we can identify what these underlying dimensions may represent in reality.

Discriminant Analysis

This is the classical approach to the identification of the underlying dimensions first proposed by Tatsuoka (1969) which he claimed as "probably the most significant development during the 1960's". I shall use an easier-to-understand way to present the essence of this approach in this section. A good understanding of the originally proposed discriminant analysis is a key not only to the identification of the underlying structures of MANOVA, but also the understanding of all later MANOVA follow-up developments.

Let's begin with the simplest case of two groups G_1 and G_2 from which the p outcome measurements were taken on each individual. Denote **X** to be the vector of p outcome measurements. Fisher (1936) proposed that finding a linear combination **b'** of **X** so that the ratio of the difference in the means of the linear combinations in G_1 and G_2 to its variance is maximized. Thus we maximize the criterion:

$$\lambda = \frac{\mathbf{b}'\mathbf{u}_1 - \mathbf{b}'\mathbf{u}_2}{\mathbf{b}' \sum \mathbf{b}}$$

The solution of **b** is not unique. Instead, it is proportional to $\sum^{-1}(\mathbf{u}_1 - \mathbf{u}_2)$. In applications, we would just replace parameters with their respective sample estimates:

$$\hat{\mathbf{b}} = \mathbf{S}^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)$$

where $\mathbf{S}^{-1}$ is the inverse of the pooled within sample variance-covariance matrix.

The vector $\hat{\mathbf{b}}$ is a linear composite of the original outcome variables on which the difference between two group means is maximal, relative to the within-group variance. Therefore it acquired the name as a " linear discriminant function" (LDF), which indicates this linear function maximally discriminates the concerned two groups. Consequently, to explain how two group means differ, it would then be reasonable to examine what the resulting LDF represent instead of examining individual outcome variables separately. The mean value of the LDF is commonly referred to as the group "centroid", which is the geometric center of the group in the p-variate multivariate space. Denote the group centroids by $\overline{\mathbf{Y}}_\mathbf{i}$, we can easily see how the Mahalanobis distance (difference in group centroids) is related to the LDF:

$$\overline{\mathbf{Y}}_1 - \overline{\mathbf{Y}}_2 = \overline{\mathbf{b}}'(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) = (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2)'\mathbf{S}^{-1}(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2)$$

In the case of two groups, one LDF will account for all the differences. In the case of several groups, however, more than one LDF will be needed as we shall see in the following.

For the number of groups exceeding two (say k groups), Fisher's criterion is analogous to the the ratio of the between-groups to within-groups sums-of-squares for the k groups on the first linear composite of the p outcome variables, namely

$$\hat{\lambda} = \frac{\hat{\mathbf{b}}'\mathbf{B}\hat{\mathbf{b}}}{\hat{\mathbf{b}}'\mathbf{E}\hat{\mathbf{b}}}$$

The vector $\hat{\mathbf{b}}$ is calculated based on the maximization rule for $\hat{\lambda}$, by taking the partial derivative with respect to $\hat{\mathbf{b}}$ and set it equal to zero, namely $\partial\hat{\lambda} / \partial\hat{\mathbf{b}} = \mathbf{0}$. After some simplication we have

$$(\mathbf{E}^{-1}\mathbf{B} - \hat{\lambda}\mathbf{I})\hat{\mathbf{b}} = \mathbf{0}$$

In this form, it is easy to see that $\hat{\lambda}$ as computed according to Fisher's criterion is the largest eigenvalue of the matrix $\mathbf{E}^{-1}\mathbf{B}$, and $\hat{\mathbf{b}}$ is its corresponding eigenvector (the first LDF), whose elements are also known as the "discriminant weights" of this linear composite. In general, the rank of $\mathbf{E}^{-1}\mathbf{B}$ will be equal to the minimum of p and k-1, m=min(p,k-1). Hence, if the eigenvalues of $\mathbf{E}^{-1}\mathbf{B}$ are distinct, there will be m=min(p, k-1) linear composites, each corresponding to a specific successive eigenvalue. In most MANOVA research context, the number of outcome variables is greater than the number of levels in the grouping variable (p > k), therefore we may get (k-1) LDFs. Again, the discriminant weights of the LDFs are unique only up to a constant of proportionality. They are printed by most of the popular statistical computing softwares. In SPSS Discriminant program, they are printed under the heading UNSTANDARDIZED CANONICAL WEIGHTS, whereas in SAS GLM MANOVA subroutine, they are printed under the heading CHARACTERISTIC VECTOR.

Not all LDFs contribute equally to grouping variable effect. Denote $\hat{\lambda}_j$ to be the jth eigenvalue of $\mathbf{E}^{-1}\mathbf{B}$, where j=1,2,3,...,m. The relative magnitude of the respective jth eigenvalue gives a descriptive index of the importance of that specific LDF toward the grouping variable effect. Similar to the interpretation for importance of a linear variate as extracted from a principal component analysis, we may explain each eigenvalue as a percentage of the total variance accounted for, that is $\hat{\lambda}_j / \sum_j \hat{\lambda}_j$. We can then determine which of the LDF capture(s) most of the grouping variable effect.

To determine how many LDF's are statistically significant, Bartlett's statistic may be employed. Since Wilk's Λ can be expressed as a function of the eigenvalues of $\mathbf{E}^{-1}\mathbf{B}$,

$$\Lambda = \prod_{\mathrm{j}=1}^{\mathrm{m}} (1+\lambda_{\mathrm{j}})^{-1}$$

we can rewrite the equation as

$$V = \{(N-1)-.5(p+k)\}\sum_{\mathrm{j}=1}^{\mathrm{m}} \ln(1+\hat{\lambda}_j)$$

Thus, the significance of the jth LDF can be assessed by computing

$$V_j = \{(N-1)-.5(p+k)\}\ln(1+\hat{\lambda}_j)$$

This statistic is checked against the specified percentiles of a Chi-square variate with p+k-2j degrees of freedom. It is noted that two major programs of SAS, PROC GLM and PROC CANDISC, provide the signficance test information for all LDFs using an F statistic based on Rao's (1973, p. 556) approximation to the distribution of the likelihood ratio.

Associated with each eigenvalue is a canonical correlation, which is the sqaure root of the complement of the Wilks lambda, is also commonly reported in computer printouts. The canonical correlation is an indicator of the extent of relationship between the grouping variable and the LDF itself. All of the information mentioned above in this section can be obtained using the following SAS program commands (partial), coding the grouping variable to be GRP here for the ease of reference:

```
PROC GLM; CLASS GRP;
   MODEL Insert outcome variables here = GRP;
   MANOVA H=GRP/CANONICAL;
```

or using SAS PROC CANDISC procedure,

```
PROC CANDISC ALL;
```

CLASS GRP;
VAR Insert outcome variables here;

Identififying Underlying Constructs

What on earth do these LDFs may represent in reality? The identification of the LDFs is a state-of-the-art task in itself since there are an infinite number of ways to rotate the resulting LDF axes mathematically (Bentler and Huba, 1980; Tatsuoka, 1973). Fortunately, in most of the instances, the resulting LDFs are readily interpretable by experienced researchers. In general, researchers should refrain from making interpretations based on discriminant weights directly (see Bray and Maxwell, 1982, p. 344-345). Tatsuoka (1971) recommended the use of standardized LDFs weights. However, it was pointed out by many other methodologists (Finn, 1974; Timm, 1975; Borgen and Seling, 1978; Wilkinson, 1975) that the these coefficients are highly susceptible to the multicollinearity condition of the outcome measures in much the same way as the interpretation of beta weights in multiple regression may suffer from the multicollinearity condiiton (Darlinton, 1968).

For interpretations of LDFs, it has been recommended (with a sound rationale) to use what are called "discriminant loadings" (Huberty and Wisenbaker, 1992; p. 77). These loadings are the simple product moment correlations between each LDF and each of the outcome variables. It should be noted that these correlations are computed as the within-groups correlations, computed as

$$r = \frac{\sum_j \sum_i (X - \overline{X}_j)(Y - \overline{Y}_j)}{\sqrt{\sum_j \sum_i (X - \overline{X}_j)^2 \sum_j \sum_i (Y - \overline{Y}_j)^2}}$$

where Y indicates the LDF scores computed for each individual. It should not be confused with the more familiar total-group correlations. Just as in common factor analysis and principal components analysis, the substantive psychological contruct/underlying-dimension defined by each LDF maybe identified based on the pattern of loadings. The variables that correlate highly with an LDF are conceptually put together to arrive at a name for that LDF, now representing a construct (or underlying dimension). In SPSS, these loadings are contained in the "Structure Pattern Matrix". In the PROC GLM program of SAS, on the other hand, three kinds of structure matrix are produced: TOTAL CANONICAL STRUCTURE, BETWEEN CANONICAL STRUCTURE, AND WITHIN CANONICAL STRUCTURE. The loadings of the WITHIN CANONICAL STRUCTURE matrix should be used for interpreting the underlying construct.

Thorndike and Weiss (1973) recommended the use of discriminant loadings over the standardized weights in the argument of their perceived property of sampling stability. However, Huberty (1975) found that neither standardized weights nor discriminant loadings were very stable under repeated sampling conditions. This fight over which of these two coefficients should be the one to use lingered on for quite some time until later literature shed more light into the nature of the arguement (Bray and Maxwell, 1982; Huberty, 1984; Huberty and Wisenbaker, 1992). Actually, these two coefficients represent different aspects of the MANOVA question. The discriminant weights are analogous to regression weights and thus represent the unique contribution of a variable in addition to other variables in the model. On the other hand, the discriminant loadings, being the correlations between LDF and the outcome variables, are more useful for interpreting the substantive nature of the LDF. Consequently, if we intend to find out what underlying dimensions are present which can be held responsible for the grouping variable effect, discriminant loadings provide the answer to the nature of the LDFs.

Relative Variable Importance

When a significant grouping effect is detected by the MANOVA test statistic, it certainly seems to be a natural response in many circumstances to ask questions like which outcome variables are more important than others in explaining the grouping variable effect. The above mentioned controversy of choosing either discriminant weights or loadings reflects different views of the relative importance issue. When a researcher poses a research question about whether some variables are more important than others in the MANOVA context, he/she should specify what aspect the relative importance relates to. In fact, different views of relative variable importance yield different lists of "important" variables (Huberty and Wisenbaker, 1992). In the following, I shall present some commonly encountered concepts of relative variable importance and discuss their applications.

One seemingly attractive index for relative variable importance is discriminant loading, which describes the correlation between the outcome variable and the LDF. This view is based on the argument that outcome variables which give more weight to the resultant underlying construct should be considered as more important than other variables which give less weight. The structure loadings (computed as shown under the previous subsection heading of Identifying Underlying Construct) are typically squared to indicate the strength of association for rank ordering purpose. In the two-group case, there is only one single LDF, and therefore only one set of r^2 values. It turns out that, a variable ordering based on these r^2 values is identical to an ordering based on the univariate Student t values — see Huberty (1972) for a mathematical proof. In the case of more than two groups (k > 2), we generally have k-1 sets of r^2 values (suppose the number of outcome variables exceed the number of groups). Since an eigenvalue λ_j reflects a proportion of variation in the p-variable system accounted for by the jth LDF, Tatsuoka (1988, p. 213) suggested a weighted value as

$$c_i = \sum_{j=1}^{m} \lambda_j r_{ij}^2$$

to derive at an ordering index for the ith variable across the m LDFs. However, it also turns out, similar to the two-group case, that an ordering based on these c_i values is identical to an ordering by the p univariate F values — see Rencher (1988) for proof. Thus, the use of c_i values as variable ordering indices ignores the intercorrelations among other p-1 outcome variables, and therefore is generally not recommended in a multivariate context, Nevertheless, if one is interested in the ranking of variables in terms of relative contributions to the specific LDF construct, then rank ordering r^2 values for that given LDF may be a reasonable assessment approach (Lambert, Wildt, and Durand, 1991).

Some researchers take the relative importance issue as the extent of contribution to the LDF scores. This view of importance is based on the argument that scores on the first LDF yields maximum group separation, and the standardized weight, b_{ji}^*, should be used in this respect. As mentioned earlier, the values of b_{ji}^*'s are problematic in the presence of multicollinearity among p outcome variables, a transformed version of b^*_{ji}'s which takes into account the correlations among the variables have been proposed (Pedhazur, 1982, pp. 81, 207). The suggested index of variable importance for variable i on LDF j is:

$$u_{ji} = b^*_{ji} / r^{ii}$$

where r^{ii} is the ith diagonal element of the inverse of the within-groups correlation matrix, commonly called "variance inflation factor" in the regression jargon. It is computed as

$$r^{ii} = [1 - R_i^2]^{-1}$$

where R_i^2 is the square of the multiple correlation between variable i and the remaining (p-1) outcome variables. If such an approach is taken, it would be reasonable to explain the relative variable importance with respect to each LDF. For an instance of k=3, we may have two significant LDFs. The first LDF may distinguish group 1 from groups 2 and 3, whereas the second LDF may distinguish between group 2 and group 3. One would then obtain two lists of relative variable importance for LDF 1 and LDF 2 separately.

The b^* values can be obtained from SPSS-X or SAS softwares under similar headings involving "STANDARDIZED COEFFICIENTS". The r^{ii} values can be obtained via SAS in a two-step procedure. First we obtain the within-groups correlation matrix from the SAS GLM program (printed under the heading PARTIAL CORRELATION COEFFICIENTS FROM THE ERROR SS&CP MATRIX). Then we invert this matrix using SAS IML program. The diagonal elements of this inverted matrix are the computed r^{ii} values.

Another view reflected on the relative importance issue is the employment of the "stepdown" analysis in the context of MANOVA (Roy, 1958). It is in effect a form of analysis of covariance in which outcome variables are entered according to some a priori order to test the relative contribution of the later-entered variables toward the grouping variable effect. Stevens (1972) recommended this procedure instead of the the MANOVA-ANOVAs approach as a follow-up procedure to MANOVA, because at least some consideration of the variable inter-correlations is considered in the successive steps. However, Finn (1974) warns that if there is no a priori or logical ordering of the outcome variables, the step-down procedure is of little use and may even lead to erroneous conclusions.

What if no a prior order can be specified by the researcher? A stepwise discriminant analysis (backward or forward) is probably by far the most popular approach under such a situation. Here we are confronted with a problem of how large a number of outcome variables is acceptable in a MANOVA situation. In general, methodologists warn against the use of a large number of outcome variables (Stevens, 1986; p. 187; Olson, 1974, p. 906). In the following discussion, we shall assume that researchers have exercised discretion on all outcome measures involved in the MANOVA, and have judged that their inter-correlations have to be taken into account in the explanation of the grouping variable effect.

All major statistical softwares provide programs for carrying out the stepwise discriminant analysis. The concepts of two statistics are important in a stepwise discriminant analysis: F-to-enter and F-to-remove. The F-to-enter value, denoted as F_i, for a sequentially entered outcome variable X_i tests the null hypothesis of the quality of k means on X_i when the previously entered q variables are partialed out. Let $\Lambda(X_i|\mathbf{X})$ denote the partial Wilks lambda value reflecting separation yielded by X_i in addition to the previously entered variables denoted by vector **X**, then

$$R_i^2 = 1 - \Lambda(X_i|\mathbf{X})$$

represents an η^2-like index. A larger F-to-enter value is equivalent to a larger R_i^2, indicates that X_i contributes significantly in addition to **X**. The R_i^2 values are produced by SAS STEPDISC procedure under the heading "squared partial correlation". It should be noted that the Wilks's lambda SPSS Discriminant program produced next to the F-to-remove is not a partial lambda, but the lambda that would result if the corresponding variable would be entered in the analysis.

In contrast to the F-to-enter value, the F-to-remove value, denoted as $F_{(i)}$, for X_i reveals the drop in separation effect of the grouping variable when X_i is removed from the model. Again, a larger $F_{(i)}$ value would indicate more importance. Suppose we would like to judge the relative importance of each variable in the presence of <u>all</u> other variables (as we have pressumed already), the $F_{(i)}$ values in the <u>last</u> step of the SAS STEPDISC procedure should then be used as indicators, which are listed under the heading of just "F" (The F values in the summary table of the SAS STEPDISC procedure are actually the F-to-enter values, which is <u>not</u> recommended for use for the ordering purpose). When used in this perspective, each $F_{(i)}$ is equivalent to the test statistic produced by analysis of covariance (ANCOVA) with X_i as the criterion variable and the remaining p-1 variables as the covariates. In fact, in a two-group context, an ordering according to $F_{(i)}$ values is equivalent to ordering according to u_i, (see Urbakh, 1971, p. 532). Again, it should be noted that the Wilks's lambda next to the F-to-remove on an SPSS Discriminant printout is not a partial lambda, but the lambda that would result if X_i was deleted. The partial program commands for SAS PROC STEPDISC are given below:

```
PROC STEPDISC SLE=.9999;
   CLASS GRP;
   VAR Insert outcome variables here;
```

The SLE=.9999 options specifies the significance level to enter to be .9999, which guarrantees the entry of all outcome variables. Since all outcome variables have been entered in the final model, the option of forward, backward, or stepwise makes no distinctions whatsoever. Again, what we are interested is the F-to-remove values (under the "F" heading on the printout) as produced in the last stage of the stepwise procedure, not to be confused with the summary information table.

Multivariate Group Contrast

Apparently, multivariate contrasts appear to be very seldom of a concern to researchers who use MANOVA (at least so in research publications). This is, perhaps, not too surprising, because very few multivariate methods textbooks discuss them in any great detail. Of the 15 applied multivariate textbooks collectively reviewed by Huberty and Barton (1990), only six even hint at multivariate group contrasts. Of these six, one (Dunteman, 1984, pp. 97-103) suggests only trend contrast, and two (Harris, 1985, pp. 402-405; Marasciulo & Levin, 1983, pp.324-329) group contrasts of means on linear composites. Two textbooks (Cliff, 1987, pp.402-405, 414-416; Lunneborg & Abbott, 1983, pp.346-353) discuss contrasts in non-specific terms, while another (Stevens, 1992, pp.211-225) goes into some detail on conducting multivariate contrast statistical significance test. The most recent comprehensive coverage (and perhaps the only one ever made) of the multivariate contrast topic in terms of "multivariate" interpretation of the contrast effect was given by Huberty, Chou, & Benitez (1994).

The notion of a multivariate group contrast is the same as that of a univariate group contrast except that vectors of outcome variable means are involved. The focus here is on contrasting groups simultaneously across all outcome variables, not only on deriving the test statistic, but also on the interpretations of the contrast effect. The test statistic may be, equivalently, the Hotelling two-group T^2, the Wilks's lambda, or the Mahalanobis D^2 in all situations. Of course, the error covariance matrix typically used with any of these statistics is the pooled-within covariance matrix. As for making simultaneous interpretations for the specific contrast effect, two approaches mentioned earlier may also apply to the group contrast, namely, identifying contrast definition and relative variable importance with respect to their contribution to the specific contrast effect.

We first of all have to explain what construct/dimension it is which underlies the contrast effect. Any contrast, be it a pairwise comparison or a complex contrast, is conceptually viewed as a two-group analysis. Therefore, there is only one possible LDF associated with the contrast effect. It is this linear composite of the outcome variables that is a basis for the construct. The statistic useful for labeling this construct is structure correlation (called structure loading) － the correlation between scores on the outcome variable and scores on the LDF. With respect to the variable importance issue, the interpretation considered here is relative contribution to the grouping variable effect, as reflected in the complement of the Wilks's lambda. Thus, for p outcome variables, one would have to conduct p contrast analyses each with p-1 variables. The analysis that yields the largest lambda (means smallest separation), indicates the deleted variable is the most important one. In the following, we shall see how statistical softwares can be used to help us find infomation in these two perspectives concerning a multivariate group contrast.

Two statistical computer packages considered here are SAS (Version 6.07) and SPSS (Release 4.1). As discussed above, two sets of information are desirable in addition to the test statistic to complete a contrast analysis. For the sake of illustration ease, we will consider the one-factor data situation with three levels (grouping variable is coded as GRP, k=3). The two contrasts considered are: $\mathbf{u}_1 - \mathbf{u}_2$ and $\mathbf{u}_1 + \mathbf{u}_2 - 2\mathbf{u}_3$. The partial program commands to conduct construct analyses for the two packages are given in Table 2.

Table 2. SAS and SPSS program commands for contrast analyses:

SAS	SPSS
PROC GLM; CLASS GRP;	MANOVA (variable list here)
(Model statement here)	CONTRAST (GRP) =
CONTRAST '1 VS 2' GRP 1 -1 0;	SPECIAL(1 1 1 1 -1 0 1 1 -2)/
CONTRAST '12 VS 3' GRP 1 1 -2;	PARTITION (GRP)/
MANOVA H=GRP/CANONICAL;	DISCRIM=COR/
	DESIGN=GRP(1),GRP(2)/

Neither SAS nor SPSS packages provide a subroutine which produces the F-to-enter or F-to-remove values directly for contrast analyses. Here we show some feasible ways to obtain information about relative variable importance. If we do not have a priori hypothesis about the order of importance for p outcome variables, F-to-remove values are preferred over F-to-enter values for rank ordering relative variable importance. To get this information with SPSS Discriminant program, two runs are needed. The first run is used to obtain some matrix information: a pooled within correlation matrix, group sizes, and group means and standard deviations for all outcome variables. This matrix can then be used in the second run to obtain the ordering information for the contrast analysis as shown below:

```
DISCRIMINANT GROUPS=(1,3)/
   VARIABLES (list here)
   METHOD=Wilks/FIN=.00001/
      FOUT=.00001/
   MATRIX=OUT(*)
DISCRIMINANT GROUPS=(1,2)/
   VARIABLES (list here)
   METHODS=WILKS/FIN=.00001
      FOUT=.00001/
   MATRIX=IN(*)
```

The above program contrasts group 1 to group 2. There is a problem with the direct use of the second run. It is that the resulting Wilks lambda and the F-to-remove statistics may be incorrect. The reason for this is that the computational algorithm transforms the error correlation matrix to an error SSCP matrix using the individual group sizes for only those groups involved in the contrast, rather than the total group size. The structure loadings obtained with the direct use of the two runs are, however, correct (same as those obtained using MANOVA program as described in Table 2). The correct test and ordering information may be obtained by editing the output file so that the sizes of the groups involved are adjusted to obtain a correct multiplier to transform the correlation matrix. The use of the two runs for conducting a complex contrast analysis is even more troublesome. It requires additional editing of the output matrix file. This editing involves the construction of two "groups" by combining (according to the specified contrast) original group sizes, means, and standard deviations. Such editing is not detailed in this paper.

An Example[1]:

The present example illustrates the attempts to discover, by the use of MANOVA technique, the real views of three categories of people: community residents (parents of current elementary students), school principals, and school teachers, on certain issues which have drawn public concerns in Taiwan as the educational systems and practices are now finally starting to pay attention to the long-forgotten aspect of the school-community relations. These issues take on the form of nine questions as listed in Table 3, and they represent the nine outcome variables in our example. The data were collected recently from a large scale

---

[1]The author wishes to thank Dr. Lin Ching-Dar for generously supplying his data for the illustration purpose of the current paper. Inquiries about the data source used in this section should be sent to Dr. Lin Ching-Dar, whereas requests for reprints and technical details of the methodology should be addressed to the author. Both are teaching at the same college.

mail survey sent out to randomly sampled elementary school principals, teachers, and parents in the entire Hualien area (林清達, 民 83).

Table 3. List of variable descriptions

Code	Variable Descriptions
V1	Can the role of the principals' participation promote the development of the community?
V2	Should principals serve as liaisons with the community?
V3	Should principals assist in communicating government policies to community residents?
V4	Should teachers maintain contacts with student's families?
V5	Should teachers actively participate in the community activities?
V6	Should schools open their facilities to community residents?
V7	Should schools provide services such as extended education and assist in sponsoring cultural and recreational events?
V8	Should shools periodically sponsor activities that would allow interactions between teachers, parents, and the students?
V9	Should the community residents be involved in the school administration?

It is worthwhile to mention that our approach differs from the usual logic of using MANOVA simply as a way of controlling overall "alpha" (experimentwise type I error rate), and following significant MANOVA with a series of univariate ANOVAs to locate the variables where the differences among three categories of people might exist. Instead, we used MANOVA because we treated all nine variables of concern as an integral part of the picture and wished to extract some useful information which could not otherwise have been made known by examining the difference of these groups of people on each variable alone. Putting results based on univariate analyses together in our view is inadequate since they ignore the correlations which may very well exist among the dependent variables. Our approach, therefore, is to essentially make use of information given by the correlation matrix of the nine variables through the innovative techniques associated with MANOVA mentioned in earlier sections. Specifically, we wish to provide answers to the following research questions:

Are there some simpler underlying dimensions which may beaccountable for the observed differences among three categories of people on the nine dependent variables? If there are, can we make sense for these underlying dimensions? How are the three categories of people differ on these underlying dimensions?

We shall first use the omnibus MANOVA significance test to determine if systematic differences exist among three categories of people in the nine-variable multivariate data space. Then, we shall pursue to answer our first research question through the help of discriminant analysis as we have discussed earlier. The basic idea of using discriminant analysis involves the identification for the linear composite(s) of the original nine outcome variables. In the context of our one-way MANOVA layout here, we may have a maximum number of two LDFs. Then, we shall consider the issue of variable importance for explaining the differences among three groups. Finally, contrast analyes were made for community residents versus teachers, principals versus teachers, and principals versus teachers and residents combined. All statistical computing was done using SAS 6.07 version run on an IBM-compatible 486-DX33 PC computer.

Data Description and Assumption Tenability

The original questionnaire was designed in three forms for three different categories of people to complete. After careful deliberation was made to sort out items useful for our illustration, a revised version of 32 items was formed to assess people's opinions on the nine questions with each question consisting of two to five inter-related individual items. Each individual item took on a Likert-type scale of 1 to 5. The scoring of the items was made in a way that higher scores indicate a higher extent of responding 'Yes' to the question being asked. The nine outcome variables were formed as subscales by grouping similarly-natured questions together. The score to each question was then computed as the sum of its individual items. We obtained data from 38 elementary school principals, 300 elementary school teachers, and 326 community residents with children in elementary school. Although we had a total sample size of 664, 11 observations were found to have missing values on at least one of the nine measures and were thus unusable for later multivariate analyses. Consequently, we decided to retain only 653 complete records. This data formed the sample of our study upon which our conclusions were made. The summary information on nine measures for the three categories of people was presented in Table 4.

Table 4. Descriptive statistics for the nine variables

Code	k	Community Residents (n=326)			Principals (n=38)			Teachers (n=300)		
		Mean	STD	Alpha	Mean	STD	Alpha	Mean	STD	Alpha
V1	4	16.44	2.22	.83	18.05	1.75	.81	17.29	1.93	.85
V2	3	11.36	1.75	.61	12.11	1.83	.72	11.74	1.75	.70
V3	3	11.57	1.79	.72	12.47	1.96	.77	11.71	1.50	.55
V4	4	15.82	1.98	.54	17.68	1.45	.46	15.86	1.84	.44
V5	4	17.96	2.94	.76	18.84	3.00	.77	17.52	3.23	.81
V6	2	7.72	1.34	.62	8.42	1.31	.67	7.82	1.23	.59
V7	3	11.52	1.86	.74	11.95	1.64	.57	10.96	1.97	.78
V8	4	16.13	2.46	.90	17.61	1.87	.85	15.96	2.31	.90
V9	5	17.38	2.80	.69	16.95	3.03	.66	16.96	3.09	.74

Note. k is the number of items in each scale that are ummedover to represent the variable. Alpha is the internal consistency index commonly known as the Cronbach's Alpha.

Total-Sample Correlation Coefficients

Variable	V1	V2	V3	V4	V5	V6	V7	V8	V9
V1	1	0.52	0.42	0.32	0.31	0.31	0.30	0.33	0.30
V2	0.52	1	0.52	0.30	0.40	0.35	0.37	0.35	0.30
V3	0.42	0.52	1	0.35	0.33	0.34	0.34	0.37	0.19
V4	0.32	0.30	0.35	1	0.36	0.32	0.36	0.44	0.14
V5	0.31	0.40	0.33	0.36	1	0.40	0.56	0.39	0.48
V6	0.31	0.35	0.34	0.32	0.40	1	0.50	0.45	0.32
V7	0.30	0.37	0.34	0.36	0.56	0.50	1	0.50	0.45
V8	0.33	0.35	0.37	0.44	0.39	0.45	0.50	1	0.41
V9	0.30	0.30	0.19	0.14	0.48	0.32	0.45	0.41	1

Note. All of the correlations are significant at .001 level.

A closer preliminary examination of the data indicated that most of the nine dependent variables displayed slight- to medium-degree negative skewness in their distributions. This came as no surprise to us because we knew beforehand from many public hearings concerning educational policies and from discussions in parent-teacher meetings that the majority of people tend to respond somewhat positively to these nine questions as listed in Table 3. Also, as one could expect, a good proportion of the observations (greater than .10) appeared as outliers in the multivariate data space and thus made the multivariate normality theoretically required for parametric multivariate methods seem untenable (various transformation procedures also failed to produce even close-to-multinormality distributions). We were thus confronted with considering the option of removing outliers so that our data analysis could proceed more smoothly. However, after sound deliberation it appeared to us that these so called "outliers" were in reality part of the general population we wish to make our inferences to. Their responses simply represent the commonly seen pattern of diversity of opinions over these issues. Removing these outliers would mean reducing the complexity of the real world situation into some artificially simplified unrealistic situation. That would be of little value to the practice-oriented researchers. Taking these concerns into consideration, we decided not to remove these outliers in favor of dealing with the complicated situation in its real form, while realizing that the consequence under such a limitation could produce at worst some loss in the statistical power of our MANOVA. It is well known that in almost all situations, deviation from multinormality has only negligible effect on the researcher's preset type I error rate.

Another assumption associated with the statistical test of MANOVA is that MANOVA assumes that the covariance matrices of the original variables are the same for all groups. The violation of the covariance homogeneity assumption is usually much more serious than its multinormality counterpart, with the extent of seriousness being dependent upon sample sizes. Generally speaking, the violation of covariance homogeneity assumption is not serious when sample sizes are approximately equal and very serious as the ratio of sample sizes gets large. Under the latter situation, empirical type I error rate tends to be liberal when larger variance is associated with smaller variance and conservative otherwise. The well-known Box's test which uses the determinants of the within covariance matrices as the generalized multivariate variances is probably by-far the most popular method for testing heterogeous covariance matrices. However, it has been pointed out that Box's test is extremely sensitive to multivariate non-normality, and in such cases it is likely to produce significant results even if covariance matrices are indeed equal (Stevens, 1992). In our case, since we have discovered the presence of multivariate non-normality in our data, it should come as no surprise that Box's test would produce a significant result which indicates it failed the test. Indeed, this was what happened to our data (computed Chi-square was 184 with 90 degrees of freedom, p-value less than .0001). It might superficially lead one to question the tenability of the covariance homogeneity assumption. However, a closer examination of the figures in Table 4 shows that the standard deviations for three groups of people are not different by much. Moreover, the natural logarithms of the computed generalized variances for three groups (as obtained in the process of computing Box's M statistic) are: $\ln|S_1| = 10.310$, $\ln|S_2| = 7.352$ and $\ln|S_3| = 9.343$, respectively. We noted that the generalized variances for the three groups are quite different. However, the smaller variance is assciated with the group of elementary school principals, which has the smallest sample size of all three groups (n_2=38). Therefore, the MANOVA test statistic will be conservative, meaning we will be safe with respect to the type I error rate should the MANOVA statistic turn out to be significant.

What Can We Know From MANOVA Results?

We obtained a Wilks' lambda of .84 ($p<.0001$), indicating that about 16% of the variance in the 9-variable system could be attributed to their membership in different groups (this can be translated into an multiple correlation coefficient of .40, a medium-sized effect). Had our research purpose simply been to answer questions such as: Do community residents, school teachers, and school principals differ in their views concerning the first question (V1)? the second question (V2)? ..., etc., then performing nine simple one-way univariate analyses of variance would wrap up our study and our need be sufficed. However, such a study would reveal only a trivial and uninteresting part of the real world phenomenon, since it completely ignores the correlations among the original nine variables.

In addition to the results of omnibus test, more information can be easily obtained in the process of performing MANOVA procedures (with the help of some computer softwares, of course). Here, the two linear composites derived from discriminant analysis had their canonical correlations with the grouping variable as .33 and .24, respectively, with p-values less than .0001. This indicates both underlying dimensions are significant for explaining the effect of the grouping variable. Therefore, understanding, identifying, and naming these two underlying dimensions becomes our next major task.

Identification of the Underlying Dimensions

The underlying dimensions are psychological constructs by themselves. Identifying what these constructs are, in our view, is a state-of-the-art task which has a lot more to do with the researcher's self-discretion and expert judgements than with one's statistical expertise. The approach we took for identifying the two underlying dimensions was by examining the correlations between each original variable and each LDF. For our purpose, we need only to examine the pooled within-group correlations. The structure of correlations between nine dependent variables and the LDFs are conveniently summarized in Table 5, in which the first LDF is labeled as CAN1 and the second LDF CAN2.

Table 5. Pooled Within Canonical Structure

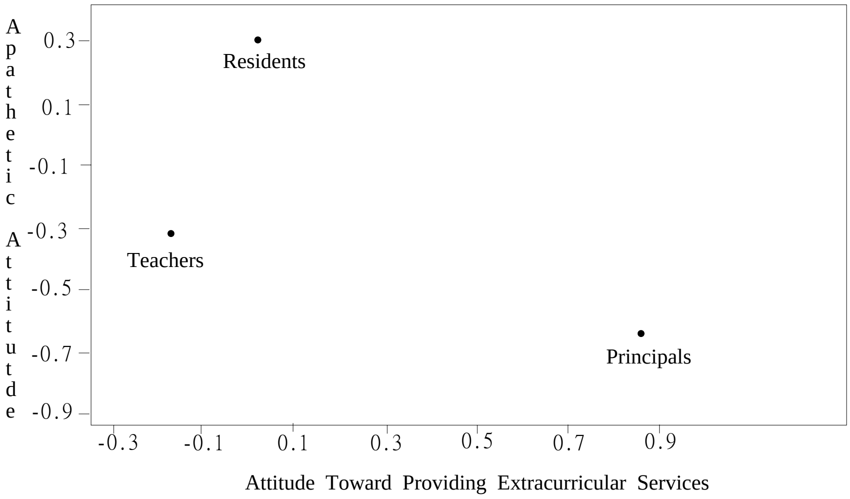
Code	CAN1	CAN2
V1	-0.702981	0.116842
V2	-0.375296	0.024473
V3	-0.250730	0.350703
V4	-0.335991	0.788516
V5	0.040151	0.386318
V6	-0.255543	0.361995
V7	0.252508	0.571419
V8	-0.123305	0.620032
V9	0.170716	0.037847

As seen from Table 5, V1 has largest loading (in absolute term) on the first LDF, followed by V2, V4, V3, V6, and V7. Note that most of the loadings on the first LDF are negative in sign, indicating a reverse relationship exists between these variables and this LDF. Referring back to Table 3 for the descriptions of these variables, V1 conveys the extent of agreement over whether the role of the principal's partcipation in the local community activities can benefit the community development. The first LDF has the highest negative loading on V1, r = -.78, which indicates a higher score with V1 tends to have a lower score on the first LDF. The second highest loading is associated with V2, which concerns the role of the principals as liaisons in the community. It also shows a negative correlation. Similar loadings were found with V3, V4, V6, and V8. It is obvious that the first LDF has a construct which is in opposition to the positive responses of these variables. Putting these negative pieces of information together, it appeared to us that this LDF displays a negative construct which shows a lack of interest in relating the school and local community together. Therefore, we treated this underlying dimension as an apathetic attitude toward the relationship between community and the local school.

The second LDF has strongest correlations with V4, V7, and V8 (.78, .57, and .62, respectively) and moderate correlations with V3, V5, and V6 (.35, .39, and .36 respectively). Note that all signs are positive. We regarded this underlying dimension as the attitude toward schools' providing extra services for the community residents. Now we have successfully reduced the original nine variables to two underlying dimensions which are responsible for the differences among the community residents, principals, and school teachers. Although from a purely mathematical perspective, being uncorrelated does not exactly mean orthogonality, yet under any circumstances, the angle between two LDFs as derived from discriminant analysis will be very close to 90 degrees (the deviation is usually less than one or two degrees). For all practical purposes, we plotted the means of three categories of people in the perpendicular two-dimensional coordinate system as shown in Figure 1.

In Figure 1 we see a distinct pattern of means for three groups in the plot. Principals have a high interest in providing extra services for the local community residents and appeared lowest on the apathetic attitude (a lower apathetic score indicates more positive attitude here). This means that the principals generally hope to play a positive role in the school-community relations and also have high aspirations for such a role. Teachers, on the other hand, showed lowest desire for schools to provide extra services to the local community residents and also showed a somewhat apathetic attitude toward the school-community relationship. Because the Taiwan's elementary school system puts the burden of administrative duties on the shoulders of its teachers, most teachers already have to take care of their seemingly endless

Figure 1. a plot of group means in two underlying dimensions.



assigned chores in addition to their normal regular teaching jobs. It is thus quite understandable why teachers showed such a low desire for schools to provide extra services since providing extra services would mean another increase in their workload. As for community residents, our analysis showed they had the highest apathetic attitude toward the school-community relationship in comparison to the principals and teachers, and that they also had a somewhat lower interest in whether schools should provide extra services for them.

Relative Variable Importance

We also investigated a question which was quite intersting to us. On which of the nine original questions did the three categories of our survey respondents disagree more than other questions? We think that providing a list of order ranking on these questions with respect to their relative importance would give a satisfactory answer to such a question. The relative importance here referred to the extent that the three groups of respondents differ in their opinions on one question relative to the other questions. Here we operationalize the relative variable importance as the relative variable contribution to the effect of the grouping variable.

There are two ways to view this relative variable contribution. The first one is to treat each variable independently without considering its correlations with other variables. Ranking the univariate F values for the grouping variable effect on each original variable would provide us with the answer. The second one is to allow all variables to enter into the picture and assess the contribution of each variable to the grouping variable effect in addition to all other variables. The latter view of the relative variable importance, as we have pointed out in the previous section, is what we believe a better way to understand their relative importance issue because it takes the entire picture into consideration. What's worth mentioning here is that two views may produce quite different lists of variable ranking order, as evidenced in Table 6.

Table 6.Variable importance ranking based on two different approaches

Code	Univariate	Multivariate

	F-Value	Rank	F-to-remove	Rank
V1	19.543	1	14.598	1
V2	5.505	5	2.630	6
V3	4.883	7	0.444	8
V4	16.691	2	6.291	3
V5	3.012	8	0.249	9
V6	5.137	6	2.738	5
V7	8.940	3	9.959	2
V8	8.191	4	1.759	7
V9	1.165	9	3.213	4

The most important variable was found to be with V1 by both univariate and multivariate analyses, which concerns the functional role of the principals. This means that the three categories of people disagree most on whether the functional roles of the principals can promote the development of the community, even after all other variables were taken into account. The second most important variable was found to be with V4 based on the univariate analyses. However, after removing the effect of all other variables, V7 becomes the second most important variable to the grouping variable effect, far ahead of V4. The most noteworthy variable is V9, which was found to be least important by way of univariate analyses, turns out to be the fourth important variable when other variables are involved.

Multivariate Group Contrasts

Three contrasts were made: residents versus teachers, principals versus teachers, and principals versus residents and teachers combined. The results of the contrast analyses were summarized in Table 7.

Table 7. Construct information (Structure Loadings) and significance test information for three Contrast Analyses

	Residents vs Teachers (1 0 -1)	Principals vs Teachers (0 1 -1)	Principals vs Others (-1 2 -1)
V1	-.6185	.3346	.5022
V2	-.3434	.1427	.2372
V3	-.1131	.4123	.4311
V4	-.0408	.8545	.8375
V5	.1720	.3534	.2918
V6	-.1137	.4245	.4431
V7	.4355	.4613	.3197
V8	.1000	.6270	.5769

V9	.1732	-.0185	-.0680
Wilks $\hat{\lambda}_j$	.9030	.9398	.9321
Canonical Correlation	.3114	.2453	.2388
F(9,642)	7.6586	4.5661	5.1970
Pr > F	< .0001	< .0001	< .0001

It can be seen from Table 8 that all three contrasts show statistical significance (with all p-values less than .0001). Our job is to explain what constructs these contrasts represent. For the contrast of residents versus teachers, relative high negative loadings were associated with principals' functional roles in the community (V1 and V2), and positive loadings with teachers' involvement in the community (V5), schools providing services to the community residents (V7), and residents' involvement in the school administration (V8). Residents tend to less concerned about principals' functional roles in the community than teachers, whereas residents' are much more interested in the services which may be provided by the school than their teachers counterpart. For the contrast of principals versus teachers, principals are seen to be much more excited about prividing outside-class services than the teachers themselves (high loadings on V4 and V8). For the contrast of principals versus residents and teachers combined, the structure pattern is similar to that of principals versus teachers, in that principals are generally more positive about all the nine questions than the residents and teachers.

Summary of Results

The community residents, teachers, and principals were found to differ in two underlying dimensions: apathetic attitude toward school-community relationship and attitude toward whether schools should provide extracurricular services for community residents and students. To the suprise of the researchers, as well as the educational administrators, and despite all the efforts to promote the school-community relations in recent years, community residents still seem to have somewhat higher apathetic attitude than school principals and teachers. Based on the results of this study, principals seem to be distinctly different from the rest of the people. These results may be translated to mean that aspirations for changes from principals to promote the school community relations in the past somehow neither elicited enough support from teachers nor gained enough attention from the community residents.

The investigation of the variable importance issue revealed that community residents, principals, and teachers differ more on issues such as the functional role of the principals, sponsoring recreational or cultural events, maintaining contacts with students' families, and whether the community residents should be involved in the management of school affairs. We believe that all these issues are important for the educational administrators to think about in the process of making decisions. After considering these issues, community residents, principals, and teachers appeared homogeneous in their views toward issues such as openning up facilities, promoting government policies, and teachers' participation in community activities.

It is our belief that the answers to our research questions shall provide true understanding of the real multivariate picture for our data situation here. With the understanding of true underlying structure, any single piece of information as seen from a univariate analysis will become a sheer minor aspect of the whole, and will appear self-explanatory under the derived multivariate structure.

Concluding Remarks

Although we have just introduced MANOVA in a one-way layout, two-way or higher dimensions layout can also be accomplished via a one-way design through recoding levels of the grouping variables involved. For instance, a two-way layout (A with two levels, B with three levels), a new grouping variable called GRP can be formed by the following SAS Data Step commands:

```
If A=1 and B=1 then GRP=1;
If A=1 and B=2 then GRP=2;
If A=1 and B=3 then GRP=3;
If A=2 and B=1 then GRP=4;
If A=2 and B=2 then GRP=5;
If A=2 and B=3 then GRP=6;
```

This recoding scheme changes the two-way layout to a one-way layout (with six levels) without affecting the overall two-way test statistic. Moreover, any main effect and interaction effect can be obtained through properly setting up the set(s) of contrast coefficients. For example, the A main effect can be tested by specifying the following contrast: (1 1 1 -1 -1 -1) and the main effect B can be tested by specifying two non-redundant contrasts (in general, the number of non-redundant contrasts required is equal to the ANOVA model degrees of freedom, contrasting different levels of the B factor) in a single run as follows: (1 -1 0 1 -1 0, 1 0 -1 1 0 -1). The major advantage of reparameterizing a two-way model into a one-way model is that researchers can better conceptualize the interaction effect which is of concern to them with a single freedom test. This helps the researcher understand the interaction effect in a down-to-earth fashion, instead of wondering where the individual effect(s) might exist. For instance, if the researcher wonders whether the difference between two levels of A in the first level of B is different from that in the the seond level of B, he may specify the following contrast: (1 -1 0 -1 1 0). The analysis of a such contrast (or sets of contrasts) is much more easier to understand than the overall tests of main or interaction effects in a higher dimension model.

It is disappointing that no single statistical computer program package may be used to readily obtain all the information discussed in the current paper. It is, however, relatively easy for the experienced researchers to obtain the additional information as we discussed by using the information provided on the standard package printouts. It goes well as an old saying puts it, Where there is a will, there is a way; we put, Where there is a will to ask a question in MANOVA, there is a way (or some ways at times) to the solution. It is the hope of the author that a new realm of research questions, as raised in the current paper, will broaden the minds of the researchers in the future employment of MANOVA for their research endeavors.

References

林清達 (民 83). **國民小學與社區互動之相關因素研究**。 花蓮: 真義出版社。

Bentler, P. M., & Huba, G. J. (1980). Symmetric and assymetric rotations in canonical correlation analysis: New methods with drug variable examples. In N. Hirschberg (Ed.), *Multivariate methods in the social sciences*: Applications. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bibb, R. & Roncek, D. W. (1976). Investigating group differences. *Sociological Methods and Research*, 4.

Bird, K. D. (1983). Simultaneous test procedures and the choice of a test statistic in MANOVA. *Psychological Bulletin,* 93, 167-178.

Blalock, H. M. (1961). Evaluating the relative importance of variables. *American Sociological Review,* 26, 866-874.

Borgen, F. H., & Seling, M. J. (1978). The use of discriminant analysis following MANOVA: Multivariate statistics for multivariate purposes. *Journal of Applied Psychology*, 63, 689-697.

Bray, J. H. & Maxwell, S. E. (1982). Analyzing and interpreting significant MANOVAs. *Review of Educational Research.* 52, 340-367.

Cliff, N. (1987). *Analyzing multivariate data*. New York: Harcourt Brace Jovanovich.

Darlinton, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin,* 69, 161-182.

Dillon, W. R. & Goldstein, M. (1984). *Multivariate Analysis: Methods and Applications*. New York: John Wiley & Sons.

Dunteman, G. H. (1984). *Introduction to multivariate analysis*. Beverly Hills, CA:Sage.

Eisenbeis, R. A., Gilbert, G. G., & Avery, R. B. (1973). Investigating the relative importance of individual variables and variable subsets in discriminant analysis. *Communications in Statistics,* 2, 205-219.

Finn, J. D. (1974). *A general model for multivariate analysis*. New York: Holt, Rinehart, & Winston.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics,* 7,179-188.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate Data Analysis with Readings*. New York: MacMillan.

Harris, R. J. (1985). *A primer of multivariate statistics*. Orlando, FL: Academic Press.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statsitics,* 6, 65-70.

Huberty, C. J (1972). Regression analysis and 2-group discriminant analysis. *Journal of Experimental Education*, 41, 39-41.

Huberty, C. J (1975). The stability of three indices of variable contribution in discriminant analysis. *Journal of Experimental Education,* 44, 59-64.

Huberty, C. J, & Smith, J. D. (1982). The study of effects in MANOVA. *Multivariate Behavioral Research.* 17, 417-432.

Huberty, C. J (1984). Issues in the use and interpretation of discriminant analysis. *Psychological Bulletin*. 95, 156-171.

Huberty, C. J (1986). Questions addressed by multivariate analysis of variance and discriminant analysis. *Georgia Educational Researcher*, 5, 47-60.

Huberty, C. J & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin,* 105, 302-308.

Huberty, C. J, & Barton, R. M. (1990). Applied multivariate statistics textbooks. *Applied Psychological Measurement*, 14, 95-101.

Huberty, C. J, & Wisenbaker, J. M. (1992). Variable importance in multivariate group comparisons. *Journal of Educational Statistics,* 17, 75-91.

Huberty, C. J, Chou, T., & Benitez, E. B. (1994). The study of multivariate group contrasts. *Advances in Social Science Methodology*, 3, 123-138.

Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1991). Bias approximations for complex estimators: An application to redundancy analysis. *Educational and Psychological Measurement.*, 51, 1-14.

Lohnes, P. R. (1961). Test space and discriminant space classification models and related significance tests. *Educational and PsychologicalMeasurement*. 21, 559-574.

Lunneborg, C. E. & Abbott, R. D. (1983). *Elementary multivariate analysis for the behavioral sciences.* New York: North-Holland.

Marasciulo, L. A., & Levin, J. R. (1983). *Multivariate statistics in the social sciences*. Monterey, CA: Brooks/Cole.

Morrison, D. F. (1990). *Multivariate Statistical Methods.* New York: McGraw-Hill.

Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of American Statistical Association,* 69, 894-908.

Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin.* 83, 579-586.

Olson, C. L. (1979). Practical considerations in choosing a MANOVA test statistic: A rejoinder to Stevens. *Psychological Bulletin*, 86,1350-1352.

Pedhazur, E. (1982). *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart and Winston.

Rao, C. R. (1973). *Linear statistical inference*. New York: John Wiley & Sons.

Rencher, A. C. (1988). On the use of correlations to interpret canonical functions. *Biometrika,* 75, 363-365.

Roy, J. (1958). Step-down procedure in multivariate analysis. *Annals of Mathematical Statistics,* 29, 1177-1187.

Stevens, J. P. (1972). Four methods of analyzing between variation for the K-Group MANOVA problem. *Multivariate Behavioral Research,* 7, 499-522.

Stevens, J. P. (1979). Comment on Olson: Choosing a test statistic inmultivariate analysis of variance. *Psychological Bulletin.* 86, 355-360.

Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences*. New Jersey: Lawrence Erlbaum Associates, Inc..

Tatsuoka, M. M. (1969). Multivariate analysis. *Review of Educational Research*. 39, 739-743.

Tatsuoka, M. M. (1971). *Multivariate analysis: Techniques for educational and psychological research*. New York: Wiley.

Tatsuoka, M. M. (1973). Multivariate analysis in educational research. In F. N. Kerlinger (Ed.), *Review of Research in Education* (Vol 1). Itasca, Ill: Peacock.

Tatsuoka, M. M. (1988). *Multivariate analysis*. New York: Macmillan.

Thorndike, R. M., & Weiss, D. J. (1973). A study of the stability of canonical correlations and canonical components. *Educational and Psychological Measurement*, 33, 123-134.

Timm, N. H. (1975). *Multivariate Analysis with Applications in Education and Psychology*. Monterey, California: Brookes-Cole.

Urbakh, V. Y. (1971). Linear discriminant analysis: Loss of discriminating power when a variable is omitted. *Biometrics*, 27, 531-534.

Wilkinson, L. (1975). Response variable hypotheses in the multivariate analysis of variance. *Psychological Bulletin*, 82, 408-412.

Zwick, R. (1985). Nonparametric one-way multivariate analysis of variance: A computational approach based on the Pillai-Bartlett trace. *Psychological Bulletin*, 97, 148-152.

Some New Ways of Analyzing MANOVA Data in a Post Hoc Context

周東山

摘 要

多變項變異數分析在教育，心理，行為，或社會科學方面的研究上，極為廣汎地被研究者採用來做計量效果差異分析的依據，然而，在所有介紹多變項變異數分析法的國外教科書中，大多僅就總括性的考驗(Omnibus test)做介紹，僅有少數提及在總括性考驗測得顯著差異後，進行事後追蹤比較的方法。又在這些介紹多變項事後追蹤比較方法的書籍中，幾乎千篇一律地集中在多重比較考驗的統計值本身，專注在控制第一類型錯誤的情境下，是否能夠正確地拒絕多重比較的虛無假設。至於在拒絕虛無假設後，到底吾人應該怎樣解釋平均數間的差異狀況呢？在回答這樣的問題時，時下通行的事後追蹤方法多為使用單變項變異數分析(ANOVA)針對各別依變項加以檢驗，即或使用Roy-Bose(1953)所提出的多變項同時推論(simultaneous inference)手續來計算多重比較之統計考驗值，下結論時仍是就各依變項逐一解釋，完全喪失多變項變異數分析要將組間之差異情形在多維依變項空間中做一個系統化解釋的原意。

本篇論文先就時下盛行的後續追蹤方式加以批判，而後提出一套能夠將組間差異狀況在真正多變項分析(truely multivariate)的架構下具體化地解釋出來的方法，使用本篇論文所介紹的方法將可具體化地回答下列三個問題:

(1)組間差異狀況是否可用一個或數個潛在特質(underlying construct)來加以解釋？如果可以，那麼這樣的潛在特質所代表的意義又是什麼？

(2)個別依變項在全體其他依變項存在的狀況下，它們的相對重要性如何？

(3)吾人是否可將多變項多重比較的效果，就前兩個問題加以解釋？

以上的問題，都不是時下所採用的後續追蹤法所能顯示的。最後，本篇論文以一實際研究之資料情境，説明如何在多變項變異數分析的架構下進行資料分析與詮釋。