Performing Multivariate Group Comparisons Following a Statistically Significant MANOVA

Craig K. Enders

This article illustrates 2 follow-up procedures that can be used to examine multivariate analysis of variance (MANOVA) group differences: the univariate analysis of a linear composite variable and multivariate contrasts. A heuristic data set is used to demonstrate the procedures, and it is shown that the follow-up methods will not always yield identical substantive interpretations.

More than two decades ago, Huberty and Smith (1982) encouraged researchers to "think multivariately" (p. 429) when conducting follow-up analyses of a statistically significant multivariate analysis of variance (MANOVA), and they outlined a strategy for conducting multivariate group contrasts in conjunction with a descriptive analysis of the linear discriminant function (LDF). Some time later, Huberty and Morris (1989) conducted a content analysis of five prominent behavioral science journals published by the American Psychological Association and found multiple univariate analyses of variance (ANOVAs) to be the predominant MANOVA follow-up method; 88 of 91 (96.7%) studies used a univariate follow-up strategy, and only 4 studies (4.4%) addressed variable importance using the LDF. Clearly, applied practice was at odds with recommendations given by Huberty and Smith.

One early (and still lingering) misconception regarding the use of univariate follow-ups is that MANOVA acts as a "gatekeeper," protecting against Type I error inflation in subsequent univariate analyses—the so-called protected *F* procedure. However, Maxwell (1992) noted that the logic of the protected *F* procedure is faulty, because honest Type I error rates are maintained under a very limited set of conditions. As noted by Maxwell, the MANOVA analysis only protects against Type I error rate inflation in a set of univariate follow-ups when (a) the MANOVA null hypothesis is completely true (in which case the follow-ups are performed only 5% of the time), (b) the MANOVA null hypothesis is entirely false (in which case there is no possibility of a Type I error), and (c) when the MANOVA null hypothesis is false for all but one outcome variable (because it is possible to make a Type I error for only a single variable, the univariate Type I error rate is maintained at α). Clearly, if the goal is to protect against Type I error inflation, the most straightforward approach is to apply a Bonferroni adjustment and forgo the MANOVA analysis entirely. Huberty and Petoskey (2000) forcefully made this point, stating that "a MANOVA/MANCOVA is *not* [italics in original] a necessary preanalysis!" (p. 205).

More fundamental than the issue of Type I error control is the fact that univariate and multivariate procedures simply address different research questions, and should be viewed as incompatible on this basis alone. Huberty and Morris (1989) provided an excellent discussion of this topic and delineated situations that require "univariate questions" ver-

Craig K. Enders, School of Education, University of Miami. Correspondence concerning this article should be addressed to Craig K. Enders, University of Miami, School of Education, PO Box 248065, Coral Gables, FL 33124-2040 (e-mail: cenders@miami.edu).

sus "multivariate questions." They argued that univariate analyses are warranted if (a) outcome variables are "conceptually independent," (b) the research is exploratory in nature, (c) previous studies have used univariate analyses, and (d) the goal is to establish equivalence among comparison groups in a nonexperimental design. In contrast, their primary criterion for performing a multivariate analysis is met when the outcome variables constitute a "system"—"a collection of conceptually interrelated variables that, at least potentially, determines one or more meaningful underlying variates or constructs" (Huberty & Morris, 1989, p. 304).

Although analytic choices are clearly dependent on one's substantive goals, there are statistical considerations as well. Tabachnick and Fidell (2001, p. 357) explained that MANOVA is most desirable if correlations among outcome variables are strong and negative or moderate in either direction (e.g., |.60|); less desirable is the situation in which correlations are strong and positive or near zero. When strong positive correlations result from using indicators of the same construct (e.g., a latent variable system; Bollen & Lennox, 1991), it may be preferable to perform a multivariate comparison of latent means using structural equation modeling (Cole, Maxwell, Arvey, & Salas, 1993).

Although the practice of following a significant MANOVA with a series of univariate ANOVAs was based on early recommendations from Cramer and Bock (1966) and became "standard operating procedure as of the 1980s" (Maxwell, 1992, p. 138), recent methodological literature appears to be unequivocally opposed to this practice. For example, in an invited address to the American Educational Research Association, Thompson (1999) wrote "when you do a multivariate analysis, you must *not* use a univariate method *post hoc* [italics in the original] to explore the multivariate effects " (p. 18). Similarly, Huberty and Petoskey (2000) stated that "researchers should conduct a multivariate analysis only when they are interested in *multivariate* [italics in the original] outcomes" (p. 205).

Despite the clear message from the methodological literature, multiple univariate tests are still the predominant method for examining group differences following a significant MANOVA. Keselman and colleagues (1998) published a content analysis of 17 educational and behavioral science journals published in 1994 and 1995 and found that approximately 84% of 208 MANOVA analyses were followed by univariate post hoc procedures. Similarly, Kieffer, Reese, and Thompson's (2001) review of 756 articles in the *American Educational Research Journal* and *Journal of Counseling Psychology* revealed that more than 70% of the multivariate analyses were followed by univariate ANOVAs.

As pointed out by Keselman and colleagues (1998), the continued use of multivariate and univariate procedures in tandem with one another is not surprising given that popular multivariate textbooks advocate this strategy. The presentation of univariate follow-ups in textbooks is problematic from a pedagogical perspective, because those who are new to the area of multivariate statistics may develop bad analytic habits long before they gain the expertise necessary to consult more thorough, yet perhaps less accessible, texts (e.g., Huberty, 1994, p. 196). In addition, the content analyses of Huberty and Morris (1989) and, more recently, Keselman and colleagues (1998) and Kieffer and colleagues (2001) have suggested that even seasoned researchers frequently conduct multivariate analyses that are inconsistent with what might be viewed as statistical "best practice."

PURPOSE

If one accepts that univariate post hoc analyses are inappropriate following a statistically significant MANOVA, the obvious question is "What *is* the correct procedure for conducting a multivariate examination of group differences?" In addressing this question, the purpose of the present article is threefold. First, I outline two existing analytic procedures (the analysis of linear composite variable scores and multivariate contrasts) that can be used to investigate group differences following a statistically significant MANOVA. I argue that the former method is well suited for situations in which a multivariate omnibus test is of interest, whereas the latter approach is most appropriate when a priori planned contrasts are specified. Second, detailed instructions are given for carrying out these two procedures in both single-factor and factorial designs. Whereas the former situation is addressed in other sources (e.g., Huberty & Smith, 1982), less has been written on follow-up procedures for factorial MANOVA designs. Finally, the analytic steps are modeled using a heuristic data set.

Unlike univariate ANOVA, there are two lines of inquiry relevant to MANOVA analyses: the examination of group differences and a description of the relative importance of each outcome variable in differentiating the *k* groups. Although the latter is obviously unique to the multivariate case, the former appears to represent a significant methodological gap in many multivariate textbooks and applied practice in general. Although the primary goal of this article is to outline group comparative procedures, it is impossible to avoid a concurrent examination of the LDF as a vital adjunct for understanding group differences. Thus, descriptive discriminant analysis (DDA) results are presented with each heuristic analysis. Readers who seek more detailed discussions of DDA are encouraged to consult works by Huberty and his colleagues (e.g., Huberty, 1984, 1994; Huberty & Morris, 1989; Huberty & Petoskey, 2000; Huberty & Smith, 1982).

HEURISTIC DATA SET

To illustrate the use of the two follow-up procedures, a subset of data was drawn from the National Household Survey on Drug Abuse, 1998 (U.S. Department of Health and Human Services, 2000). The subset consisted of 814 respondents who reported using marijuana within 30 days prior to completion of the survey. The mean age of these respondents was $M = 25.7$ years ($SD = 8.46$), and approximately 60% were male. Because of the pedagogical nature of this article, no sample weights were used in the analyses.

The ages at which respondents began using cigarettes, alcohol, and marijuana (*CIGAGE*, *ALCAGE*, and *MJAGE*, respectively) served as outcome variables in the heuristic analyses. These variables are collectively referred to as the onset of substance use. Note that this set of outcome variables constitutes a variable system as defined by Huberty and Morris (1989).

The following research question was posed for the purpose of illustrating a single-factor MANOVA analysis: Do groups defined by four levels of educational attainment (*EDUC;* less than high school, high school, some college, and college graduate) differ with respect to the onset of substance use, as defined by the set of three outcome variables? The means and standard deviations of the outcome variables are given by *EDUC* in Table 1, as is the pooled within-group correlation matrix.

The factorial MANOVA analysis was illustrated by adding a second classification variable that denoted whether or not a respondent had previously been arrested *(ARRESTED)*, resulting in a 2×4 MANOVA. Descriptive statistics for each design cell are presented in Table 2. The pooled within-group correlation matrix for this analysis was quite similar to that shown in Table 1.

OVERVIEW OF FOLLOW-UP METHODS

Univariate Analysis of a Linear Composite Variable

As noted by Bray and Maxwell (1985), a MANOVA significance test with p outcome variables can be alternatively conceptualized as a univariate test of mean differences on a new variate, V, which is a linear combination of the original p outcome variables (henceforth referred to simply as the composite variable). When $p = 3$, as is the case in the heuristic data, V is computed as

TABLE 1

Outcome Variable Descriptive Statistics by Education Level

Variable	Less Than High School (n = 238)	High School (n = 304)	Some College (n = 204)	College Graduate (n = 68)
CIGAGE				
M	15.59	16.73	17.26	18.93
SD	2.81	3.05	3.15	4.55
ALCAGE				
M	14.62	15.06	14.61	14.68
SD	2.89	2.77	2.78	2.70
MJAGE				
M				
SD	14.97	15.60	15.98	17.57
	2.59	3.02	3.14	5.58
	Pooled correlation matrix			
	1	**2**	**3**	
1. CIGAGE	3.16	—		
2. ALCAGE	0.25	2.80	—	
3. MJAGE	0.29	0.52	3.23	

Note. CIGAGE = age that respondents first used cigarettes; ALCAGE = age that respondents first used alcohol; MJAGE = age that respondents first used marijuana. The square root of the mean square within (MSw) values (i.e., the pooled standard deviations) are listed on the diagonal of the pooled correlation matrix.

$$V = b_1Z_1 + b_2Z_2 + b_3Z_3 \quad (1)$$

where the *Z*s are the outcome variables expressed in standardized form (the standardization method is important and is discussed below), and the corresponding *b*s are standardized discriminant function weights that reflect the contribution of each *Z* to the composite after covariation with the remaining *Z*s is partialled out. The subscripts attached to the *b*s in Equation 1 denote the weight associated with each outcome variable. Note that the resulting composite is expressed on a standardized metric with a mean of zero and unit standard deviation for the entire sample. In the context of the current analysis, the composite would be computed as follows:

$$ONSET = b_1CIGAGEz + b_2ALCAGEz + b_3MJAGEz \quad (2)$$

The weights in Equation 1 are derived such that the ratio SS_B/SS_W is maximized on the composite, V. That is, no other set of weights would yield a composite variable that provides greater discrimination among the k groups being compared. When $k > 2$, additional linear combinations can be derived using the formula in Equation 1, with the constraint that the V_cs are orthogonal. In general, the number of composites that can be formed is equal to the lesser of $k - 1$ and p.

Before going further, it is important to discuss the standardization of the outcome variables in the equations above. Although computer packages routinely create and save composite variable scores for single-factor designs (e.g., SPSS DISCRIMINANT, SAS PROC DISCRIM), these scores can be generated for factorial designs by translating Equation 1 into the appropriate command syntax (e.g., using the COMPUTE statement in SPSS).

TABLE 2

Means and Standard Deviations of Dependent Variables by Education Level and Arrest Record

	CIGAGE		ALCAGE		MJAGE		
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*n*
No arrests							
Less than high school	15.84	2.94	15.33	2.84	15.59	2.39	114
High school	16.51	2.73	15.20	2.84	15.65	2.93	204
Some college	17.61	3.07	15.03	2.84	16.32	3.30	148
College graduate	19.86	4.59	14.80	2.98	17.89	5.69	44
Previously arrested							
Less than high school	15.35	2.67	13.97	2.80	14.40	2.64	124
High school	17.16	3.60	14.76	2.62	15.51	3.22	100
Some college	16.36	3.21	13.52	2.30	15.09	2.47	56
College graduate	17.21	4.02	14.46	2.13	17.00	5.45	24

Note. CIGAGE = age that respondents first used cigarettes; ALCAGE = age that respondents first used alcohol; MJAGE = age that respondents first used marijuana.

However, it is important to note that the appropriate standardization of the outcome variables is dependent on the software package that generated the standardized discriminant function weights. If using SPSS MANOVA, SPSS DISCRIMINANT, and SAS PROC DISCRIM, Equation 1 should be applied using outcome variables that have been standardized using the pooled standard deviation of the *k* groups (i.e., the positive square root of the mean square within [MS_W] for each outcome variable). However, the standardized canonical (i.e., discriminant function) coefficients produced by PROC GLM in SAS require a different standardizer: the positive square root of the mean square total (MS_T) for each outcome variable (i.e., the sample standard deviation). In either case, these quantities can be obtained from the univariate ANOVA results that are produced by most multivariate software procedures. For single-factor analysis, these values are given on the diagonal of the within-group correlation matrix presented in Table 1.

Given that a MANOVA can be recast as a univariate ANOVA of a linear composite of the original *p* outcome variables, it is my contention that this analytic strategy provides applied researchers with a general, as well as a familiar, method for conducting group comparison following a statistically significant MANOVA. As is demonstrated later in this article, the method can be applied to a variety of MANOVA contexts, including factorial designs. Furthermore, this analytic strategy may be particularly useful when researchers are interested in an omnibus MANOVA effect as opposed to a priori planned contrasts, the situation that seems to dominate analytic practice.

It should be noted that this follow-up method has received little attention in multivariate textbooks and the literature in general. For example, Tabachnick and Fidell (1996, p. 405) mentioned in a footnote that post hoc comparisons can be performed using the composite variable, but gave no explanation of the procedure. Maxwell (1992) outlined the computation of linear composites to test hypotheses about group differences, but the procedure was somewhat different than that outlined here (see the discussion of moderately restricted contrasts below). Finally, Sheehan-Holt (1998) proposed a "partially restricted contrast" procedure (p. 875) that is equivalent to the one I outline.

Although somewhat tangential to the goals of this article, a brief discussion of simultaneous confidence interval (i.e., simultaneous test) procedures is warranted at this point. Roy and Bose (1953) proposed a follow-up method that is the multivariate analog to Scheffé (1953) follow-ups in univariate ANOVA. After rejecting the MANOVA omnibus test, the Roy-Bose procedure allows one to conduct all possible pairwise and complex comparisons using individual dependent variables, as well as linear combinations of dependent variables, while maintaining a specified Type I error rate. Roy and Bose developed the simultaneous confidence interval procedure using Roy's greatest root, but the procedure was later extended for use with other multivariate test statistics by Gabriel (1968). Given the enormous number of contrasts that are allowed, it is not surprising that the procedure has been criticized for being overly conservative (e.g., Stevens, 1996, p. 199), particularly when applied to univariate contrasts (e.g., Barcikowski & Elliott, 1991; Bird & Hadzi-Pavlovic, 1983).

Variations on the simultaneous testing procedure have been proposed that are similar to the composite variable analysis outlined in this section. For example, Bird and Hadzi-Pavlovic (1983) described *moderately restricted* contrasts in which groups are compared on a composite variable computed using standardized weights of 0, 1, or –1 (see also Barcikowski & Elliott, 1991; Maxwell, 1992). In the current context, such a composite might be computed as

$$ONSET = 1(CIGAGEz) - 1(ALCAGEz) + 1(MJAGEz). \quad (3)$$

Although the intent of this procedure is to simplify the interpretation of the composite variable, Sheehan-Holt (1998) reported that moderately restricted contrasts lack power, and instead suggested *partially restricted* contrasts. Using partially restricted contrasts, a simultaneous test procedure is used to examine all pairwise and complex comparisons using composite variable scores computed from Equation 1, a procedure that is equivalent to the one I outline in this section. Sheehan-Holt's preliminary simulation results indicated that partially restricted contrasts have adequate power in most situations—particularly when based on Roy's greatest root—and are fairly robust to assumption violations.

Multivariate Group Contrasts

Huberty and Smith (1982) proposed a MANOVA follow-up strategy that involves a series of two-group multivariate contrasts (either simple or complex). Of the two follow-up procedures outlined herein, this approach has been given the most attention in multivariate texts (e.g., Huberty, 1994, p. 196; Tabachnick & Fidell, 2001, p. 352). In the context of the current demonstration, reconsider the comparison of four educational attainment groups (less than high school = 1, high school = 2, some college = 3, college graduate = 4) on the onset of substance use. It may be of interest to compare the mean vectors of high school and college graduates. This contrast would be

$$\psi = \mu_2 - \mu_4,$$

where m_2 and m_4 are the mean vectors of the p outcome variables for high school and college graduates, respectively. Similarly, a complex contrast comparing the college graduates with the combined mean vectors of the remaining three groups would be

$$\psi = 3\mu_4 - \mu_1 - \mu_2 - \mu_3.$$

As pointed out by Huberty (1994, p. 197), the Hotelling T^2 statistic (or equivalently Wilks's lambda) can be used to test such contrasts for statistical significance using the error covariance matrix from the omnibus MANOVA under the assumption of homogeneity of covariance matrices. Consistent with their univariate counterparts, multivariate contrasts may be performed in lieu of the omnibus test if they constitute a priori planned comparisons.

It is important to note that the composite variable getting tested by each multivariate contrast is not necessarily the same as (or even similar to) the linear combination of p outcome variables associated with the MANOVA omnibus test. When $k > 2$, a unique set of structure coefficients and standardized discriminant function weights are derived for each of the c contrasts. This implies that a series of multivariate contrasts can potentially yield significance tests that are inconsistent with the omnibus MANOVA results. For example, a multivariate contrast could produce a statistically significant difference between two groups that had nearly identical centroids (i.e., composite variable means) in the corresponding omnibus analysis. Again, this is because the two analyses maximize group differences using a different linear combination of the p outcome variables. Such inconsistencies are not problematic if one's analytic focus is on a set of a priori contrasts, because the omnibus test would likely be of no interest. However, it is for this reason that I later argue against the use of multivariate contrasts when the MANOVA omnibus test *is* of interest—in that situation it may make more sense to investigate group differences using the outcome variable weights that generated the significant omnibus effect (i.e., analyze the composite variable scores from the omnibus analysis).

Finally, multivariate contrasts can be applied to more complex designs such as factorial MANOVA (Huberty, Chou, & Benitez, 1994, p. 130). In the context of a factorial design, contrasts could be specified that are akin to multivariate simple main effect tests. This can easily be accomplished using common statistical software packages (e.g., the SPSS MANOVA procedure) and is demonstrated later in the article.

ILLUSTRATIVE ANALYSES

Having outlined two methods for examining group differences following a statistically significant MANOVA, I now present a series of heuristic analyses. The SPSS syntax for these analyses is given in the Appendix, and the raw data are available on request. Follow-ups that use the composite variable scores computed from the first LDF (i.e., the largest root) are probably most appropriate when Roy's greatest root is used to test the omnibus effect (e.g., Sheehan-Holt, 1998). However, I use Wilks's lambda for the purposes of illustrating tests of dimensionality in DDA.

Single-Factor MANOVA Design

To illustrate the two follow-up procedures in the context of a one-factor MANOVA, the substance use onset variables were compared across the four educational attainment groups (less than high school, high school, some college, and college graduate).

Omnibus analysis. Prior to examining the multivariate omnibus test, the homogeneity of covariance matrices assumption was inspected following procedures outlined by Huberty and Petoskey (2000). Although the Box F test was statistically significant at $p < .005$ (the alpha level suggested by Huberty and Petoskey for this test), the natural logarithms of the $k + 1$ covariance matrices were "in the same ballpark" (p. 193); the four group log determinants ranged between 5.61 and 7.99, and the log determinant of the pooled covariance matrix was 6.29. Thus the analysis proceeded as if the equal-covariance-matrix condition was met on the basis of the relative similarity of the log determinants.

With four groups and $p = 3$ outcome variables, there are a total of three (i.e., the minimum of p and $k - 1$) dimensions, or composite variables, that can differentiate the four groups. The simultaneous test of all three dimensions was statistically significant, $\Lambda = .88$, $\chi^2(9, N = 814) = 105.78$, $p < .01$. The remaining two tests, which examine the combination of the second and third dimension and the third dimension in isolation, respectively, were not significant: $\Lambda = .99$, $\chi^2(4, N = 814) = 6.96$, $p = .14$ and $\Lambda = .99$, $\chi^2(1, N = 814) = .68$, $p = .41$.

To further understand the nature of this multivariate effect, a DDA was performed. The structure coefficients and standardized discriminant function weights from this analysis are given in Table 3. The four group centroids from the first composite variable are shown in the LDF plot in Figure 1. As seen in the Figure, the centroids increase in value as the

TABLE 3

Significant Test Results and LDF Weights for the One-Factor MANOVA Illustration

Effect	Structure *r*	Std. Weight	Wilks's Λ	*F*	*p*
Omnibus					
CIGAGE	0.80	0.75	0.82	12.34	< .001
ALCAGE	–0.02	–0.55			
MJAGE	0.59	0.66			
Less than high school vs. high school					
CIGAGE	–0.96	–0.28	0.98	6.19	< .001
ALCAGE	–0.41	–0.03			
MJAGE	–0.52	–0.07			
High school vs. some college					
CIGAGE	0.51	0.56	0.98	4.47	0.004
ALCAGE	–0.48	–0.98			
MJAGE	0.35	0.70			
Some college vs. college graduate					
CIGAGE	0.74	0.65	0.97	8.67	< .001
ALCAGE	0.03	–0.54			
MJAGE	0.69	0.79			

Note. LDF = linear discriminant function; CIGAGE = age that respondents first used cigarettes; ALCAGE = age that respondents first used alcohol; MJAGE = age that respondents first used marijuana. The *df* values for the omnibus MANOVA were (9, 1967) and (3, 808) for all contrasts.

level of education increases; the centroids for the first three groups are spaced in roughly equidistant intervals, whereas the college graduate centroid is considerably higher.

Because the centroids are means on a variate that is a linear combination of the three outcome measures, it is necessary to interpret the LDF—for example, what is the interpretation of a "high" mean score on the composite variable? To facilitate this interpretation, both the structure coefficients (i.e., bivariate correlations between each outcome and the composite) and standardized weights (i.e., the partialled relationship between each out-

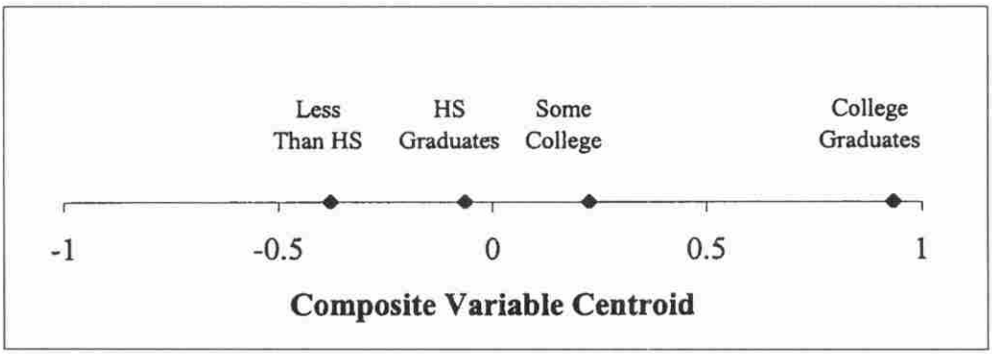

FIGURE 1

LDF Plot of Composite Variable Score Centroids for Single-Factor MANOVA Analysis

Note. HS = high school; LDF = linear discriminant function.

come and the composite) can be examined. As seen in Table 3, the structure *r*s for *CIGAGE* and *MJAGE* are strong and positive (.80 and .59, respectively), whereas the coefficient for *ALCAGE* is virtually zero (–.02). This suggests that higher centroid values (see Figure 1) are associated with later onset of smoking (both cigarettes and marijuana). In a bivariate sense, the onset of alcohol use is unrelated to the composite variable that produced the significant MANOVA effect. Thus, the LDF plot in Figure 1 suggests that increasing levels of educational attainment are associated with a delay in the onset of smoking behavior, although this relationship obviously cannot be interpreted in a causal sense.

An examination of the standardized weight for *ALCAGE* ($b = -.55$) yields a slightly different interpretation, however. After partialling out the effects of *CIGAGE* and *MJAGE*, *ALCAGE* makes a negative contribution to the LDF, the magnitude of which is similar to that of *MJAGE* and *ALCAGE*. Because *ALCAGE* had a near-zero bivariate correlation with the composite but had a large standardized weight, this suggests the presence of a suppressor variable effect (e.g., Courville & Thompson, 2001; Horst, 1966, p. 355).

More central to the purpose of this discussion is the examination of group contrasts. The significant omnibus effect suggests that the four educational attainment groups can be differentiated using a linear combination of outcome variables, but follow-up tests are needed to further delineate the group differences. Although it is unnecessary—and probably inadvisable—to follow the omnibus test with group contrasts *and* an analysis of the composite variable, I present both approaches in an attempt to demonstrate and highlight differences between the two.

Composite variable analysis. Although statistical software packages routinely save composite variable scores for single-factor designs (e.g., the SPSS DISCRIMINANT procedure), it is useful to illustrate the process, because the same procedure can be applied to factorial designs. Prior to computing the composite, the outcome variables must be standardized using the pooled standard deviation (i.e., the positive square root of MS_W for each outcome variable). For example, *CIGAGE* would be standardized using the grand mean and pooled standard deviation values as follows:

$$CIGAGEz = (CIGAGE - \bar{X}_{CIGAGE}) / (MSW_{\text{CIGAGE}})^{.5}. \quad (4)$$

Using the standardized outcome variables and the standardized discriminant function weights obtained from the DDA, the composite variable scores are computed as follows:

$$ONSET = .752(CIGAGEz) - .547(ALCAGEz) + .656(MJAGEz). \quad (5)$$

The resulting composite variable, *ONSET*, was next submitted to a single-factor univariate ANOVA for further analysis. The univariate F test for *ONSET* was obviously significant, but it is the post hoc group comparisons that are of interest in this analysis. However, it is important to point out that the univariate analysis of the composite variable coincides with the omnibus MANOVA analysis. Specifically, the ratio of the SS_B/SS_T (i.e., η^2) is identical to Roy's greatest root (θ) from the MANOVA omnibus test. This is not surprising, because θ is simply the first eigenvalue (i.e., SS_B / SS_T ratio) of the $\mathbf{BT}^{-1}$ matrix from the multivariate analysis. Note that SAS expresses the Roy criterion (θ) as an eigenvalue of $\mathbf{BW}^{-1}$, as does the SPSS GLM procedure; the SPSS MANOVA procedure expresses θ as an eigenvalue of $\mathbf{BT}^{-1}$. Thus, in cases where θ is expressed as $\mathbf{BW}^{-1}$ (i.e., SAS and SPSS GLM), it is the SS_B / SS_W ratio from the univariate analysis of the composite that will coincide with the value of θ from the multivariate omnibus test. Furthermore, the η^2 values ($SS_B / [SS_B + SS_W]$) for each of the (min[p, $k - 1$]) composites are equivalent to the squared canonical correlations obtained from a discriminant analysis. The sum of the η^2 values (or squared canonical correlations) is equal to the Pillai trace. These relationships can be used to verify that the composite outcome variable has been computed correctly. Furthermore, the univariate

F test for *ONSET*, $F(3, 801) = 35.06$, $p < .01$, is identical to the *F* test for Roy's greatest root provided by multivariate software packages (e.g., SPSS GLM, SAS PROC GLM, SAS PROC DISCIM).

For the purposes of this demonstration, pairwise comparisons were conducted between adjacent levels of the grouping variable (e.g., high school graduates versus some college), and a Bonferroni adjustment was used to protect against Type I error inflation (i.e., $\alpha = .05 / 3 = .017$). All three comparisons were statistically significant, suggesting that each increment in educational attainment is associated with a statistically significant delay in the onset of smoking behavior. I made this interpretation of the LDF (see Figure 1) on the basis of the structure coefficients.

Multivariate contrasts. Following the method originally outlined by Huberty and Smith (1982), pairwise multivariate contrasts were conducted to examine differences between adjacent levels of the educational attainment grouping variables (i.e., the same comparisons conducted previously). The SPSS syntax for these contrasts is given in the Appendix. Again, these contrasts could be conducted in lieu of the omnibus test if they constituted a priori hypotheses, but are presented here as post hoc tests for demonstration purposes. Consistent with the univariate follow-ups of the composite variable scores, all three contrasts were statistically significant after implementing a Bonferroni adjustment. Table 3 gives Wilks's lambda, *F*, and *p* values for the constrasts as well as the structure coefficients and standardized weights for each.

At first glance, it may appear that the contrast results were consistent with those from the composite variable analysis (i.e., all contrasts were statistically significant), but the follow-up procedures have actually diverged in an important way. Table 3 shows the structure coefficients and standardized weights for three multivariate contrasts, as well as those from the omnibus MANOVA effect (i.e., the weights used to compute the composite variable scores in the preceding section). As seen in the table, the coefficients and weights associated with the omnibus test are quite similar to those obtained from the comparison of college graduates and individuals who attended but did not finish college. However, the weights are dramatically different for the remaining two contrasts. For example, the strong negative structure coefficients generated from two contrasts suggest that *ALCAGE* exhibits a bivariate relationship with the composite variable—the omnibus test produced a structure coefficient close to zero for this outcome variable. Thus, although both follow-up procedures produced the same statistical conclusions up to this point (i.e., all group comparisons were statistically significant), the substantive interpretation of these results is quite different from an LDF perspective. It is for this reason I later argue that the choice of follow-up method is dependent on whether contrasts constitute a priori hypotheses versus post hoc exploration, and DDA is a vital adjunct to a significant MANOVA.

Factorial Design

To further illustrate the use of these two follow-up procedures in more complex factorial MANOVA designs, a second classification variable, *ARRESTED,* was added to the analysis, resulting in a 2×4 MANOVA design. This additional variable was a dichotomous indicator of whether or not an individual had been arrested prior to the interview.

Omnibus analysis. Consistent with the previous analysis, the equal-covariance-matrix assumption was investigated prior to examining the MANOVA results. The Box *F* test was statistically significant at $p < .005$, but the natural logarithms of the covariance matrices were somewhat similar; the log determinants of the eight design cells ranged between 5.43 and 8.17, and the log determinant of the pooled covariance matrix was 6.25. As such, the following analyses were performed as though the assumption were met.

Results indicated the presence of a statistically significant multivariate interaction effect, $\Lambda = .97$, $F(9, 1957) = 2.72$, $p < .01$. Note that the multivariate effect size, $\eta^2 = .03$, is much smaller than that observed in the single-factor design. In a factorial MANOVA, the outcome variables are recombined for each design effect, and both the main effect for *ARRESTED* and *EDUC* were statistically significant at $p < .01$ ($\Lambda = .97$ and .90, respectively). In this case, only the interaction effect is interpreted, but if it was of interest to examine the main effects, the same procedures outlined previously would be followed. Note that the LDF's for the main effects could be quite different from that of the interaction effect.

To further understand the multivariate interaction effect, a DDA was performed using the first of three possible interaction composite variables. The group centroids (i.e., the composite variable means) were obtained for each design cell and are shown in Figure 2. For ease of illustration, a separate LDF plot is presented for each level of the *ARRESTED* classification variable.

From the centroids in Figure 2, the nature of the interaction can be seen. For individuals with no arrest record, group centroids increased in value as the level of education increased, and college graduates clearly had the highest composite variable mean. In contrast, the four educational attainment groups were not as disparate within the sample of individuals with an arrest record.

Again, it is necessary to provide an interpretation of these centroids using the structure coefficients and standardized weights. On the basis of the structure coefficients it was determined that two variables, *CIGAGE* and *MJAGE,* were salient in defining the interaction composite: structure *r*s were .99 and .35, respectively, whereas the structure coefficient for *ALCAGE* was .22. This interpretation was based on a|.30| cutoff frequently suggested in the context of factor analysis. Although |.30| is a frequently cited rule of thumb, higher values are certainly appropriate in many situations. It is not recommended that this cutoff be blindly applied in every analytic context. Thus, consistent with the single-factor design results, the composite can be characterized as the onset of smoking behavior and is most strongly aligned with the age one begins using cigarettes. Returning to the centroid plots in Figure 2, the onset of smoking behavior generally occurs later in groups with higher educational attainment (i.e., groups with "high" centroid values), and it appears that college graduates with no arrest record begin smoking much later than the remaining groups.

An examination of the standardized weights provides a similar interpretation. Weights for *CIGAGE* and *MJAGE* were .98 and .12, respectively, whereas the weight for *ALCAGE* was −.08. Clearly, the most salient outcome was *CIGAGE*, and the weight for *MJAGE* was also positive, albeit weak. The weight for *ALCAGE* was negative and also quite weak.

Although these interpretations are purely descriptive in nature, a series of significance tests are now illustrated to demonstrate how one might follow up the multivariate effect. Again, both a composite variable analysis and multivariate contrasts were performed. The following analyses are akin to simple comparisons in which differences among arrest groups were examined within each level of *EDUC*. Alternatively, simple effects tests could have been performed to examine differences among the four *EDUC* groups within each of the two arrest conditions. Although simple effect tests are ubiquitous following interaction effects, they may not always be an appropriate follow-up strategy. Levin and Marascuilo (1972) discussed alternative strategies (e.g., interaction contrasts), and these methods could readily be applied to the multivariate case using either of the two follow-up strategies outlined here.

Composite variable analysis. Although three composite variables were associated with the interaction effect in this example, only the first dimension is examined here. Using the standardized weights associated with the first composite, the interaction composite scores were computed as follows:

$$INTCOMP = .980(CIGAGEfz) - .080(ALCAGEfz) + .122(MJAGEfz). \quad (6)$$

Again, the three outcome variables were standardized using the square root of MS_W for each outcome variable prior to computing the composite. Note also that the variable names

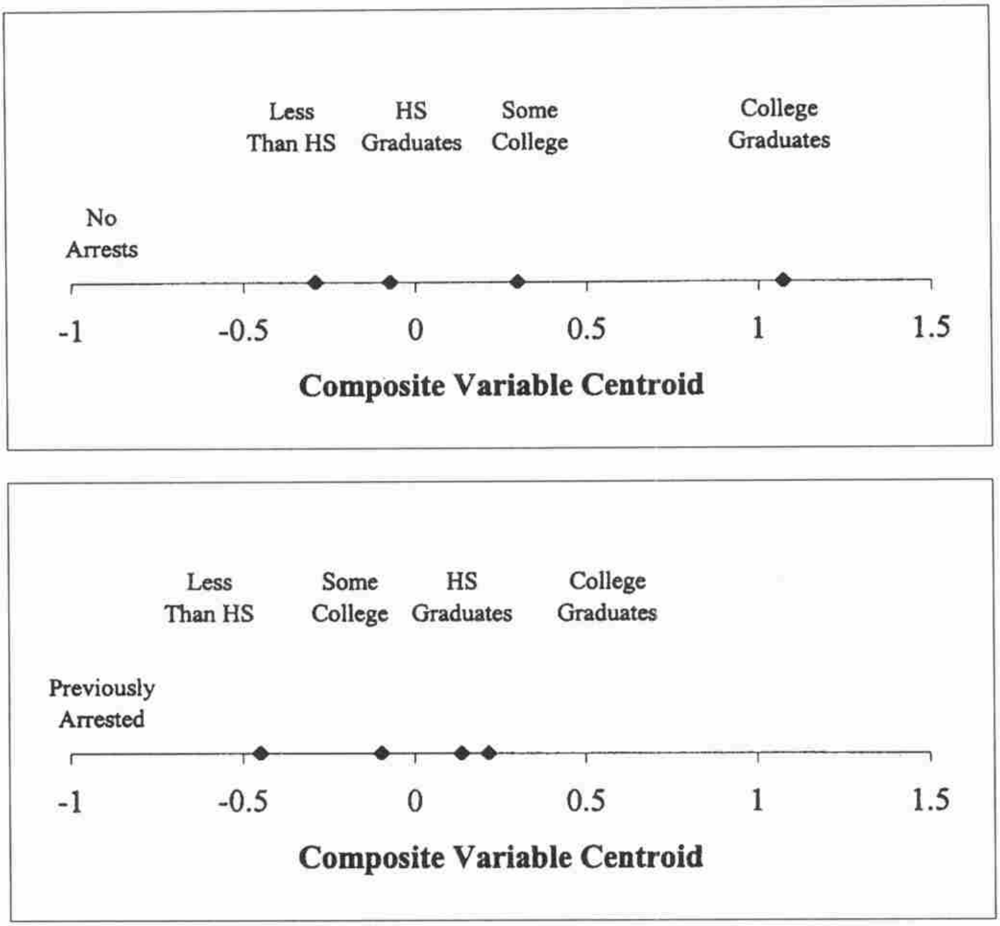

FIGURE 2

LDF Plot of Composite Variable Score Centroids for Two-Factor MANOVA Analysis

Note. HS = high school; LDF = linear discriminant function.

above differ from those in Equation 5 because a new set of standardized outcome variables was computed using the error terms from the factorial design (see the syntax in the Appendix).

To follow up the significant MANOVA interaction, the composite variable scores (*INTCOMP*) were submitted to a 2×4 ANOVA for further analysis. If it was of interest to follow up one of the main effects from the factorial MANOVA design, the main effect composite scores would be submitted to a univariate factorial ANOVA that corresponded to the original MANOVA design—in this case a 2×4 ANOVA. The significance of the interaction was already established by the MANOVA, so the purpose of this analysis is to obtain follow-up tests. Nevertheless, it is of interest to note that the partial η^2 effect size (i.e., $SS_{INT} / [SS_{INT} + SS_W]$) from the univariate analysis is equivalent to Roy's greatest root from the multivariate interaction. As noted earlier, both SAS and the SPSS GLM procedure express the Roy criterion (θ) as a function of **B** and **W**, so the univariate analysis of a composite from a factorial design will yield a SS_B / SS_W ratio equivalent to θ—this is true regardless of which design effect is being examined. Because SPSS MANOVA expresses θ as a function of **B** and **T**, where **T** is equal to **B** + **W** for each design effect, it is the partial η^2 value ($SS_B / [SS_B + SS_W]$) that is equivalent to θ in this case. Similarly, the *F* statistic from the univariate interaction, $F(3, 806) = 6.26$, $p < .01$, is identical in value to the *F* statistic used to test Roy's greatest root in a multivariate analysis (e.g., SPSS GLM, SAS PROC DISCRIM). This is useful because it verifies the computations for the composite variable scores.

Simple comparisons were conducted to examine differences among arrest groups within each level of *EDUC*, and a Bonferroni adjustment was used to control Type I error inflation across the set of four comparisons (i.e., $\alpha = .05 / 4 = .0125$). Statistically significant differences were observed among arrest conditions for the college graduates, $F(1, 806) = 11.38$, $p < .01$, as well as the individuals who attended but did not finish college, $F(1, 806) = 6.34$, $p = .01$. In both cases, the composite means (i.e., centroids) were higher for the no arrest group, indicating later onset of smoking behavior. Note that these significance tests are consistent with the pattern of group centroids shown in Figure 2. This should not be a surprise given that each comparison was performed using the same outcome variables weights (i.e., composite variable) that generated the significant MANOVA interaction effect in the first place.

Multivariate contrasts. Following the same analytic logic used in the composite analysis above, a series of contrasts akin to multivariate simple comparisons were performed, the syntax for which is shown in the Appendix. That is, mean vector differences were examined between the two arrest groups separately within each level of *EDUC*. It is important to reiterate that, although the same pairs of groups were being compared in the composite variable analysis above, a new set of outcome variable weights are derived that maximize group differences on each multivariate contrast. As such, there is no reason to expect the two follow-up procedures to produce identical (or even similar) results.

After performing a Bonferroni adjustment (i.e., $\alpha = .0125$), three of the multivariate contrasts between the two arrest groups were statistically significant: (a) the college graduates, $F(3, 804) = 3.81$, $p = .01$; (b) the group that attended some college, $F(3, 804) = 5.13$, $p < .01$; and (c) the less than high school group, $F(3, 804) = 5.20$, $p < .01$. Not only was a different contrast statistically significant in this case, but the substantive interpretation of these contrasts was quite different in some cases. For example, consider the group that attended some college. The structure coefficients for *CIGAGE, ALCAGE,* and *MJAGE* were –.65, –.89, and –.62, respectively. The standardized weights were as follows: *CIGAGE* = –.44, *ALCAGE* = –.71, and *MJAGE* = –.14. Recall that the *ALCAGE* structure coefficient for the omnibus interaction effect was .22, whereas the corresponding standardized weight was –.08. Clearly, *ALCAGE* was not a salient outcome variable if one considers the composite variable associated with the omnibus test, but was salient in this particular contrast. Again, this underscores the fact that the two follow-up procedures can yield substantive conclusions that are quite different.

DISCUSSION AND CONCLUSION

Two decades ago, Huberty and Smith (1982) encouraged researchers to "think multivariately" (p. 429) when using MANOVA techniques and outlined a method for conducting multivariate group contrasts in conjunction with DDA. However, subsequent content analyses of published studies (e.g., Keselman et al., 1998; Kieffer et al., 2001) indicate that researchers routinely *think univariately* when conducting follow-ups of statistically significant MANOVA effects, despite admonitions from methodologists about such analytic behavior (e.g., Huberty & Petoskey, 2000; Keselman et al., 1998; Thompson, 1999). Unfortunately, the practice of mixing and matching analytic approaches seems to be tacitly, if not explicitly, encouraged by many popular multivariate texts. As such, the goal of this article was to outline and illustrate two follow-up procedures that can be used to explore significant group differences in the MANOVA context: the analysis of composite variable scores and multivariate group contrasts.

From the preceding illustrations, it is clear that the two follow-up strategies may or may not yield results that are substantively consistent with one another, and the reason for this is straightforward. When computing and analyzing the composite variable scores computed from an omnibus MANOVA effect, the set of outcome variable weights is constant across all subsequent follow-up analyses. That is, the same linear combination of the outcome variables that produced the significant omnibus effect is also being examined in the subsequent comparisons. Parallels between the multivariate significance test and the corresponding univariate analysis were also pointed out (e.g., the ratio of explained variance and error variance was identical, the F statistics used to test the largest root are identical in value). In contrast, when using multivariate contrasts such as those outlined by Huberty and Smith (1982), each of the c contrasts is based on a different linear combination of the outcome variables. Although the c linear composites may be substantively similar to those obtained from the omnibus test, this will not always be the case.

In my view, the differences between follow-up procedures outlined above have at least three important implications for applied practice. First, the choice of multivariate follow-up procedure is not arbitrary and depends on the substantive question of interest. Like univariate ANOVA analyses, the use of MANOVA in applied practice can occur in one of two mutually exclusive situations: (a) those in which the researcher has no a priori contrasts of interest, and thus is interested in the omnibus test, and (b) those in which the researcher is interested in a set of a priori group contrasts. In the former case, the rejection of the multivariate null hypothesis necessitates further analyses to uncover which group differences are contributing to the statistically significant findings. In this situation, it seems eminently sensible to proceed with a follow-up strategy that investigates the same substantive phenomenon uncovered by the omnibus test, namely the analysis of the composite variable scores. However, in the case of a priori contrasts the omnibus effect would be of no interest to the researcher. This being the case, the use of multivariate contrasts is clearly the preferred procedure. In this context, the possibility of obtaining c unique linear combinations of the outcome variables is not problematic and, in fact, reflects the substantive goal of the analysis: to describe the nature of the differences between specific groups. The presence of multiple linear composites should not be viewed as inconsistent in this case, but rather as illuminating to the researcher's a priori hypotheses.

Second, the routine use of the univariate omnibus F test is frequently criticized by authors who favor the use of planned, or focused, comparisons (e.g., Olejnik & Hess, 1997; Thompson, 1994). Although it is beyond the scope of this article to extend specific criticisms to the multivariate context, the points I have made here should underscore the need for researchers to *think carefully* when thinking multivariately. Concerning the use of the omnibus F test in ANOVA, Olejnik and Hess stated, "We believe that most researchers have something more specific in mind when they design their investigations" (p. 229). That important substantive differences can result from the use of the two multivariate follow-up procedures only emphasizes the notion that researchers should think carefully about, and explicitly state, the specifics of their investigations. More so than in the univariate context, the nature of the research question and ensuing follow-up strategy can have a dramatic impact on the substantive interpretation of MANOVA results.

The third implication for applied practice concerns the vital importance of interpreting the LDF (i.e., DDA; reporting and interpreting structure coefficients and standardized weights) in conjunction with the exploration of group differences. This is not a new idea and is found in numerous works by Huberty and his colleagues (e.g., Huberty & Morris, 1989; Huberty & Petoskey, 2000; Huberty & Smith, 1982). When conducting post hoc tests using the composite variable, it makes little sense to discuss multivariate group differences without also describing the nature of the construct or LDF on which the groups differ. Perhaps even more important is the use of DDA in conjunction with multivariate contrasts. Clearly, if each of the *c* contrasts has the potential to yield a unique linear combination of the outcome variables, it is essential to describe each LDF and the associated outcome variable weights from each contrast. Without such a description, the group contrast results are of little scientific and descriptive utility.

In summary, it is hoped that applied researchers begin to "think multivariately" when choosing a follow-up strategy. It is important that researchers think carefully about whether or not their investigation can be cast using several a priori group contrasts, so that an appropriate follow-up analysis may be used. Finally, LDF weights should be presented and interpreted whenever performing MANOVA follow-ups, regardless of whether a priori contrasts or post hoc tests of the multivariate composite variable are used.

REFERENCES

Barcikowski, R. S., & Elliott, R. (1991, October). *Exploratory multivariate analysis of variance: Contrasts and variables.* Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago. (ERIC Document Reproduction Service No. ED339716)

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

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110,* 305–314.

Bray, J. H., & Maxwell, S. E. (1985). *Multivariate analysis of variance.* Newbury Park, CA: Sage.

Cole, D. A., Maxwell, S. E., Arvey, R., & Salas, E. (1993). Multivariate group comparisons of variable systems: MANOVA and structural equation modeling. *Psychological Bulletin, 114,* 174–184.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement, 61,* 229–248.

Cramer, E. M., & Bock, R. D. (1966). Multivariate analysis. *Review of Educational Research, 36,* 604–617.

Gabriel, K. R. (1968). Simultaneous test procedures in multivariate analysis of variance. *Biometrika, 55,* 489–504.

Horst, P. (1966). *Psychological measurement and prediction.* Belmont, CA: Wadsworth.

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

Huberty, C. J. (1994). *Applied discriminant analysis.* New York: Wiley & Sons.

Huberty, C. J., Chou, T. F., & Benitez, E. B. (1994). The study of multivariate group contrasts. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 3, pp. 123–148). Greenwhich, CT: JAI Press.

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

Huberty, C. J., & Petoskey, M. D. (2000). Multivariate analysis of variance and covariance. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of multivariate statistics and mathematical modeling* (pp. 183–208). San Diego, CA: Academic Press.

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

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donohue, B., et al. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA techniques. *Review of Educational Research, 3,* 350–386.

Kieffer, K. M., Reese, R. J., & Thompson, B. (2001). Statistical techniques employed in *AERJ* and *JCP* articles from 1988 to 1997: A methodological review. *Journal of Experimental Education, 69,* 280–309.

Levin, J. R., & Marascuilo, L. A. (1972). Type IV errors and interactions. *Psychological Bulletin, 78,* 368–374.

Maxwell, S. E. (1992). Recent developments in MANOVA applications. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 2, pp. 137–168). Stamford, CT: JAI Press.

Olejnik, S., & Hess, B. (1997). Top ten reasons why most omnibus ANOVA *F*-tests should be abandoned. *Journal of Vocational Education Research, 22,* 219–232.

Roy, S. N., & Bose, R. C. (1953). Simultaneous confidence interval estimation. *Annals of Mathematical Statistics, 24,* 513–536.

Scheffé, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika, 40,* 87–104.
Sheehan-Holt, J. K. (1998). MANOVA simultaneous test procedures: The power and robustness of restricted multivariate contrasts. *Educational and Psychological Measurement, 58,* 861–881.
Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Erlbaum.
Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins.
Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). New York: HarperCollins.
Thompson, B. (1994). Planned versus unplanned and orthogonal versus nonorthogonal contrasts: The neoclassical perspective. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 3, pp. 3–27). Stamford, CT: JAI Press.
Thompson, B. (1999, April). *Common methodology mistakes in educational research, revisited, along with a primer on both effect sizes and the bootstrap.* Invited address presented at the annual meeting of the American Educational Research Association, Montreal. (ERIC Document Reproduction Service No. ED429110)
U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2000). *National household survey on drug abuse, 1998.* Retrieved April 2, 2003, from the Inter-university Consortium for Political and Social Research Web site: http://www.icpsr.umich.edu:8080/ICPSR-STUDY/02934.xml

APPENDIX

One-Factor MANOVA Analyses

```
*OMNIBUS ONE-FACTOR MANOVA ANALYSIS.
manova
  cigage alcage mjage BY educ (1 4)
 /print = cellinfo(means) homogeneity(box)
 /discrim = all.

*STANDARDIZE DV'S USING SQRT OF MS WITHIN FOR EACH DV.
compute cigagez = (cigage - 16.71253) / sqrt(9.99299).
compute alcagez = (alcage - 14.78624) / sqrt(7.86087).
compute mjagez = (mjage - 15.67690) / sqrt(10.41751).

*COMPUTE THE COMPOSITE DV FOR THE ONE-FACTOR MANOVA.
compute onset = (.752 * cigagez)  + (- .547 * alcagez) + (.656 * mjagez).

*CONDUCT PAIRWISE CONTRASTS OF THE ONE-FACTOR MANOVA COMPOSITE DV.
unianova
   onset BY educ
  /emmeans = tables(educ) compare adj(LSD)
  /print = descriptive etasq
  /design = educ

*CONDUCT PAIRWISE MULTIVARIATE CONTRASTS FOR THE ONE-FACTOR MANOVA.
*THE CONTRAST SUBCOMMAND SPECIFIES EDUC 1 VS. 2, 2 VS. 3, AND 3 VS. 4.
manova
  cigage alcage mjage BY educ (1 4)
 /error = w
 /discrim = all
 /contrast (educ) = special(1 1 1 1, 1 -1 0 0, 0 1 -1 0, 0 0 1 -1)
 /design = educ (1) educ (2) educ (3).
```

(Continued on next page)

APPENDIX (Continued)

Two-Factor MANOVA Analyses

```
*OMNIBUS TWO-FACTOR MANOVA ANALYSIS.
manova
  cigage alcage mjage BY educ (1 4) arrested (0 1)
 /print = cellinfo(means) homogeneity(box)
 /discrim = all alpha (.05).

*STANDARDIZE DV'S USING SQRT OF MS WITHIN FOR EACH DV.
compute cigagezf = (cigage - 16.71253) / sqrt(9.77569).
compute alcagezf = (alcage - 14.78624) / sqrt(7.62927).
compute mjagezf = (mjage - 15.67690) / sqrt(10.27308).

*COMPUTE THE COMPOSITE DV FOR THE TWO-FACTOR MANOVA INTERACTION.
compute intcomp = (.980 * cigagezf) + (- .080 * alcagezf) + (.122 * mjagezf).

*CONDUCT FOLLOW-UPS ON INTERACTION COMPOSITE VARIABLE.
unianova
   intcomp BY educ arrested
  /emmeans = tables(educ*arrested) compare(arrested)
  /print = descriptive etasq
  /design = arrested educ arrested*educ.

*CONTRASTS TO OBTAIN MULTIVARIATE SIMPLE COMPARISONS.
manova
  cigage alcage mjage BY educ (1 4) arrested (0 1)
 /discrim = all
 /design = educ arrested within educ(1) arrested within educ(2)
  arrested within educ(3) arrested within educ(4).
```