A COMPARISON OF TWO FOLLOW-UP ANALYSES AFTER MULTIPLE ANALYSIS OF VARIANCE, ANALYSIS OF VARIANCE, AND DESCRIPTIVE DISCRIMINANT ANALYSIS: A CASE STUDY OF THE PROGRAM EFFECTS ON EDUCATION-ABROAD PROGRAMS

Alvin H. Yu
St. Cloud State University
hyu@stcloudstate.edu

Garry Chick
The Pennsylvania State University

Abstract.—This study compared the utility of two different post-hoc tests after detecting significant differences within factors on multiple dependent variables using multivariate analysis of variance (MANOVA). We compared the univariate *F* test (the Scheffé method) to descriptive discriminant analysis (DDA) using an educational-tour survey of university study-abroad students. Differences on five dependent measures of study-abroad program outcomes were compared by students' academic major and in-country residency arrangements (residence hall, home stay, apartment). Univariate *F* test and DDA were used to detect group differences in the dependent variables after the overall significant *F* test in MANOVA. The results indicated that the two post-hoc tests were congruent. However, DDA served better than the univariate *F* test to determine which outcome variables contributed most to separating the independent variables. We conclude that both post-hoc analyses should be utilized after a significant MANOVA to obtain accurate and full insight into the data.

1.0 INTRODUCTION

Multiple analysis of variance (MANOVA) is a statistical test that captures the effects of multiple independent variables on more than one dependent variable (Bray and Maxwell 1985). An important question in multivariate research is how best to examine and identify significant differences, if any, of varying values of multiple independent variables on multiple dependent variables. The univariate *F* test and descriptive discriminant analysis (DDA) are two major post-hoc approaches utilized to identify the differences between groups (Tabachnick and Fidell 2007). In reviewing the literature, Kieffer et al. (2001) found that more than 80 percent of reviewed studies utilizing MANOVA also employed univariate *F* tests, such as the Scheffé and Tukey tests, as post-hoc analyses. This common approach to post-hoc analysis fails to consider the shared variance between dependent variables, as a series of univariate *F* tests examines group differences only individually on each dependent variable. In light of this discrepancy in the research literature, our primary concern is the effectiveness of post-hoc analysis for examining group differences following a significant MANOVA.

DDA, a multivariate post-hoc procedure, may serve as an alternative approach to identify group differences because it takes all dependent variables into account and follows MANOVA ethos. As Huberty and Smith (1982) suggest, researchers might need to "think multivariately" (p. 429) as they conduct multivariate studies. Specifically, DDA identifies which variables contribute the most to separating predictor groups with one procedure, instead of individually evaluating a predictor group's unique effect on each of the dependent variables. The purpose of this study was to compare the utility of the univariate *F* test with DDA using data from a survey of study-abroad students on an educational tour. Comparisons of the two post-hoc procedures were made after significant results were found using MANOVA.

2.0 METHODS

We used data from a survey of students in a study-abroad program to observe the differences between two post-hoc analyses after MANOVA. Specifically, we attempted to identify the relationships of independent predictors to the combined program effects as evidenced by examining the relationships with multiple dependent variables. In summer 2005, the Web-based survey collected data on participants' perceptions of program effects on educational tours. The advantages of this survey program included lower publication costs (Nahm et al. 2004), convenient access (Schutt 2001), and an autonomous database system.

Table 1.—Characteristics of participants in effect-modeling study

Demographic Characteristics	Frequency	Percentage
Majors, N=263		
Art	28	11
Business	68	26
Liberal Arts	108	41
Science	59	22
Residency options, N=265		
Residency hall	104	39
Home stay	60	23
Apartment	101	38

In this example, 25 program effect items were generated from an intensive literature review and entered into an exploratory factor analysis. The five factors extracted from this analysis served as dependent variables in MANOVA and subsequent post-hoc tests. MANOVA was used to test for significant mean differences across the five program-effect variables for individuals who varied in terms of two independent predictors: academic majors and residency arrangements. MANOVA is used to control the possibilities of Type I error inflation as it examines the effects of independent variables on the dependent variables simultaneously instead of examining each dependent variable separately (Tabachnick and Fidell 2007).

When significant overall *F*-test values were identified in each MANOVA, we compared the two post-hoc procedures, Scheffé method for univariate *F* (see Hair et al. 2006) and DDA, to contrast their different approaches in determining the effects of group differences on the dependent variables. Unless otherwise noted, the criterion for statistical significance was set at .05 for these analyses.

3.0 RESULTS

Participating in this online study were 265 students, a sample size that provides adequate statistical power (Mertler and Vannata 2002). Participants' academic majors were categorized as art, business, liberal arts, and science. Residency options were sorted into three categories: apartment, home stay, and residence hall. Table 1 illustrates sample characteristics. Most participants were liberal arts or business majors. The home-stay option was not very common as most students lived in apartments and residence halls during their stay abroad.

A value of more than .60 levels in the Kaiser-Meyer-Olkin measurement and a significant Bartlett's test of sphericity suggested that this dataset was suitable for exploratory factor analysis (Tabachnick and Fidell 2007). The analysis produced a five-factor solution that was evaluated on the basis of three criteria: Eigenvalues, loading values, and scree plot (Mertler andVannatta 2002). Specifically, an item was deleted before the next statistical test if any of the following conditions were met: item with Eigenvalue lower than 1, factor loading value lower than .45 (with 20 percent overlapping variance (Comrey and Lee 1992), or being outside of the sharply descending line in a scree plot. As for cutoff levels for loading values, Comrey and Lee (1992) suggest that 0.40 levels with about 25 percent overlapping variance are appropriate criteria for discriminating factor loadings. In this study, two items (i.e., "I discovered that local people have opinions that differ from mine on some issues" and "I am more willing to interact with people with different cultural backgrounds than I was before my trip abroad") failed to meet the above criteria. These items were deleted before the next statistical test, MANOVA. About 51 percent of the total variance was explained in this analysis. The five factors served as scales measuring program outcomes for language learning, personal development, foreign connection, cultural immersion, and career development, which were each dependent variables for this study. Reliability analyses demonstrated that all scales met acceptable levels of reliability as Cronbach's alpha was 0.6 or higher for each, and this study represented exploratory research (Robinson et al. 1991). Table 2 lists the items associated with each scale and the corresponding factor loading values and Cronbach's alpha coefficients.

MANOVA was conducted to determine differences in students' majors with respect to the combined program outcomes. Because of its robustness, Pillai's Trace was utilized as the test statistic because the Box's *M* test showed that equal variances could not be assumed, $M = 66.262$, $F(45, 42863.66) = 1.406$, $p < .05$ (Mertler and Vannatta 2002). MANOVA results revealed

Table 2.—Factor loadings and Cronbach's alpha values of five program-effect model by exploratory factor analysis

Factor Names and Items	Loading	α
Factor 1, Language learning		
I think my destination language proficiency was improved after participating in the program.	.86	.84
I think this program gave me a better chance to practice the destination language than the classroom alone.	.85	
I really enjoyed learning the destination language in this program.	.80	
I could read some foreign newspaper without too much difficulty after I participated in the study-abroad program.	.57	
I could read some foreign newspaper without too much difficulty after I participated in the study abroad program.	.54	
Factor 2, Personal development		
I am more independent than I was before the program.	.67	.71
I feel more confident in myself after finishing the program.	.59	
I have different perceptions of what my life should be since my return from the other country.	.57	
I have a new outlook on the world since my study-abroad experience.	.56	
The program changed the perceptions about the local people at my study-abroad destination.	.34	
Factor 3, Foreign connection		
I felt that I was part of the local community during this trip.	.73	.72
I felt at home during this trip.	.60	
I feel I am still connected with local people even after the program was over.	.59	
I will visit this destination again if possible.	.52	
Factor 4, Cultural immersion		
Different cultural encounters enriched my study-abroad program.	.75	.64
The program gave me a chance to learn about a different culture.	.60	
The program helped me to increase my knowledge of the local culture at my study abroad destination.	.41	
To engage myself in a different cultural environment was one of my purposes for joining this program.	.40	
Factor 5, Career development		
I will consider employment in global companies because of my experience in the study- abroad program.	.68	.70
The program has raised my interest in some overseas job.	.63	
I have different views about my future career after my trip.	.50	
This program was important to me because it will someday be useful in getting a good job.	.58	
Deleted Items		
I discovered that local people have opinions that differ from mine on some issue.	.43	
I am more willing to interact with people with different cultural backgrounds than I was before my trip abroad.	.32	

Note: All items are measured on a 5-point scale of 1= Strongly disagree to 5 = Strongly agree.

significant program effect differences in the four majors, Pillai's Trace = 0.13, $F(15, 771) = 2.41$, $p < .05$, partial multivariate $\eta^2 = .05$.

In the next step, the Scheffé method with Bonferroni correction suggested that students' majors significantly differ in cultural immersion ($F(3, 263) = 6.43$, $p < .001$, partial $\eta^2 = .07$). In particular, it appears that art majors significantly differed from business and science majors in the cultural immersion dimension. Business and science majors were more likely to endorse cultural immersion than were art majors (i.e., 4.73 and 4.73 vs. 4.38, respectively). DDA was also used as one of the post-hoc methods to identify which program effect factors would contribute most to separating the major groups (Art, Business, Liberal Arts, and Science). For that, the five dependent variables served as predictors to separate the major groups. DDA revealed that one out

Table 3.—Results of discriminant function analysis of the program effects by majors

	Discriminant Function	
Variables	Correlation Coefficients with Discriminant Function	Standardized Discriminant Function
Language learning	.17	-.01
Personal development	-.04	-.37
Foreign connection	.21	-.13
Cultural immersion	.94	1.06
Career development	.16	.17

Note: Wilks' Lambda Λ= 0.87, $\chi^2(5, N = 265) = 35.70$, p<.05

Table 4.—MANOVA of program-effect factors by majors

		Majors				
Program Effects	Overall Mean	Art	Business	Liberal Arts	Science	F- value (df = 263)
		-------------------*Mean Agreement Score*----------------				
Language learning	3.80	3.64	3.99	3.82	3.63	2.58
Personal development	4.25	4.23	4.32	4.27	4.15	1.10
Foreign connection	3.94	3.83	4.04	3.91	3.92	0.70
Cultural immersion	4.65	$4.38_{a,\ b}$	4.73_{a}	4.64	4.73_{b}	6.43***
Career development	3.85	3.78	4.06	3.81	3.72	2.75*

Note: Overall agreement score was measured on a 5-point scale of 1 = Strongly disagree to 5 = Strongly agree. Means sharing the same subscript differ at p < 0.01 in the Scheffé comparison, two-tailed.
* p <.05; ** p <.01; *** p <.001

of three generated functions was significant [Λ = 0.87, χ^2 (5, N = 265) = 35.70, p <.05], indicating that the construct differentiating majors was primarily the cultural immersion scale. It had a canonical correlation coefficient of .94 (see Table 3). Results of the Scheffé's post-hoc method (see Table 4) with a Bonferroni adjustment (.01 significance levels) were congruent with findings from DDA; that is, art majors significantly differed from business and science majors in the cultural immersion dimension.

As for residency arrangement and program effects, MANOVA was conducted to determine the location differences, (i.e., including apartment, home stay, and residence hall), in the combined five program-outcome factors. Pillai's Trace statistic was used to evaluate the significant relationship between residency location and effects since equal variances could not be assumed in the Box's M test, Pillai's Trace = .31, $F(10, 518) = 9.30$, p = .002, partial multivariate $\eta^2 = 0.15$. Based on univariate F tests with Bonferroni adjustment (see Table 5), residency location category differences were significant for language learning, $F(2, 262) = 42.57$, p <.001, partial $\eta^2 = 0.25$. The home-stay group had a significantly higher mean value than the residence hall group (4.51 vs. 3.46) and the apartment group (4.51 vs. 3.73) in terms of perceived language learning effects (Table 6). DDA also indicated that the language learning effect (with a canonical correlation coefficient of 0.92) served as the best variable to distinguish residency group across the three levels, apartment, home stay, and residence hall, Λ = 0.70, $\chi^2(5, N = 265) = 91.47$, p <.05.

4.0 DISCUSSION OF THEORETICAL IMPLICATIONS

Since the 1980s, the univariate F test has been widely employed as a standard procedure following a significant MANOVA test based on the suggestion of Cramer and Bock (1966). Further illustrating its wide usage as a post-hoc procedure, Keselman and colleagues (1998) reviewed 79 MANOVA studies from 1994 through 1995 and found only four studies that used DDA as a

Table 5.—Results of discriminant function analysis of the program effects by residence arrangements

	Discriminant Function	
Variables	Correlation Coefficients with Discriminant Function	Standardized Discriminant Function
Language learning	0.92	1.00
Personal development	-0.23	-0.25
Foreign connection	-0.12	-0.04
Cultural immersion	-0.03	-0.15
Career development	-0.08	-0.10

Note: Wilks' Lamda Λ=.70, $\chi^2(10, N = 265) = 91.47$, $p < .001$

Table 6.—MANOVA of program-effect factors by residence arrangements

		Residency Location			
Program Effects	Overall Mean	Residence Hall	Home Stay	Apartment	F- value (df = 263)
		--------------*Mean Agreement Score*------------			
Language learning	3.80	3.46_a	$4.51_{a,b}$	3.73_b	42.57***
Personal development	4.26	4.30	4.10	4.31	3.04*
Foreign connection	3.94	4.03	3.87	3.89	1.16
Cultural immersion	4.65	4.68	4.65	4.64	0.25
Career development	3.86	3.96	3.84	3.76	2.04

Note: Overall agreement score was measured on a 5-point scale of 1 = Strongly disagree to 5 = Strongly agree. Means sharing the same subscript differ at p<0.01 in the Scheffé comparison, two-tailed.
* p <.05; ** p <.01; *** p <.001

follow-up test. This predominant use occurred despite a major concern about utilizing a univariate approach as post-hoc analyses for a multivariate procedure and despite fundamental flaws with this approach. Specifically, a univariate post-hoc procedure does not take the correlations between dependent variables into account. Considering these limitations, Field (2005) strongly recommends the use of DDA as it accounts for correlations between dependent variables and meets the ethos of MANOVA .

The results for this study indicated that the post-hoc tests from the univariate F test (Scheffé method) and DDA were congruent when we tested for significant differences between factors in MANOVA. In a similar study, Enders (2003) suggested that the choice of post-hoc analysis depends on the research question of interest: exploring group differences or identifying the contribution of variables to separating the predictor groups. Specifically, the univariate test is recommended when researchers emphasize understanding the differences of group means on the dependent variables. Alternatively, a multivariate approach, such as DDA, could be employed to determine which group contributes most to the significant MANOVA results. DDA can also help understand which variable might capture group differences when all univariate F tests are not significant. In this exploratory study, we were not able to identify a priori which approach to utilize as the more appropriate post-hoc analysis procedure, and both approaches were employed and compared to determine how each identified group differences on the dependent measures.. The comparisons in this study suggested that both types of post-hoc analyses, the univariate F test and DDA, were useful in providing a full understanding of the differences in the data, and offered a case where employing both methods is recommended.

5.0 CITATIONS

Babbie, E. 2001. **The practice of social research (9th ed.).** Belmont, CA: Wardsworth/Thomas Learning.

Bray, J.H.; Maxwell, S.E. 1985. **Multivariate analysis of variance.** Beverly Hills, CA: Sage.

Comrey, A.L.; Lee, H.B. 1992. **A first course in factor analysis (2nd ed.).** Hillside, NJ: Lawrence Erlbaum.

Cramer, E.M.; Bock, R.D. 1966. **Multivariate analysis.** Review of Educational Research. 36: 604-617.

Enders, C.K. 2003. **Performing multivariate group comparisons following a statistically significant MANOVA.** Measurement and Evaluation in Counseling and Development. 36: 40-56.

Field, A.P. 2005. **Discovering statistics using SPSS (2nd ed.).** London, UK: Sage.

Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E.; Tatham, R.L. 2006. **Multivariate data analysis (6th ed.)** Upper Saddle River, NJ: Pearson.

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.; Donahue, B.; Kowalchuk, R.K.; Lowman, L.L.; Petoskey, M.D.; Keselman, J.C.; Levin, J.R. 1998. **Statistical practices of education researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses.** Review of educational research. 68: 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.

Mertler, C.A.; Vannatta, R.A. 2002. **Advanced and multivariate statistical methods: Practical application and interpretation (2nd ed.).** Los Angeles: Pyrczak.

Nahm, E.; Preece, J.; Resnick, B. 2004. **Usability of health web sites for older adults: A preliminary study.** Computers, Informatics, Nursing. 22: 326-334.

Robinson, J.P.; Shaver, P.R.; Wrightsman, L.S. 1991. **Criteria for scale selection and evaluation.** In: Robinson, J.P.; Shaver, P.R.; Wrightsman, L.S., eds. Measures of personality and social psychological attitudes. New York: Academic Press.

Schutt, R.K. 2001. **Investigating the social world (3rd ed.).** Thousand Oaks, CA: Pine Forge Press.

Tabachnick, B.G.; Fidell, L.S. 2007. **Using multivariate statistics, (5th ed.).** Boston, MA: Allyn and Bacon.