

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Copyright is retained by the first or sole author, who grants right of first publication to *Practical Assessment, Research & Evaluation*. Permission is granted to distribute this article for nonprofit, educational purposes if it is copied in its entirety and the journal is credited. PARE has the right to authorize third party reproduction of this article in print, electronic and database forms.

Volume 19, Number 17, November 2014

ISSN 1531-7714

A Primer on Multivariate Analysis of Variance (MANOVA) for Behavioral Scientists

Russell T. Warne, *Utah Valley University*

Reviews of statistical procedures (e.g., Bangert & Baumberger, 2005; Kieffer, Reese, & Thompson, 2001; Warne, Lazo, Ramos, & Ritter, 2012) show that one of the most common multivariate statistical methods in psychological research is multivariate analysis of variance (MANOVA). However, MANOVA and its associated procedures are often not properly understood, as demonstrated by the fact that few of the MANOVAs published in the scientific literature were accompanied by the correct post hoc procedure, descriptive discriminant analysis (DDA). The purpose of this article is to explain the theory behind and meaning of MANOVA and DDA. I also provide an example of a simple MANOVA with real mental health data from 4,384 adolescents to show how to interpret MANOVA results.

One of the most common multivariate statistical procedures in the social science literature is multivariate analysis of variance (MANOVA). In one review of educational psychology journals, MANOVA was used in 23.0% quantitative articles, making it the most common multivariate procedure (Warne, Lazo, Ramos, & Ritter, 2012, p. 138)—results that agree with Bangert and Baumberger (2005), who found that MANOVA was the most commonly used multivariate statistical procedure in a clinical psychology journal. Other reviewers of statistical procedures in psychological research (e.g., Kieffer, Reese, & Thompson, 2001; Zientek, Capraro, & Capraro, 2008) have found MANOVA to be the most popular multivariate statistical method in the published literature.

Despite the popularity of MANOVA, the method is persistently misunderstood, and researchers are particularly unfamiliar with the proper statistical procedures after rejecting a multivariate null hypothesis. Warne et al. (2012), for example, found that

in only 5 out of 62 articles that used MANOVA (8.1%) in educational psychology journals included the correct post hoc procedure: discriminant descriptive analysis (DDA). Tonidandel and LeBreton (2013) found that almost every article in the *Journal of Applied Psychology* in which MANOVA was used did not include a follow-up DDA. In my scrutiny of three major psychology journals—the *Journal of Clinical Psychology*, *Emotion*, and the *Journal of Counseling Psychology*—I found a total of 58 articles in which researchers used MANOVA between 2009 and 2013. However, in none of these articles did researchers use a DDA. Huberty and Morris (1989) found similar results when investigating six leading psychology and education research journals. The purpose of this article is to explain the principles of MANOVA and how to correctly perform and interpret a MANOVA and its associate post hoc procedure, DDA. A simple example of a MANOVA and DDA using real data relevant to

behavioral scientists follows the explanation to make the article more applicable.

What is MANOVA?

MANOVA is a member of the General Linear Model—a family of statistical procedures that are often used to quantify the strength between variables (Zientek & Thompson, 2009). Many of the procedures in the General Linear Model are hierarchically organized—i.e., more specific procedures are often special cases of general procedures (Zientek & Thompson, 2009). MANOVA, specifically, is an analysis of variance (ANOVA) that has two or more dependent variables (Fish, 1988). Because most behavioral scientists already have a firm grasp of ANOVA (Aiken, West, & Millsap, 2008), I will use it as the starting point in this article for explaining MANOVA.

In an ANOVA the independent variable is a nominal variable that has two or more values, and the dependent variable is interval or ratio scaled.¹ The null hypothesis is that the mean score on the dependent variable will be statistically equal for every group. As with any null hypothesis statistical significance testing procedure, an observed statistic is calculated and then compared to a sampling distribution. If the observed statistic is found to be a more extreme value in the sampling distribution than the critical value (as shown in Figure 1), then the null hypothesis will be rejected; otherwise, the null hypothesis is retained. Afterwards, an effect size—usually η^2 in an ANOVA—quantifies the relationship between the independent and dependent variables (Thompson, 2006).

MANOVA is merely an ANOVA that has been mathematically extended to apply to situations where there are two or more dependent variables (Stevens, 2002). This necessitates several changes to the logic of ANOVA. The first is displayed in Figure 2. In contrast to Figure 1, which is a one dimensional number line, Figure 2 shows the rejection region of a MANOVA as a region outside of a circle on a two-dimensional Cartesian plane. Therefore, instead of the observed statistic being expressed as a point on a number line, it can be graphically expressed as a vector (Thomas, 1992). If the vector extends into the rejection region,

then the null hypothesis is rejected (as would be the case for vectors A, C, and D in Figure 2). However, if a vector does not extend into the rejection region (e.g., vector B in Figure 2), then the null hypothesis is retained.

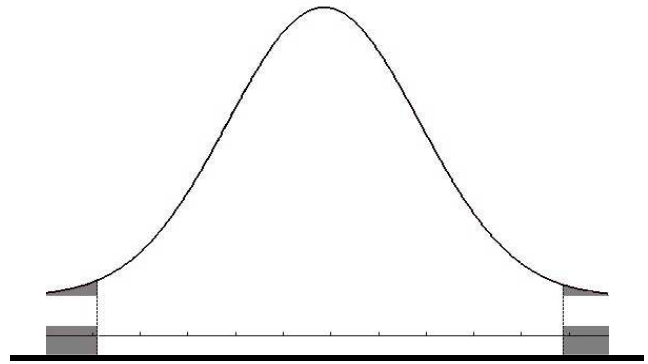


Figure 1. Example of the logic of univariate hypothesis testing.

The shaded area is the rejection region. If the test statistic is in the rejection region (shown both on the histogram and on the number line below), then the null hypothesis is rejected.

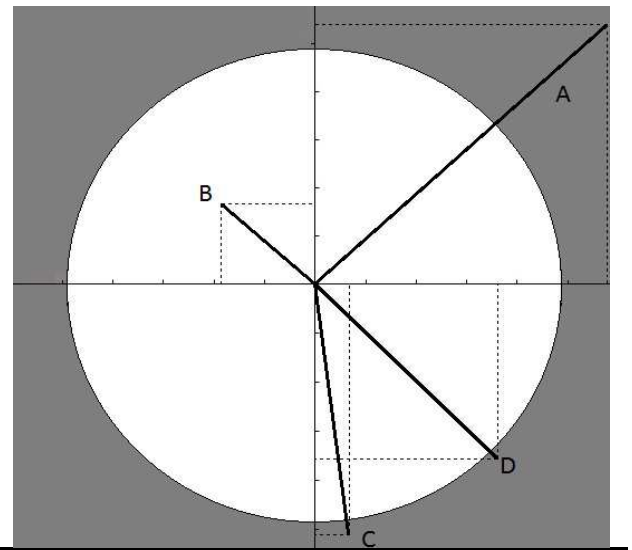


Figure 2. Example of the logic of multivariate analysis of variance hypothesis testing for perfectly uncorrelated variables.

The shaded area is the rejection region. If the vector for the test ends in the rejection region, then the null hypothesis is rejected. If the dotted lines originating at the end of a vector meet an axis in the shaded region, then the ANOVA for that axis's dependent variable would also be rejected.

¹ Please note that a one-way ANOVA simplifies to a *t*-test if the independent nominal variable has only two groups (Thompson, 2006).

Figure 2 also shows how the results of a MANOVA can vary from the results of two ANOVAs. Each vector in the figure can be decomposed into two pieces (which are represented in the figure by the dotted lines), one for each dependent variable (Saville & Wood, 1986). Notice how vector A extends beyond the point on both axes where the unshaded region ends. This means that if an ANOVA were conducted on either dependent variable (which visually would collapse the two-dimensional plane into a number line resembling the lower portion of Figure 1), the null hypothesis of the ANOVA would be rejected. Likewise, analyzing vector B would produce a non-statistically significant MANOVA and two ANOVAs that were both also non-statistically significant. On the other hand, vector C would produce a statistically significant MANOVA and one statistically significant ANOVA (for the vertical axis). Finally, vector D would produce a statistically significant MANOVA, but in both ANOVAs the null hypotheses would be retained.

Why Use MANOVA?

Given MANOVA's more complicated nature, some researchers may question whether it is worth the added complexity. The alternative to using MANOVA is to conduct an ANOVA for each dependent variable. However, this approach is not advantageous because (a) conducting multiple ANOVAs increases the likelihood of committing a Type I error, and (b) multiple ANOVAs cannot determine whether independent variable(s) are related to combinations of dependent variables, which is often more useful information for behavioral scientists who study correlated dependent variables.

Avoiding Type I Error Inflation

When multiple dependent variables are present, conducting a MANOVA is beneficial because it reduces the likelihood of Type I error (Fish, 1988; Haase & Ellis, 1987; Huberty & Morris, 1989). The probability of Type I error at least once in the series of ANOVAs (called experiment-wise error) can be as high as $1 - (1 - .05)^k$, where k is the number of ANOVAs conducted. Therefore, if a researcher chooses the traditional α value of .05 for two ANOVAs, then the experiment-wise Type I error can be as high as .0975—not .05—even though the α for each ANOVA is .05. However, the Type I error for a MANOVA on the

same two dependent variables would be only .05. The severity of Type I error inflation in multiple ANOVAs depends on how correlated the dependent variables are with one another, with Type I error inflation being most severe when dependent variables are uncorrelated (Hummel & Sligo, 1971). Readers will recognize the Type I error inflation that occurs with multiple ANOVAs because experiment-wise Type I error inflation also occurs when conducting multiple independent sample t-tests. In fact, one of the main rationales for conducting ANOVA is to avoid conducting multiple t-tests, which inflates the Type I error that occurs when researchers conduct many t-tests (Thompson, 2006).²

Figure 2 represents an ANOVA with two perfectly uncorrelated dependent variables. However, if dependent variables are correlated, then the MANOVA is better represented in Figure 3. Notice how the axes are no longer perpendicular with each other. Rather, the correlation between the dependent variables is represented by the angle of the axes, with a more oblique angle indicating a higher correlation between dependent variables (Saville & Wood, 1986). Although the axes in Figures 3 look strange, it is important for readers to remember that in statistics all axes and scales are arbitrary. Therefore, having nonperpendicular axes is acceptable. In exploratory factor analysis, for example, factor solutions are almost always rotated so that the pattern of factor loadings is more interpretable (Costello & Osborne, 2005). Indeed, in oblique rotation methods, such as promax rotation, the axes are not only rotated, but can be nonperpendicular—just like the axes in Figure 3 (Thompson, 2004).

Although the examples in Figures 2 and 3 represent MANOVAs with two dependent variables, an extension to three dependent variables can be imagined without difficulty. For three perfectly uncorrelated dependent variables, the graph would be three dimensional, with the unshaded region being a

² Some researchers who choose to conduct multiple t-tests instead of an ANOVA control Type I error inflation with a Bonferroni correction (Thompson, 2006). The analogous relationship between t-tests and ANOVA and between ANOVA and MANOVA extend even to this point, and applying a Bonferroni correction would also control Type I error inflation in a group of ANOVAs. However, Bonferroni methods are overly conservative, especially with a large number of tests (Stevens, 2002) or correlated variables (Hummel & Sligo, 1971). Moreover, a Bonferroni correction to a series of ANOVAs would not provide the benefits of MANOVA described elsewhere in the article.

perfect sphere. As the variables became more correlated, the axes would become more oblique, and the sphere would become more distorted and resemble something like a football. If two of the three dependent variables were perfectly correlated (i.e., $r = +1$), then the graph would collapse into a two-dimensional graph like in Figure 3. A MANOVA with four or more dependent variables requires more than three dimensions and axes—something that most people have difficulty imagining visually, but which a computer can easily handle mathematically.

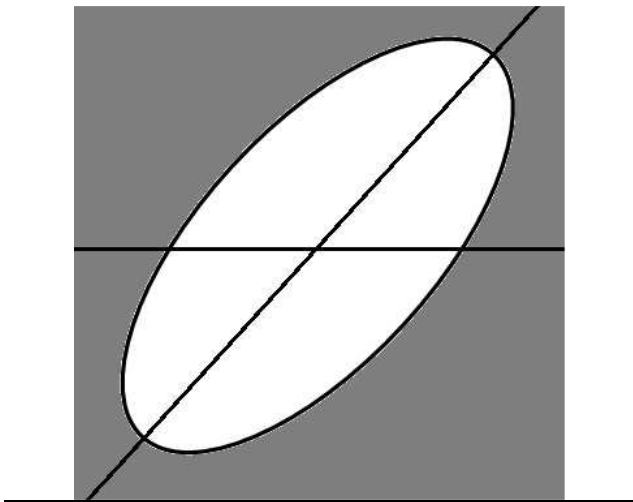


Figure 3. Example of the logic of multivariate analysis of variance hypothesis testing for moderately correlated variables.

The shaded area is the rejection region.

Examining Combinations of Variables

ANOVA and MANOVA also differ in that they investigate completely different empirical questions. Zientek and Thompson (2009, p. 345) explained that, “. . . two ANOVAs actually test the differences of the means on the observed or measured variables . . . whereas the MANOVA actually tests the differences of the mean of the DDA function scores” of the groups. In other words, MANOVA tests the differences between underlying unobserved latent variables (derived from the variables in the dataset), while ANOVA only tests differences among groups on an observed variable (Haase & Ellis, 1987; Huberty & Morris, 1989; Zientek & Thompson, 2009). MANOVA is therefore often more useful to social scientists than ANOVA because most topics they research are latent constructs that are not directly observable, such as beliefs and attitudes. With ANOVA it is assumed that these constructs are measured without error and with a

single observed variable—an unrealistic assumption for many constructs in the behavioral sciences. Therefore, MANOVA is a statistical procedure that is more in accordance than ANOVA with behavioral scientists’ beliefs about the topics they study.

Post Hoc Procedures

Post hoc procedures are often necessary after the null hypothesis is rejected in an ANOVA (Thompson, 2006) or a MANOVA (Stevens, 2002). This is because the null hypotheses for these procedures often do not provide researchers with all the information that they desire (Huberty & Morris, 1989; Thomas, 1992). For example, if the null hypothesis is rejected in an ANOVA with three or more groups, then the researcher knows that at least one group mean statistically differs from at least one other group mean. However, most researchers will be interested in learning which mean(s) differ from which other group mean(s). Many ANOVA post hoc tests have been invented to provide this information, although Tukey’s test is by far the most common test reported in the psychological literature (Warne et al., 2012).

Just like ANOVA, MANOVA has post hoc procedures to determine why the null hypothesis was rejected. For MANOVA this is usually a DDA, which is a statistical procedure which creates a set of perfectly uncorrelated linear equations that together model the differences among groups in the MANOVA (Fish, 1988; Stevens, 2002). The benefit of having uncorrelated equations from a DDA is that each function will provide unique information about the differences among groups, and the information can be combined in an additive way for easy interpretation.

Unfortunately, most researchers who use MANOVA do not use DDA to interpret their MANOVA results (Huberty & Morris, 1989; Kieffer et al., 2001; Warne et al., 2012). This may be because when researchers use SPSS to conduct a MANOVA, the computer program automatically conducts an ANOVA for each dependent variable. However, SPSS does not automatically conduct a DDA when the null hypothesis for a MANOVA has been rejected, even though DDA is a more correct post hoc procedure.

Zientek and Thompson (2009) explained that because MANOVA analyzes latent variables composed of observed variables and ANOVA is only concerned with observed variables, using ANOVA as a post hoc

procedure following an ANOVA is non-sensical (see also Kieffer et al., 2001). This is because ANOVA and MANOVA were developed to answer completely different empirical questions (Huberty & Morris, 1989). Indeed, Tonidandel and LeBreton (2013) explained that, . . . *multivariate theories yield multivariate hypotheses which necessitate the use of multivariate statistics and multivariate interpretations of those statistics. By invoking univariate ANOVAs as follow-up tests to a significant MANOVA, researchers are essentially ignoring the multivariate nature of their theory and data.* (p. 475)

Moreover, Fish (1988) and Huberty and Morris (1989) showed that the same data can produce different results when analyzed using a MANOVA or a series of ANOVAs—a situation that could potentially be misleading to researchers who follow MANOVA with a series of post hoc ANOVAs.

The purpose of the remainder of this article is to provide a real-life example of a MANOVA with a single nominal independent variable and two dependent variables and a post hoc DDA as a simple blueprint for future researchers who wish to use this procedure. By doing so, I hope to make conducting and interpreting MANOVA and DDA as easy as possible for readers.

Conducting a MANOVA

The example data in this study is taken from the National Longitudinal Study of Adolescent Health (Add Health), a longitudinal study of adolescent and adult development that started in the 1994-1995 school year when its participants were in grades 7-12. Data for this example were taken from a public use file of variables collected during the initial data collection period and downloaded from the Inter-University Consortium for Political and Social Research web site. For this example the only independent variable is grade, and the two dependent variables are participants' responses to the questions (a) "In the last month, how often did you feel depressed or blue?" and (b) "In the last month, how often did you have trouble relaxing?" Both responses were recorded on a 5-point Likert scale (0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Often, and 4 = Every day). For convenience these variables will be called "depression" and "anxiety," respectively. As previous researchers have found that anxiety and depressive disorders are often comorbid (e.g., Mineka,

Watson, & Clark, 1998), it is theoretically sound to conduct a MANOVA to investigate the relationship between grade and both outcomes at the same time. The preference of a MANOVA over a univariate test is further strengthened by the moderately strong correlation between depression and anxiety ($r = .544$, $p < .001$) in the dataset. The null hypothesis for this MANOVA is that the six independent variable groups are equal on both dependent variables. For a MANOVA with additional groups or more than two dependent variables, the null hypothesis would be that all groups are equal on all dependent variables (Stevens, 2002).

Table 1 shows descriptive statistics for the variables in the example data from the Add Health study. A MANOVA was conducted to determine the relationship between grade and the combined variables of depression and anxiety. The results from SPSS 19.0 are shown in Table 2, which I modified slightly to make it easier to read.

Table 1. Descriptive Statistics of Variables Used in the Example MANOVA

Grade	n	Depression		Anxiety	
		M	SD	M	SD
7	662	0.88	1.114	0.75	1.048
8	664	1.08	1.188	0.80	1.058
9	778	1.17	1.188	0.93	1.080
10	817	1.27	1.232	0.96	1.109
11	790	1.37	1.204	1.12	1.156
12	673	1.34	1.140	1.10	1.105

Note. Listwise deletion in all analyses, so all results are based on the 4,384 cases that had data on all three variables.

Readers will notice a few things about Table 2. First, there are two sections: intercept (containing four rows) and grade (containing another four rows). The intercept section of the table is necessary to scale the results and does not provide any substantive information for interpretation. The grade section, however, displays results from the hypothesis test. Second, there are four rows, each of which display four statistical test statistics: (a) Pillai's Trace, (b) Wilks' Lambda, (c) Hotelling's Trace, and (d) Roy's Largest Root. All four of these are test statistics for the same null hypothesis, although their formulas differ (Olson, 1976). All four can be converted to an F-statistic, which can then be used to calculate a p-value, which are displayed in Table 2.

Table 2. Example MANOVA Results (Modified from SPSS Output)

Effect	Test Statistic	Value	F	df	Sig. (<i>p</i>)	Partial η^2	Noncentrality Parameter
Intercept	Pillai's Trace	0.533	2470.323	2, 4337	< .001	.533	4940.647
	Wilks' Lambda	0.467	2470.323	2, 4337	< .001	.533	4940.647
	Hotelling's Trace	1.139	2470.323	2, 4337	< .001	.533	4940.647
	Roy's Largest Root	1.139	2470.323	2, 4337	< .001	.533	4940.647
Grade	Pillai's Trace	0.022	9.834	10, 8676	< .001	.011	98.340
	Wilks' Lambda	0.978	9.873	10, 8674	< .001	.011	98.726
	Hotelling's Trace	0.023	9.911	10, 8672	< .001	.011	99.112
	Roy's Largest Root	0.021	18.388	5, 4338	< .001	.021	91.939

All four of the MANOVA test statistics for the Add Health data are statistically significant ($p < .001$). This indicates that the null hypothesis is rejected, and grade has a statistically significant relationship with the combined dependent variables of depression and anxiety. In many cases the results from the four statistical tests will be the same (O'Brien & Kaiser, 1985). In fact, these statistics will *always* produce the same result when there are only two groups in the independent variable (Haase & Ellis, 1987; Huberty & Olejnik, 2006) or when there is a single dependent variable (because this simplifies the test to an ANOVA; see Olson, 1976).

However, it is possible that the four statistical tests will produce contradictory results, with some statistics lacking statistical significance and others being statistically significant (Haase & Ellis, 1987). This is because the four tests differ in their statistical power, with Roy's great root having the most power when dependent variables are highly correlated—possibly because they all represent a single construct—and the other three having the more power for more disparate outcomes (Huberty & Olejnik, 2006). Additionally, Pillai's trace is highly robust to many violations of the assumptions of MANOVA (Olson, 1976). Regardless of the test statistic(s) that a researcher uses, he or she should explicitly report which test statistic(s) he or she has chosen and not merely state that the results are statistically significant (Haase & Ellis, 1987; Olson, 1976). Readers can see that in this example all four test statistics provided evidence that the null hypothesis should be rejected, even though the values of the statistics and their degrees of freedom vary because of their varying formulas and theoretical distributions (Haase & Ellis, 1987).

Table 2 contains additional information of note. First, the effect sizes (i.e., partial η^2 values) are all either .011 and .021—indicating that grade does not have a particularly powerful statistical relationship with the dependent variables. Second, the noncentrality parameter for each test statistic is displayed in Table 2. This noncentrality parameter is a statistical parameter necessary to estimate a noncentral distribution that models the distribution of test statistics if the null hypothesis is not true (Steiger & Fouladi, 1997). When the noncentrality parameter is equal to zero, the null hypothesis perfectly fits the data and therefore should not be rejected. Therefore, it makes sense that all four test statistics have high noncentrality parameter values (between 91.939 and 99.112) in Table 2 because all test statistics suggest that the null hypothesis should be rejected, and each one has a very small p-value ($p < .001$). Finally, the column labeled “Observed Power” gives the a posteriori statistical power of the MANOVA design for each test statistic. The numbers in this column will always vary inversely with the corresponding p-value, so they provide no information beyond that provided by the p-values in the table (Hoenig & Heisey, 2001).

Post Hoc DDA

With the results of the MANOVA in Table 2 showing that the null hypothesis should be rejected, it is necessary to conduct a post hoc DDA in order to determine the mathematical function(s) that distinguish the grade groups from one another on dependent variable scores. The number of functions that are created will vary, but the minimum will always be the number of dependent variables or the number of groups minus 1—whichever value is smaller (Haase &

Ellis, 1987; Stevens, 2002; Thomas, 1992). Therefore, in the Add Health example there will be two discriminant functions because there are two dependent variables.

In SPSS, a descriptive discriminant analysis can be conducted through the “discriminant” tool. The computer program produces a table labeled “Wilks’ Lambda,” and displays the results of the significance test for the two discriminant functions. In this example the first discriminant function is statistically significant ($p < .001$), while the second is not ($p = .125$), despite the large sample size ($n = 4,384$). Therefore, it is not necessary to interpret the second discriminant function.

SPSS displays a number of other tables that aid in the interpretation of discriminant functions, the two most important of which show the standardized discriminant coefficients (SDCs) and the structure coefficients for the discriminant functions. These tables are displayed in Table 3. The SDCs can be used to calculate a score for each subject for each discriminant function. Given the numbers in Table 3, the function score for the first function can be calculated as:

$$DDA\ Score_1 = .700\textit{Depression} + .431\textit{Anxiety}$$

These SDC values are interpreted the same way as the standardized regression coefficients in a multiple regression (Stevens, 2002). Therefore, the .700 in the equation means that for every 1 standard deviation increase in examinees’ depression scores their DDA score is predicted to increase by .700 standard deviations if all other variables are held constant. More substantively, because discriminant functions maximize the differences among the grade groups, it can be seen that although both depression and anxiety contribute to group differences, depression shows more between-group variation. This is supported by the data in Table 1, which shows that from Grades 7 to 11 depression and anxiety scores increase steadily, although the changes in depression scores are more dramatic. These results also support studies showing that both anxiety disorders and depression often begin in adolescence (e.g., Cole, Peeke, Martin, Truglio, & Seroczynski, 1998).

However, it is important to recognize that the SDCs cannot indicate which variable is more “important” for distinguishing groups. The easiest way one can determine variable importance is by calculating a parallel discriminant ratio coefficient (DRC; Thomas,

Table 3. Example Standardized Discriminant Coefficients and Structure Coefficients (Modified from SPSS Output)

Standardized Discriminant Coefficients ^a		
	Function 1	Function 2
Depression	.700	-.958
Anxiety	.431	1.105
Structure Coefficients ^b		
	Function 1	Function 2
Depression	.932	-.364
Anxiety	.808	.590
Parallel Discriminant Ratio Coefficients		
	Function 1	Function 2
Depression	(.700)(.932) = .652	(-.958)(-.364) = .349
Anxiety	(.431)(.808) = .348	(1.105)(.590) = .652

^a SPSS displays this information in a table labeled “Standardized Canonical Discriminant Function Coefficients.”

^b SPSS displays this information in a table labeled “Structure Matrix.” *Note.* Parallel DRCs should always add to 1.0 (Thomas & Zumbo, 1996), but do not in Function 2 because of rounding.

1992; Thomas & Zumbo, 1996), which requires the structure coefficients for the variables. The second section of Table 3 shows the SPSS output of the structure coefficients, which are the correlations between each observed variable and the DDA scores calculated from each discriminant function. This is analogous to structure coefficients in multiple regression, which are correlations between independent variables and predicted values for the dependent variable (Thompson, 2006). To calculate a parallel DRC, one must multiply each SDC with the corresponding structure coefficient (Thomas, 1992), as is demonstrated at the bottom of Table 3. For the first function the parallel DRCs are .652 for depression and .348 for anxiety—indicating that depression is clearly the more important variable of the two in distinguishing between independent variable groups in this example.

Although parallel DRCs are highly useful, their interpretation may be complicated by the presence of a suppressor variable, which is a variable that has little correlation with a dependent variable but which can strengthen the relationship that other variables have

with a dependent variable and thereby improve model fit if included (Smith, Ager, & Williams, 1992; Thompson, 2006; Warne 2011). Because a suppressor variable and a useless variable will both have small parallel DRCs, it is often necessary to calculate a total DRC to determine if a suppressor variable is present using the following equation, adapted from Equation 21 of Thomas and Zumbo (1996):

$$\text{Total DRC} = |\text{SDC}| \sqrt{\frac{\left(1 + \frac{df_{\text{Between}}}{F}\right)}{\frac{df_{\text{Within}}}{1 + \lambda}}}$$

In this equation, df_{Between} and df_{Within} represent the degrees of freedom for the model's between- and within-group components, which are the same degrees of freedom as in an ANOVA for the same independent variable on a single dependent variable (Olson, 1976). F represents the same F-ratio for that variable, and λ is the discriminant function's eigenvalue. All information needed to calculate a total DRC for each independent variable is available to a researcher in the SPSS output for a MANOVA and DDA. Calculating total DRCs is not necessary for this example because both the depression and anxiety variables have high enough parallel DRCs that both are clearly useful variables in the discriminant function, and neither is a plausible suppressor variable.

Conclusion

Although this MANOVA and its accompanying DDAs were simple, analyses with additional independent or dependent variables would not be much more complex. Regardless of the number of variables, interpreting the MANOVA is the same as interpreting Tables 2 and 3. More complex models could produce more than two DDA functions, but interpreting additional functions is the same as interpreting the two functions in the Add Health example. Moreover, with additional functions, SPSS often produces interesting figures that can show researchers the degree of overlap among groups and how different their average members are from one another. These figures can help researchers understand the discriminant functions and how the groups differ from one another. For example, Shea, Lubinski, and Benbow (2001, pp. 607-610) presented five graphs that efficiently showed in three dimensions the differences

among average members of their DDA groups, which ranged in number from four to eight.

This article on MANOVA and DDA was designed to help social scientists better use and understand one of the most common multivariate statistical methods in the published literature. I hope that researchers will use the information in this article to (a) understand the nature of MANOVA and DDA, (b) serve as a guide to correctly performing their own MANOVA and post hoc DDA analyses, and (c) aid in the interpretation of their own data and of other studies that use these methods. Researchers in the behavioral sciences should use this powerful multivariate analysis method more often in order to take advantage of its benefits. I also encourage researchers to investigate other post hoc procedures, such as dominance analysis (Azen & Budescu, 2003, 2006), which has advantages over DDA. However, other methods are more complex and difficult to perform than a DDA, and many researchers may not find the advantages worth the extra time and expertise needed to use them. Therefore, I believe that DDA is often an expedient choice for researchers who wish to understand their multivariate analyses. Regardless of the choice of post hoc method, researchers should completely avoid using univariate methods after rejecting a multivariate null hypothesis (Huberty & Morris, 1989; Kieffer et al., 2001; Tonidandel & LeBreton, 2013; Zientek & Thompson, 2009).

References

- Aiken, L. S., West, S. G., & Millsap, R. E. (2008). Doctoral training in statistics, measurement, and methodology in psychology: Replication and extension of Aiken, West, Sechrest, and Reno's (1990) survey of PhD programs in North America. *American Psychologist*, *63*, 32-50. doi:10.1037/0003-066X.63.1.32
- Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, *8*, 129-148. doi:10.1037/1082-989x.8.2.129
- Azen, R., & Budescu, D. V. (2006). Comparing predictors in multivariate regression models: An extension of dominance analysis. *Journal of Educational and Behavioral Statistics*, *31*, 157-180. doi:10.3102/10769986031002157
- Bangert, A. W., & Baumberger, J. P. (2005). Research and statistical techniques used in the *Journal of Counseling & Development*: 1990-2001. *Journal of Counseling & Development*, *83*, 480-487. doi:10.1002/j.1556-6678.2005.tb00369.x

- Cole, D. A., Leeke, L. G., Martin, J. M., Truglio, R., & Seroczynski, A. D. (1998). A longitudinal look at the relation between depression and anxiety in children and adolescents. *Journal of Consulting and Clinical Psychology, 66*, 451-460. doi:10.1037/0022-006X.66.3.451
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9. Available online at <http://pareonline.net/getvn.asp?v=10&n=7>
- Fish, L. J. (1988). Why multivariate methods are usually vital. *Measurement and Evaluation in Counseling and Development, 21*, 130-137.
- Haase, R. F., & Ellis, M. V. (1987). Multivariate analysis of variance. *Journal of Counseling Psychology, 34*, 404-413. doi:10.1037/0022-0167.34.4.404
- Hoening, J. M., & Heisey, D. M. (2001). The abuse of power. *The American Statistician, 55*, 19-24. doi:10.1198/000313001300339897
- Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin, 105*, 302-308. doi:10.1037/0033-2909.105.2.302
- Huberty, C. J., & Olejnik, S. (2006). *Applied MANOVA and discriminant analysis*. Hoboken, NJ: John Wiley & Sons.
- Hummel, T. J., & Sligo, J. R. (1971). Empirical comparison of univariate and multivariate analysis of variance procedures. *Psychological Bulletin, 76*, 49-57. doi:10.1037/h0031323
- 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. doi:10.1080/00220970109599489
- Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377-412. doi:10.1146/annurev.psych.49.1.377
- O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method of analyzing repeated measures designs: An extensive primer. *Psychological Bulletin, 97*, 316-333. doi:10.1037/0033-2909.97.2.316
- Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin, 83*, 579-586. doi:10.1037/0033-2909.83.4.579
- Saville, D. J., & Wood, G. R. (1986). A method for teaching statistics using N-dimensional geometry. *The American Statistician, 40*, 205-214. doi:10.1080/00031305.1986.10475394
- Smith, R. L., Ager, J. W., Jr., & Williams, D. L. (1992). Suppressor variables in multiple regression/correlation. *Educational and Psychological Measurement, 52*, 17-29. doi:10.1177/001316449205200102
- Steider, J. H., & Fouladi, R. T. (1997). *Noncentrality interval estimation and the evaluation of statistical models*. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 221-257). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Thomas, D. R. (1992). Interpreting discriminant functions: A data analytic approach. *Multivariate Behavioral Research, 27*, 335-362. doi:10.1207/s15327906mbr2703_3
- Thomas, D. R., & Zumbo, B. D. (1996). Using a measure of variable importance to investigate the standardization of discriminant coefficients. *Journal of Educational and Behavioral Statistics, 21*, 110-130. doi:10.3102/10769986021002110
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
- Thompson, B. (2006). *Foundations of behavioral statistics*. New York, NY: The Guilford Press.
- Tonidandel, S., & LeBreton, J. M. (2013). Beyond step-down analysis: A new test for decomposing the importance of dependent variables in MANOVA. *Journal of Applied Psychology, 98*, 469-477. doi:10.1037/a0032001
- Warne, R. T. (2011). Beyond multiple regression: Using commonality analysis to better understand R² results. *Gifted Child Quarterly, 55*, 313-318. doi:10.1177/0016986211422217
- Warne, R. T., Lazo, M., Ramos, T., & Ritter, N. (2012). Statistical methods used in gifted education journals, 2006-2010. *Gifted Child Quarterly, 56*, 134-149. doi:10.1177/0016986212444122
- Zientek, L. R., Capraro, M. M., & Capraro, R. M. (2008). Reporting practices in quantitative teacher education research: One look at the evidence cited in the AERA panel report. *Educational Researcher, 37*, 208-216. doi:10.3102/0013189X08319762
- Zientek, L. R., & Thompson, B. (2009). Matrix summaries improve research reports: Secondary analyses using published literature. *Educational Researcher, 38*, 343-352. doi:10.3102/0013189X09339056

Acknowledgment:

The author appreciates the feedback provided by Ross Larsen during the writing process.

Citation:

Warne, Russell (2014). A Primer on Multivariate Analysis of Variance (MANOVA) for Behavioral Scientists. *Practical Assessment, Research & Evaluation*, 19(17). Available online:
<http://pareonline.net/getvn.asp?v=19&n=17>

Corresponding Author:

Russell T. Warne
Department of Behavioral Science
Utah Valley University
800 W. University Parkway MC 115
Orem, UT 84058
Email: rwarne@uvu.edu