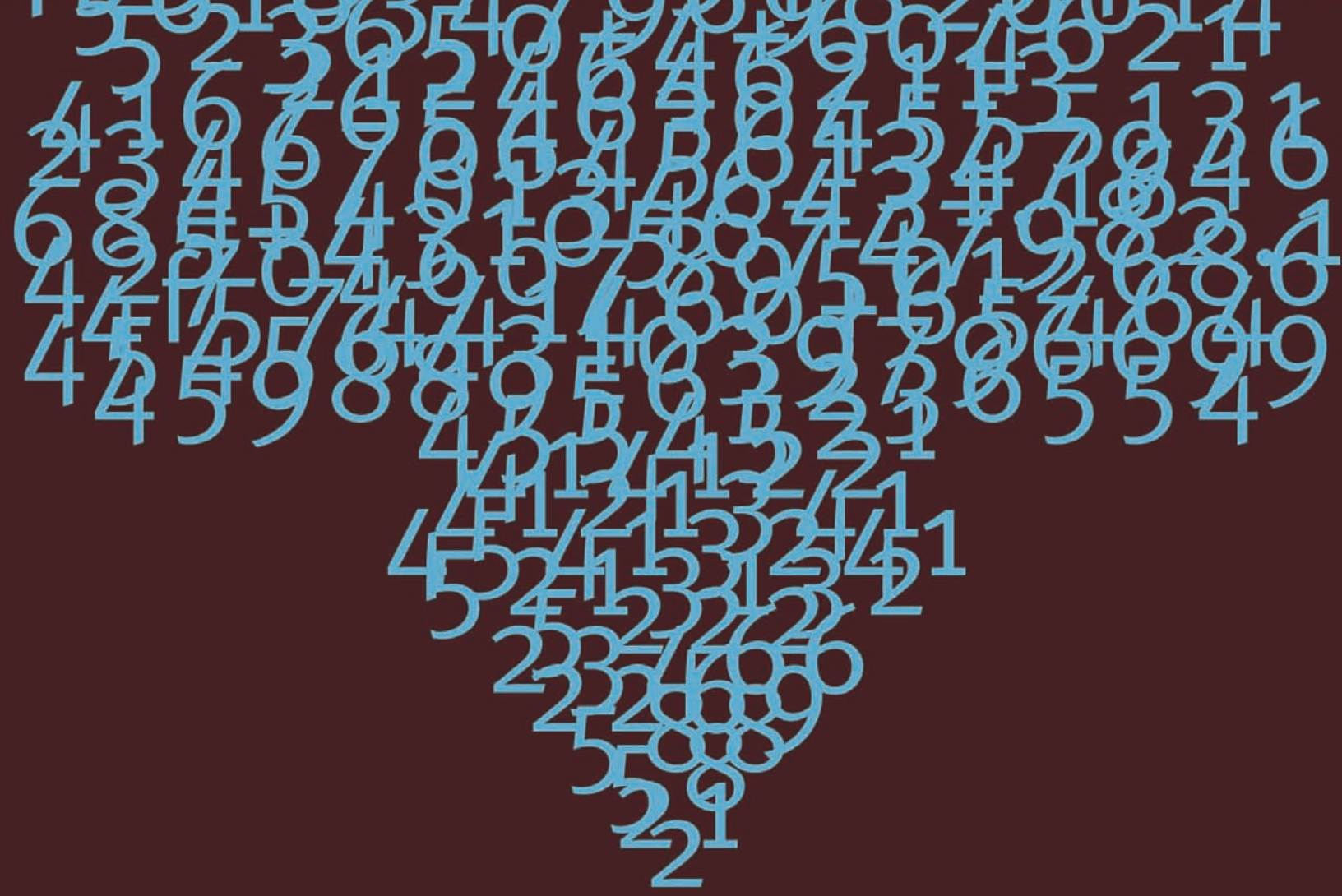

Pearson New International Edition

Using Multivariate Statistics

Barbara G. Tabachnick Linda S. Fidell

Sixth Edition

Pearson New International Edition

Using Multivariate Statistics
Barbara G. Tabachnick Linda S. Fidell
Sixth Edition

PEARSON

For the example, we assign WRAT-R scores higher priority since reading problems represent the most common presenting symptoms for learning disabled children. To keep overall alpha below .05, individual alpha levels are set at .025 for each of the two DVs. WRAT-R scores are analyzed through univariate ANOVA, as displayed in Table 7. Because the main effect of disability is not interesting and the interaction is not statistically significant in MANOVA (Table 2), the only effect of interest is treatment. The critical value for testing the treatment effect (6.55 with 1 and 12 df at $\alpha = .025$) is clearly exceeded by the obtained F of 46.1225.

WRAT-A scores are analyzed in ANCOVA with WRAT-R scores as covariate. The results of this analysis appear in Table 9.[12] For the treatment effect, critical F with 1 and 11 df at $\alpha = .025$ is 6.72. This exceeds the obtained F of 5.49. Thus, according to stepdown analysis, the significant effect of treatment is represented in WRAT-R scores, with nothing added by WRAT-A scores.

Note that WRAT-A scores show significant univariate but not stepdown F. The lack of significance for WRAT-A scores in stepdown analysis does not mean that they are unaffected by treatment, but rather that no unique variability is shared with treatment after adjustment for differences in WRAT-R. This result occurs despite the relatively low correlation between the DVs.

This procedure can be extended to sets of DVs through MANCOVA. If the DVs fall into categories, such as scholastic variables and attitudinal variables, one can ask whether there is any change in attitudinal variables as a result of an IV, after adjustment for differences in scholastic variables. The attitudinal variables serve as DVs in MANCOVA while the scholastic variables serve as covariates.

5.3.3 Using Discriminant Analysis

Discriminant analysis provides information useful in assessing DVs (DVs are predictors in the context of discriminant analysis). A structure (loading) matrix is produced which contains correlations between the linear combination of DVs that maximizes treatment differences and the DVs themselves. DVs that correlate highly with the combination are more important to discrimination among groups.

Discriminant analysis can also be used to test each of the DVs in the standard multiple regression sense; the effect on each DV is assessed after adjustment for all other DVs. That is, each DV is assessed as if it were the last one to enter an equation.

TABLE 9 Analysis of Covariance of WRAT-A Scores, With WRAT-R Scores as the Covariate

Source	SS	df	MS	*F*
Covariate	1.7665	1	1.7665	0.0361
D	538.3662	2	269.1831	5.5082
T	268.3081	1	268.3081	5.4903
DT	52.1344	2	26.0672	0.5334
S(DT)	537.5668	11	48.8679	

[12]A full stepdown analysis is produced as an option through IBM SPSS MANOVA. For illustration, however, it is helpful to show how the analysis develops.

5.3.4 Choosing Among Strategies for Assessing DVs

The choice between univariate and stepdown *F* is not always easy, and often you want to use both. When there is no correlation among the DVs, univariate *F* with adjustment for Type I error is acceptable. When DVs are correlated, as they almost always are, stepdown *F* is preferable on grounds of statistical purity, but you have to prioritize the DVs and the results can be difficult to interpret.

If DVs are correlated and there is some compelling priority ordering of them, stepdown analysis is clearly called for, with univariate *F*s and pooled within-cell correlations reported simply as supplemental information. For significant lower-priority DVs, marginal and/or cell means adjusted for higher-priority DVs are reported and interpreted.

If the DVs are correlated but the ordering is somewhat arbitrary, an initial decision in favor of stepdown analysis is made. If the pattern of results from stepdown analysis makes sense in the light of the pattern of univariate results, interpretation takes both patterns into account with emphasis on DVs that are significant in stepdown analysis. If, for example, a DV has a significant univariate *F* but a nonsignificant stepdown *F*, interpretation is straightforward: The variance the DV shares with the IV is already accounted for through overlapping variance with one or more higher-priority DVs. This is the interpretation of WRAT-A in the preceding section and the strategy followed in Section 6.

But if a DV has a nonsignificant univariate *F* and a significant stepdown *F*, interpretation is much more difficult. In the presence of higher-order DVs as covariates, the DV suddenly takes on "importance." In this case, interpretation is tied to the context in which the DVs entered the stepdown analysis. It may be worthwhile at this point, especially if there is only a weak basis for ordering DVs, to forgo evaluation of statistical significance of DVs and resort to simple description. After finding a significant multivariate effect, unadjusted marginal and/or cell means are reported for DVs with high univariate *F*s but significance levels are not given.

An alternative to attempting interpretation of either univariate or stepdown *F* is interpretation of loading matrices in discriminant analysis. This process is facilitated when IBM SPSS MANOVA or SAS GLM is used because information about the discriminant functions is provided as a routine part of the output. Alternatively, a discriminant analysis may be run on the data.

Another perspective is whether DVs differ significantly in the effects of IVs on them. For example: Does treatment affect reading significantly more than it affects arithmetic? Tests for contrasts among DVs have been developed in the context of meta-analysis with its emphasis on comparing effect sizes. Rosenthal (2001) demonstrates these techniques.

5.4 Specific Comparisons and Trend Analysis

When there are more than two levels in a significant multivariate main effect and when a DV is important to the main effect, the researcher often wants to perform specific comparisons or trend analysis of the DV to pinpoint the source of the significant difference. Similarly, when there is a

significant multivariate interaction and a DV is important to the interaction, the researcher follows up the finding with comparisons on the DV. Specific comparisons may also be done on multivariate effects. These are often less interpretable than comparisons on individual DVs, unless DVs are all scaled in the same direction, or are based on factor or principal component scores.

Comparisons are either planned (performed in lieu of omnibus *F*) or post hoc (performed after omnibus *F* to snoop the data). When comparisons are post hoc, an extension of the Scheffé procedure is used to protect against inflated Type I error due to multiple tests. The procedure is very conservative but allows for an unlimited number of comparisons. Following Scheffé for ANOVA, the tabled critical value of *F* is multiplied by the degrees of freedom for the effect being tested to produce an adjusted, and much more stringent, *F*. If marginal means for a main effect are being contrasted, the degrees of freedom are those associated with the main effect. If cell means are being contrasted, our recommendation is to use the degrees of freedom associated with the interaction.

The difference between setting up contrasts on individual DVs and setting up contrasts on the combination is that all DVs are included in the syntax. Table 10 shows syntax for trend analysis and user-specified orthogonal contrasts on the main effect of DISABLTY for the small-sample example. The coefficients illustrated for the orthogonal contrasts actually are the trend coefficients. Note that IBM SPSS GLM requires fractions in part of the LMATRIX command to produce the right answers.

Use of this syntax also provides univariate tests of contrasts for each DV. None of these contrasts is adjusted for post hoc analysis. The usual corrections are to be applied to minimize inflated Type I error rate unless comparisons are planned.

5.5 Design Complexity

When between-subjects designs have more than two IVs, extension of MANOVA is straightforward as long as sample sizes are equal within each cell of the design. The partition of variance continues to follow ANOVA, with a variance component computed for each main effect and interaction. The pooled variance–covariance matrix due to differences among subjects within cells serves as the single error term. Assessment of DVs and comparisons proceed as described in Sections 5.3 and 5.4.

Two major design complexities that arise, however, are inclusion of within-subjects IVs and unequal sample sizes in cells.

5.5.1 Within-Subjects and Between-Within Designs

The simplest design with repeated measures is a one-way within-subjects design where the same subjects are measured on a single DV on several different occasions. The design can be complicated by addition of between-subjects IVs or more within-subjects IVs.

Repeated measures analysis extends to MANOVA when the researcher measures several DVs on several different occasions. The occasions can be viewed in two ways. In the traditional sense, occasions is a within-subjects IV with as many levels as occasions. Alternatively,

TABLE 10 Syntax for Orthogonal Comparisons and Trend Analysis

Type of Comparison	Program	Syntax	Section of Output	Name of Effect
Orthogonal	IBM SPSS GLM	GLM WRATR WRATA BY TREATMNT DISABLTY /METHOD = SSTYPE(3) /INTERCEPT = INCLUDE /CRITERIA = ALPHA(.05) /LMATRIX "LINEAR" DISABLTY 1 0 -1 TREATMNT*DISABLTY 1/2 0 -1/2 1/2 0 -1/2 /LMATRIX "QUADRATIC" DISABLTY 1 -2 1 TREATMNT*DISABLTY 1/2 -2/2 1/2 1/2 -2/2 1/2 /DESIGN = TREATMNT DISABLTY TREATMNT*DISABLTY.	**Custom Hypothesis Tests:** **Multivariate Test Results**	Wilks' Lambda
	IBM SPSS MANOVA	MANOVA WRATR WRATA BY TREATMNT (1, 2) DISABLTY (1, 3) /METHOD = UNIQUE /PARTITION (DISABLTY) /CONTRAST(DISABLTY)=SPECIAL(1 1 1, 1 0 -1, 1 -2 1) /DESIGN = TREATMNT DISABLTY(1) DISABLTY(2) TREATMNT BY DISABLTY.	EFFECT... DISABLTY(2) EFFECT... DISABLTY(1)	Wilks' Lambda
	SAS GLM	PROC GLM DATA=SASUSER.SS_MANOV; CLASS TREATMNT DISABLTY; MODEL WRATR WRATA = TREATMNT DISABLTY TREATMNT*DISABLTY; CONTRAST 'LINEAR' DISABLTY 1 0 -1; CONTRAST 'QUADRATIC' DISABLTY 1 -2 1; manova h=_all_/short; run;	MANOVA Test Criteria.... No Overall linear (quadratic) Effect	Wilks' Lambda

(*continued*)

TABLE 10 Continued

Type of Comparison	Program	Syntax	Section of Output	Name of Effect
Trend Analysis	IBM SPSS GLM	No special syntax; done as any other userspecified contrasts.	EFFECT... DISABLTY(2)	Wilks' Lambda
	IBM SPSS MANOVA	MANOVA WRATR WRATA BY TREATMNT(1,2) DISABLTY(1,3) /METHOD = UNIQUE /PARTITION (DISABLTY) /CONTRAST(DISABLTY)= POLYNOMIAL (1,2,3) /DESIGN = TREATMNT DISABLTY(1) DISABLTY(2) TREATMNT BY DISABLTY.	EFFECT... DISABLTY(1)	
	SAS GLM	No special syntax for between-subjects IVs.		

each occasion can be treated as a separate DV—one DV per occasion (Section 2.8). In this latter view, if there is more than one DV measured on each occasion, the design is said to be doubly multivariate—multiple DVs are measured on multiple occasions. (There is no distinction between the two views when there are only two levels of the within-subjects IV.)

It also is possible to have multiple DVs, but treat the within-subjects IV univariately. This is useful when (1) there are only two levels of the within-subjects IV, (2) there is no concern with violation of sphericity, or (3) a trend analysis is planned to replace the omnibus tests of the within-subjects IV and any interactions with the within-subjects IV. All programs that do doubly multivariate analysis also show univariate results.

5.5.2 Unequal Sample Sizes

When cells in a factorial ANOVA have an unequal number of scores, the sum of squares for effect plus error no longer equals the total sum of squares, and tests of main effects and interactions are correlated. There are a number of ways to adjust for overlap in sums of squares (cf. Woodward & Overall, 1975). Both the problem and the solutions generalize to MANOVA.

All the MANOVA programs described in Section 7 adjust for unequal *n*. IBM SPSS MANOVA offers both Method 1 adjustment (METHOD = UNIQUE), which is default, and Method 3 adjustment (METHOD = SEQUENTIAL). Method 3 adjustment with survey data through IBM SPSS MANOVA is shown in Section 6.2. Method 1—called SSTYPE(3)—is the default

among four options in IBM SPSS GLM. In SAS GLM, Method 1 (called `TYPE III or TYPE IV`) also is the default among four options available.

6 Complete Examples of Multivariate Analysis of Variance and Covariance

In the research described in Appendix 'Research Designs for Complete Examples', Section 1, there is interest in whether the means of several of the variables differ as a function of sex role identification. Are there differences in self-esteem, introversion–extraversion, neuroticism, and so on associated with a woman's masculinity and femininity? Files are MANOVA.*.

Sex role identification is defined by the masculinity and femininity scales of the Bem Sex Role Inventory (Bem, 1974). Each scale is divided at its median to produce two levels of masculinity (high and low), two levels of femininity (high and low), and four groups: Undifferentiated (low femininity, low masculinity), Feminine (high femininity, low masculinity), Masculine (low femininity, high masculinity), and Androgynous (high femininity, high masculinity). The design produces a main effect of masculinity, a main effect of femininity, and a masculinity–femininity interaction.[13]

DVs for this analysis are self-esteem (ESTEEM), internal versus external locus of control (CONTROL), attitudes toward women's role (ATTROLE), socioeconomic level (SEL2), introversion–extraversion (INTEXT), and neuroticism (NEUROTIC). Scales are coded so that higher scores generally represent the more "negative" trait:low self-esteem, greater neuroticism, etc.

Omnibus MANOVA (Section 6.2) asks whether these DVs are associated with the two IVs (femininity and masculinity) or their interaction. The Roy–Bargmann stepdown analysis, in conjunction with the univariate *F* values, allows us to examine the pattern of relationships between DVs and each IV.

In a second example (Section 6.3), MANCOVA is performed with SEL2, CONTROL, and ATTROLE used as covariates and ESTEEM, INTEXT, and NEUROTIC used as DVs. The research question is whether the three personality DVs vary as a function of sex role identification (the two IVs and their interaction) after adjusting for differences in socioeconomic status, attitudes toward women's role, and beliefs regarding locus of control of reinforcements.

6.1 Evaluation of Assumptions

Before proceeding with MANOVA and MANCOVA, we must assess the variables with respect to practical limitations of the techniques.

6.1.1 Unequal Sample Sizes and Missing Data

IBM SPSS FREQUENCIES is run with SORT and SPLIT FILE to divide cases into the four groups. Data and distributions for each DV within each group are inspected for missing values, shape, and variance (see Table 11 for output on the CONTROL variable for the Feminine group). The run reveals the presence of a case for which the CONTROL score is missing. No datum is missing

[13]Some would argue with the wisdom of considering masculinity and femininity separate IVs, and of performing a median split on them to create groups. This example is used for didactic purposes.