Between

```
manova DV by A(1,2) B(1,3)
 /design=B, A w B(1), A w B(2), A w B(3)
 /design=A, B w A(1), B w A(2).
```

Within

```
manova V1 to V4 /wsfactors=A(2) B(2)
 /wsdesign=B, A w B(1), A w B(2).

manova V1 to V4 /wsfactors=A(2) B(2)
 /wsdesign=A, B w A(1), B w A(2).
```

Mixed

```
manova V1 V2 by A(1,2)
 /wsfactors=WSIVname(2)
 /design=mwithin A(1), mwithin A(2).

manova V1 V2 by A(1,2)
 /wsfactors=WSIVname(2)
 /wsdesign=mwithin WSIVname(1), mwithin WSIVname(2).
```

A more detailed discussion can be found in the following pages, extracted from an SPSS newsletter.
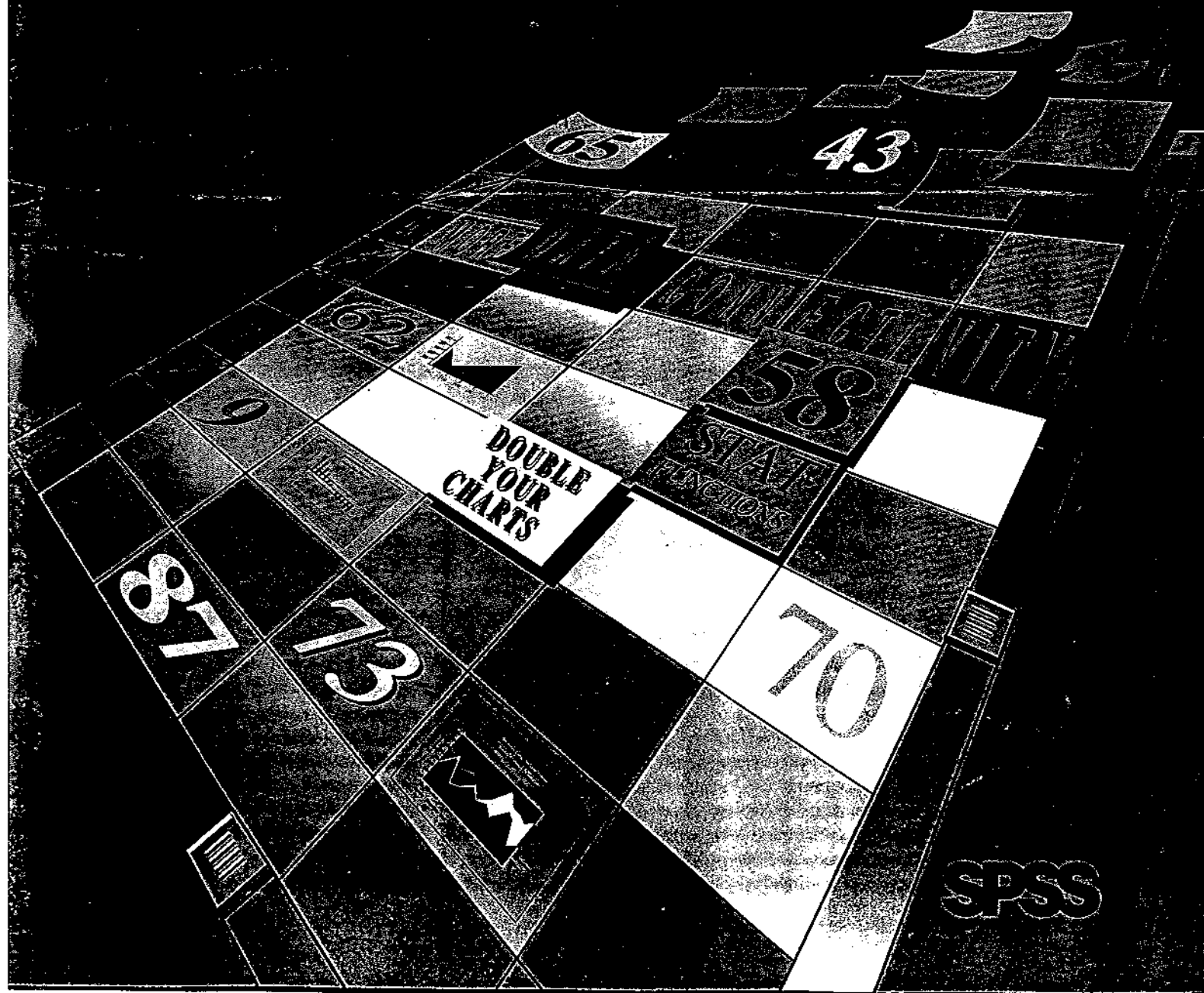
KEYW**O**RDS

TIPS AND NEWS FOR STATISTICAL SOFTWARE USERS

SPSS

Testing Simple Effects in MANOVA

Factorial designs in analysis of variance and covariance, including designs with within-subjects factors, are very common in many fields of research. The SPSS MANOVA procedure provides a powerful and flexible set of tools for performing most of the analyses that are available under the general linear model framework. A very common problem is that of an experiment in which interactions have been found and the researcher wants to explore the data more carefully to determine what statements may be made about main effects or interactions in the presence of the two-way or higher order interaction effects. Tests of such simple main effects or simple interaction effects are generally easily handled in MANOVA through the flexibility in model specification offered by the subcommands DESIGN and WSDESIGN.

Two-way between-subjects models: estimating simple main effects

Let's begin with the simplest case in which we might want to test for simple effects: a two-way factorial design in which we have found an interaction effect. If the two factors are both between-subjects factors and have two and three levels respectively, we might have the following syntax for the factorial analysis:

```
MANOVA Y BY A(1,2) B(1,3).
```

This one line of syntax will produce the full factorial analysis (MANOVA always does a full factorial model by default), equivalent to specifying either

```
MANOVA Y BY A(1,2) B(1,3)
 /DESIGN
         or
MANOVA Y BY A(1,2) B(1,3)
 /DESIGN=A, B, A BY B.
```

If the A by B interaction term is nonzero, the effects of the two factors A and B are not the same across all levels of the other factor. That is, it is possible for A to have a positive effect on the dependent variable at one level of B, no effect at another level and a negative effect at a third level. Such a situation might lead to an overall main effects test for A in which no evidence of any A effect was discovered. This is because the effect of A is confounded with the A by B interaction effect. It is also possible that A has a positive (or negative) effect at each level of B, but that this effect is stronger at some levels of B than at others. In this case it does make sense to talk about an overall positive (or negative) main ef-

fect for factor A, but discussion of the magnitude of this effect must be conditioned on the particular levels of the B factor within which these effects do not differ.

In each of these cases what is called for is to examine the effects of the A factor separately within each level of the B factor. These effects are what are known as simple main effects. Specification of such effects in MANOVA is simple, following a logical algorithm applied to our model specifications on the DESIGN subcommand. The general algorithm is as follows: To obtain the proper simple effects estimates and tests of one factor at (within) each level of a second factor, replace the main effect of the factor of interest and the two-way interaction involving these two factors with the simple effects of the factor of interest within each level of the other factor. For our example, we would replace the main effect of factor A and the A by B interaction with the simple effects of factor A at each level of factor B:

```
MANOVA Y BY A(1,2) B(1,3)
 /DESIGN=B,
         A W B(1), A W B(2), A W B(3)
```

where the W operator is an acceptable shorthand for the WITHIN keyword. Important notes to keep in mind here are the following: We have simply removed the main effect of the A factor and the A by B interaction term from a full factorial specification and have replaced them with a request for the simple effects of A within (separately for) each level of factor B. The effect of this substitution is to repartition the same overall model into different effects, but to maintain the same total model (total degrees of freedom, total sums of squares accounted for, same predicted values and residuals, etc). That is, we are estimating the same B main effect as in the original full factorial model, and repartitioning the A main effect and the A by B interaction effect into the simple main effects of A at each level of B.

This is important to note for two reasons. First, when working with data with unequal numbers of observations in the cells of the design (generally referred to as unbalanced data), the sums of squares for a particular effect such as A W B(1) will generally not be the same when specified alone on the DESIGN subcommand as when specified as part of a larger model, due to the intercorrelation among the factors in an unbalanced design. The algorithm outlined here is designed to maintain the same overall model throughout the testing of simple effects so that the simple effects estimated are logical followups to the results of the overall full factorial analysis. Second, in releases beginning with version 5.0 the default error term in MANOVA has been changed from WITHIN CELLS to WITHIN+RESIDUAL.

Thus even in balanced designs, the error term and degrees of freedom used for testing simple effects would not be the same as in the original analysis unless the same overall model was estimated or unless the user explicitly specified /ERROR=WITHIN on the ERROR subcommand.

So far we have talked only in terms of the simple main effects of A at each level of B. However, the implications of an interaction effect are completely symmetric. That is, to say that the effects of factor A are different at different levels of factor B is equivalent to saying that the effects of factor B are different at different levels of factor A. Thus we would probably also want to test the simple main effects of B at each level of A. To do this we would simply follow the same algorithm, reversing the role of factor A and factor B. That is, we remove the main effect of factor B from the full factorial specification, along with the A by B interaction and substitute the simple main effects of B at each level of A. Our syntax would thus be:

```
MANOVA Y BY A(1,2) B(1,3)
 /DESIGN=A,
        B W A(1), B W A(2).
```

One important point to note is that, since the A and B simple main effects each involve a repartitioning of the interaction term, attempting to fit both sets of simple main effects on one DESIGN subcommand would introduce redundant effects and should thus be avoided. Estimation of both sets of simple main effects in one MANOVA run can be accomplished simply by stacking two DESIGN subcommands:

```
MANOVA Y BY A(1,2) B(1,3)
 /DESIGN=B, A W B(1), A W B(2), A W B(3)
 /DESIGN=A, B W A(1), B W A(2).
```

General between-subjects models: estimating simple main effects

The algorithm outlined above generalizes immediately to cases of higher-order designs. Let's illustrate with the case of a three-way design, with factors A, B and C. For the sake of brevity we will assume that each factor has only two levels, since there is no loss of generality in our discussion and this saves us from writing out more terms in our DESIGN specifications.

If in a higher-order design we wish to estimate simple main effects, the procedure is exactly that outlined above, except that we would have other terms also listed on the DESIGN subcommand. For example, in an A by B by C design in which we wanted to estimate the simple effects of A at each level of B, we would perform the same replacement of main and interaction effect terms as before, but would maintain the model specifications involving the C factor. Thus our full factorial syntax:

```
MANOVA Y BY A B C(1,2)
 /DESIGN=A, B, C, A BY B, A BY C, B BY C,
         A BY B BY C
    would become
MANOVA Y BY A B C(1,2)
 /DESIGN=B, C, A BY C, B BY C, A BY B BY C,
         A W B(1), A W B(2).
```

Since we are using UNIQUE or regression approach sums of squares, the order of effects specified makes no difference, assuming that each cell of the design contains at least one observation (designs involving empty cells are much more complicated and require careful special handling).

Many statisticians might object to the foregoing simple effects tests because they are being conducted in a model in which a higher-order interaction is being estimated which contains the effects in question. The logic behind this objection would be that first we should test the three-way interaction. If this is significant we should then proceed to test simple, simple main effects and/or simple interaction effects. If it is not significant, remove the three-way interaction and re-estimate the model. That is, follow-up tests on simple effects should not be performed until a final model has been chosen. The algorithm outlined here is not affected by this approach. We would have first re-estimated the model without a three-way interaction term, as

```
MANOVA Y BY A B C(1,2)
 /DESIGN=A, B, C, A BY B, A BY C, B BY C
    and the same substitutions would apply, resulting in
MANOVA Y BY A B C(1,2)
 /DESIGN=B, C, A BY C, B BY C,
         A W B(1), A W B(2).
```

Others would consider this approach somewhat rigid. That is, though an interaction effect in a sample was not of sufficient magnitude to provide evidence at (say) the .05 alpha level of an interaction effect in the population, the assumption of no interaction effect as opposed to a small one might be presumptuous. Thus another strategy would be to fit the simple effects in the context of the overall factorial model, estimating them in the presence of the estimated questionable interaction effects. Each user is responsible for coming to her or his own conclusions as to what procedures should be followed in this case; MANOVA can be made to analyze the data in either case.

General between-subjects models: estimating simple, simple main effects and simple interaction effects

If the three-way interaction had been significant in the above model, we would be faced with a more complicated situation. That is, not only do the effects of factor A depend

on which level of factor B we consider, but they also depend on the level of factor C in which our A by B designation of interest is found. The logical step at this point is to examine the two-way interactions at each level of the third factor (such as A by B within each level of C) to see if within each level of the third factor the main effects of the other two factors are invariant. Generalization of the algorithm discussed in the two-way case results in the A by B and A by B by C interactions being replaced by the simple interaction effects of A by B at each level of C. Thus

```
MANOVA Y BY A B C(1,2)
 /DESIGN=A, B, C, A BY B, A BY C, B BY C,
        A BY B BY C
becomes
MANOVA Y BY A B C(1,2)
 /DESIGN=A, B, C, A BY C, B BY C,
        A BY B W C(1), A BY B W C(2).
```

If the simple interaction effects are nonzero, the next step is to estimate the simple, simple main effects of say, factor A at each level of the two-way breakdown of factors B and C. The simple, simple effects of A at each level of factors B and C involve a repartitioning of the A main effect, the A by B, A by C and A by B by C interactions:

```
MANOVA Y BY A B C(1,2)
 /DESIGN=A, B, C, A BY B, A BY C, B BY C,
        A BY B BY C
becomes
MANOVA Y BY A B C(1,2)
 /DESIGN=B, C, B BY C,
        A W B(1) BY C(1), A W B(1) BY C(2),
        A W B(2) BY C(1), A W B(1) BY C(2).
An equivalent specification would be
MANOVA Y BY A B C(1,2)
 /DESIGN=B, C, B BY C,
        A W B(1) W C(1), A W B(1) W C(2),
        A W B(2) W C(1), A W B(1) W C(2).
```

As with the more simple two-way case, the factors here are perfectly symmetric, so we could just as sensibly be using B or C in place of A. Also, the substitution rules used here generalize to designs with any number of factors.

Models involving within-subjects effects

As most users are aware, MANOVA offers the capability of using the multivariate approach to analyzing data involving within-subjects (often involving repeated measures) effects. The within-subjects part of the model is specified separately from the between-subjects part, but in an analogous manner, via the WSDESIGN subcommand. Thus a two-way, completely within-subjects design involving two two-level factors A and B could be specified as:

```
MANOVA V1 TO V4
 /WSFACTORS=A(2) B(2)
          which would be the same as
MANOVA V1 TO V4
 /WSFACTORS=A(2) B(2)
 /WSDESIGN
        or
MANOVA V1 TO V4
 /WSFACTORS=A(2) B(2)
 /WSDESIGN=A, B, A BY B.
```

The estimation of simple effects in completely within-subjects designs requires no new concepts; we simply apply the same rules to the WSDESIGN subcommand that we applied to the DESIGN subcommand. So the simple effects of A at each level of B would be specified as

```
MANOVA V1 TO V4
 /WSFACTORS=A(2) B(2)
 /WSDESIGN=B,
          A W B(1), A W B(2).
```

This is also true for the more complicated three-way and higher-order cases.

Models involving between and within subjects effects

Since the between- and within-subjects parts of the model are specified separately in MANOVA, the case of a design involving both between- and within-subjects factors presents some complications. The specifications for each part of the model are crossed by default. That is, all between-subjects factors are automatically crossed with all within-subjects factors.

Since MANOVA won't allow the specification of between-subjects factors on the WSDESIGN subcommand or within-subjects factors on the DESIGN subcommand, we need a way to tell the procedure that we want to fit the effects of a factor of one type at each level of one or more factors of the other type. Fortunately, there is a method for doing this, and the algorithm involved is generally no more complex than the earlier one, and in many cases it is even simpler.

Take the case of a two-way model involving one between-subjects factor (call it A) and a within-subjects factor (TIME). The standard syntax for the full factorial model is

```
MANOVA V1 V2 BY A(1,2)
 /WSFACTORS=TIME(2)
```

which is equivalent to specifying either TIME on the WSDESIGN or A on the DESIGN subcommand, or both. If we want to estimate the simple effects of time for each level of A, we use the MWITHIN keyword on the DESIGN subcommand, and replace the main effect of A with MWITHIN A(1) and MWITHIN A(2):

```
MANOVA V1 V2 BY A(1,2)
/WSFACTORS=TIME(2)
/DESIGN=MWITHIN A(1), MWITHIN A(2).
```

MWITHIN stands for mean within, and it effectively turns the crossing of A and TIME into the nesting of time within each level of A. This case requires some special caution in reading the output, since what we are thinking of as a simple main effect, TIME at each level of A, is listed on the output as an interaction effect. This analysis produces two tables, the first of which shown in Figure S-1 contains the between-subjects part of the analysis:

What is actually being tested by the MWITHIN A(1) and MWITHIN A(2) terms here are the null hypotheses that the average value across all time points (represented to within a constant multiple by transformed variable T1) is zero within level 1 and level 2 of A, respectively. These are, in general, not hypotheses in which we are usually interested. The hypotheses of common interest are to be found in the within-subjects section of the output, shown below in Figure S-2:

Transformed variable T2 represents a normalized difference variable comparing the two TIME points. Thus a test of MWITHIN A(1) BY TIME represents a test of the null hypothesis that the TIME differences are zero at level 1 of factor A, and MWITHIN A(2) BY TIME corresponds to a similar test at level 2 of factor A.

The substitution rule in this case is even simpler than in cases in which all factors are either between- or within-subjects in nature. That is, all we had to do was to remove the A effect from the DESIGN subcommand and replace it with MWITHIN each level of A. The same rule applies when we want to go the other way, to look at A differences at each TIME point:

```
MANOVA V1 V2 BY A(1,2)
/WSFACTORS=TIME(2)
/WSDESIGN=MWITHIN TIME(1), MWITHIN TIME(2)
```

produces tests of corresponding null hypotheses with the roles of the two factors reversed. However, in this case the tables are presented somewhat differently, as all four hypothesis degrees of freedom in the analysis are defined as within-subjects effects. In each case shown in Figure S-3 we have a constant or intercept term, followed by the term of interest, labeled essentially as an interaction term.

In this case, T1 is simply V1 and T2 is simply V2. That is, the transformation applied to the dependent variables was an identity transformation. Thus the MWITHIN TIME(1) effect tests the null hypothesis that the mean of V1 is zero, averaged across both levels of A, and MWITHIN TIME(2) tests a similar hypothesis concerning V2. As before, these tests of constant or intercept terms are not generally of interest.

The terms labeled as interactions, A BY MWITHIN TIME(1) and A BY MWITHIN TIME(2) are the effects we want, as they test the null hypotheses that there are no population differences between the levels of factor A for V1 and V2, respectively.

The general substitution rule for designs with both between- and within-subjects factors is to substitute MWITHIN each level of a particular main or interaction effect for that factor or interaction and all effects encompassed by that term. So if we wanted to estimate the effects of TIME at each level of the breakdown of an A by B between-subjects design, (the simple, simple main effects of TIME within A by B) we would specify:

Figure S-1

```
* * * * A n a l y s i s   o f   V a r i a n c e – design  1 * * * *
Tests of Between-Subjects Effects.


Tests of Significance for T1 using UNIQUE sums of squares
Source of Variation      SS       DF       MS        F      Sig of F
WITHIN+RESIDUAL         60.64     17       3.57
MWITHIN A(1)           441.80      1      441.80    123.85    .000
MWITHIN A(2)           440.06      1      440.06    123.36    .000
```

Figure S-2

```
* * * * A n a l y s i s   o f   V a r i a n c e – design  1 * * * *
Tests involving 'TIME' Within-Subject Effect.


Tests of Significance for T2 using UNIQUE sums of squares
Source of Variation        SS       DF      MS       F      Sig of F
WITHIN+RESIDUAL           78.64     17      4.63
MWITHIN A(1) BY TIME       9.80      1      9.80     2.12     .164
MWITHIN A(2) BY TIME      20.06      1     20.06     4.34     .053
```

```
MANOVA V1 V2 BY A B(1,2)
 /WSFACTORS=TIME(2)
 /DESIGN=MWITHIN A(1) BY B(1),
          MWITHIN A(1) BY B(2),
          MWITHIN A(2) BY B(1),
          MWITHIN A(2) BY B(2).
```

Thus MWITHIN A BY B specifications replace the A by B interaction and the A and B main effects, which are encompassed within A by B. If we wanted to estimate TIME effects only within the levels of A, we would specify:

```
MANOVA V1 V2 BY A B(1,2)
 /WSFACTORS=TIME(2)
 /DESIGN=B, A BY B,
          MWITHIN A(1), MWITHIN A(2).
```

The same logic applies completely when testing simple effects of between-subjects factors at different levels of within-subjects factors. The substitution algorithm here can also, as in the case involving only between- or within-subjects factors, be extended to as many factors as necessary.

As mentioned earlier, while it is possible in some designs to estimate more than one set of simple effects at a time, it is safest to do them individually, as the results of specifying redundant requests are often meaningless ANOVA tables. This is particularly true with regard to use of the MWITHIN keyword in releases prior to verion 5.0 of SPSS. In later releases only one term can be used with MWITHIN, but in earlier releases use of redundant MWITHIN requests may produce output of questionable validity that some users will not be able to properly interpret.

Finally, there is some disagreement in the ANOVA literature about the use of error terms in designs involving both between- and within-subjects factors. Specifically, it is sometimes claimed to be desirable to use a pooled-error term when fitting A within each TIME point, just as a pooled-error term is used when fitting TIME effects within each A level. However, the simple effects of A at each level of TIME are simply the A effects for the original correlated dependent variables. They are therefore not independent and cannot be pooled to obtain a test statistic with a proper F-distribution under the null hypothesis. Therefore, in this situation MANOVA uses a separate error term at each level of TIME, equivalent to a simple univariate analysis of variance on each dependent variable.

If you have a statistical question or topic that you would like us to address, please fax or mail it to:
   Keywords Editor
   SPSS Inc.
   444 North Michigan Avenue
   Chicago, IL 60611
   Fax: (312) 329-3690

Figure S-3

```
* * * * A n a l y s i s   o f   V a r i a n c e  — design  1 * * * *
Tests involving 'MWITHIN TIME(1)' Within-Subject Effect.

Tests of Significance for T1 using UNIQUE sums of squares
Source of Variation         SS      DF       MS        F      Sig of F
WITHIN+RESIDUAL            65.29     17      3.84
MWITHIN TIME(1)          408.71      1     408.71    106.42    .000
A BY MWITHIN TIME(1)      10.82      1      10.82      2.82    .112

* * * * A n a l y s i s   o f   V a r i a n c e  — design  1 * * * *
Tests involving 'MWITHIN TIME(2)' Within-Subject Effect.

Tests of Significance for T2 using UNIQUE sums of squares
Source of Variation         SS      DF       MS        F      Sig of F
WITHIN+RESIDUAL            74.00     17      4.35
MWITHIN TIME(2)          473.68      1     473.68    108.82    .000
A BY MWITHIN TIME(2)      18.95      1      18.95      4.35    .052
```