

GLM Repeated Measures: One Within, One Between

Reading: SPSS Advanced Models 9.0: 2. Repeated Measures

Homework:

Download: [glm_1w1b.sav](#) ([Download Tips](#))

1. [Overview](#)
2. [The Data](#)
3. [Select the GLM Repeated Measures Procedure](#)
4. [Transformations for the Within-Subjects Factor](#)
5. [The Basic Output](#)
 - [Univariate Tests](#)
 - [Multivariate Tests](#)
6. [Interpreting Significant Effects: Displaying the Means](#)
7. [Interpreting Significant Effects: Tests for Main Effects](#)
8. [Interpreting Significant Interactions: Simple Main Effects](#)
9. [Interpreting Significant Interactions: Tukey's HSD](#)
10. [References](#)

Content

1. Overview

This is a continuation of our discussion of within-subjects designs. In the previous set of notes, [GLM: Repeated-measures designs - One within-subjects factor](#), we discussed how to set up GLM to run a repeated measures design with one within-subjects factor. In this set of notes we extend the model to include a between-subjects factor.

top 

2. The Data

The previous set of notes described a treatment outcome study that used Eye Movement Desensitization and Reprocessing (EMDR) as a treatment for psychological trauma (Wilson, Becker, & Tinker, 1995, 1997). In this set of notes we extend the analysis of that study by comparing participants who were randomly assigned to the immediate treatment condition with those randomly assigned to the delayed treatment condition. Assignment to treatment conditions becomes the between-subjects factor in the design.

The design is shown in Table 1. Participants in the immediate treatment condition received three 90-minute EMDR sessions between the two assessments. Participants in the delayed treatment condition received no treatment between the two assessments. The delayed treatment condition acts as control for history and regression to the mean artifacts. For example, perhaps people responded to the ads for the study because they were feeling particularly disturbed by their trauma. And maybe they would naturally feel better with the mere passage of time. The delayed treatment condition controls for the possibility of that "regression to the mean" threat to validity of the study. (Participants in the delayed treatment condition received the EMDR treatment following the postdelay assessment. Those results are not analyzed in this set of notes.)

Table 1. Research Design Comparing Immediate Treatment with Delayed Treatment

	Pretreatment		Postdelay/ Posttreatment
Delayed Treatment Condition	Pretreatment Assessment (<i>n</i> = 40)	30-day delay	Postdelay Assessment (~30 days after T1) (<i>n</i> = 40)
Immediate Treatment Condition	Pretreatment Assessment (<i>n</i> = 40)	Three 90-minute EMDR Treatment Sessions	Posttreatment Assessment (~30 days after T1) (<i>n</i> = 40)

The data are stored in the file glm_1w1b.sav. The variables in that file are shown in Table 2. The dependent measure is the Avoidance score of the Impact of Events scale (IES; Horowitz, Wilner, & Alvarez, 1979). Avoidance symptoms include trying not to think about the traumatic event and staying away from situations and people that remind you of the event.

Table 2. The variables in glm_1w1b.sav

Variable Name	Variable Label / Value Label
txcond	Treatment condition 1 = immediate treatment, 2 = delayed treatment
ies_a_pre	IES Avoidance, Pretreatment
ies_a_pos	IES Avoidance, Posttreatment or Postdelay

top 

3. Select the GLM Repeated Measures Procedure

The GLM Repeated Measures dialog box is opened by clicking

Analyze

General Linear Model

Repeated Measures...

GLM - Repeated Measures Define Factor(s)

The within-subject factors are defined in the opening window. Lets call the within-subjects factor in this example *prepost*. The prepost factor has two levels. **Add** prepost(2) to the within-subject factor window.

Lets also name the measure this time so that the output will not refer to the measure as *measure_1*. Lets name the measure *ies_avd*, the avoidance score of the IES. You do not need to name the measure, it just makes the output easier to read when we see *ies_avd* rather than *measure_1*.

GLM - Repeated Measures

The next step is to define each of the two levels of the time within-subjects variable. Move each of the variables to the **Within-Subjects Variables (time):** window. Make sure that the conceptual order of the original variables is preserved when you move them to the within-subjects variables window, that is, *ies_a_pre* should be moved to the ___?__(1,**ies_avd**) space, *ies_a_pos* should be moved to the ___?__(2,**ies_avd**) space.

There is a between-subjects factor in this design so we need to move that factor, *txcond*, to the **Between-Subjects Factor(s):** window.



4. Transformations for the Within-Subjects Factor

There is one degree of freedom for the *prepost* within-subjects factor. GLM will create two transformed scores, one for the average of the repeated measures variables, *iesa_pre* & *iesa_pos*, and one for the *prepost* main effect.

The coefficients used for the creating the transformed variables can be displayed by clicking the **Display...Transformation matrix** option in the Options... dialog box.

The coefficients for creating the new, temporary, transformed variable **average** are shown in Table 3. The coefficients for each of the variables, *iesi_pre* and *iesi_pos* are .707.

The coefficients are normalized, the sums of the squares of the coefficients is equal to 1.00.

The between-subjects effect, treatment condition, will be analyzed using this transformed score.

The coefficients for the prepost main effect are -.707 and +.707. These coefficients are used to create a new, temporary variable called **prepost_linear**. This transformed variable provides the test of the prepost within-subjects effect. The contrast is the difference between the pre and post scores.

If you only have two levels for a within-subjects factor, these will always be the coefficients for that factor.

The coefficients sum to zero. If there is no difference between the means the transformed score will be zero.

The coefficients are normalized. The sum of the squares

Table 3. Average
Measure: IES_AVD
Transformed Variable: AVERAGE

IES Avoidance, Pretreatment	.707
IES Avoidance, Posttreatment/Postdelay	.707

Table 4. PREPOST(a)
Measure: IES_AVD

Dependent Variable	Prepost Linear
IES Avoidance, Pretreatment	-.707
IES Avoidance, Posttreatment/Postdelay	.707
a The contrasts for the within subjects factors are: PREPOST: Polynomial contrast	

of the coefficients equal 1.00.

top 

5. The Basic Output

The basic output is similar as for a repeated measures design with one within-subject factor. The following tables are included in the output: Within Subjects Factors, see Table 5; Between Subjects Factors, see Table 6; Mauchly's Test of Sphericity, see Table 7; Tests of Within Subjects Effects, see Table 8; Tests of Within-Subjects Contrasts, see Table 8, and Tests of Between-Subjects Effects, see Table 9, and Multivariate Tests, see Table 10.

Within- and Between Subject Factor Definition

Table 5. Within-Subjects Factors

Measure: IES_AVD

PREPOST	Dependent Variable
1	IESA_PRE
2	IESA_POS

Table 5 shows the two levels of the within-subjects factor called *prepost*. The dependent variables that make up the two levels of *prepost* are the pre-and post measures of the IES Avoidance scale, *iesa_pre* and *iesa_pos*.

Table 6. Between-Subjects Factors

		Value Label	N
TREATMENT GROUP	1	immediate treatment	39
	2	delayed treatment	37

Table 6 shows the two levels for the between-subjects factor *txcond*, treatment group, along with the *ns* for the two cells.

Univariate Tests

Univariate Assumptions

As was noted in the previous set of notes for repeated measures, the within-subject effect is analyzed in GLM by first transforming the original variables into single degree of freedom tests of the null hypothesis. In this example the prepost within-subject effect has a single degree of freedom so there is only one single degree of freedom estimate of the prepost effect.

Because GLM does not have to average multiple single degree of freedom tests to get an overall test of the effect, the assumptions about the correlations between the transformed variables being zero and the variances of the transformed variables being equal are not pertinent. The Mauchly sphericity test will still be printed, see Table 7, but the Mauchly W statistic will be 1.00, the df will be 0 and the significance level will not be printed. In addition the epsilons will all be 1.00, meaning that the corrected df are the same as the original df.

Table 7. Mauchly's Test of Sphericity(b)
Measure: IES_AVD

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon(a)		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
PREPOST	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.

b Design: Intercept+TXCOND
Within Subjects Design: PREPOST

Univariate ANOVA Results

Univariate tests of the *time* within-subject effect are shown in Table 8. The **sphericity assumed** row indicates that the statistics in that row have not been adjusted by any of the epsilons. It is appropriate in this example because with 1 df tests the sphericity assumption is not relevant.

Table 8. Tests of Within-Subjects Effects

Measure: IES_AVD

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
PREPOST	Sphericity Assumed	2556.977	1	2556.977	73.110	.000	.497
	Greenhouse-Geisser	2556.977	1.000	2556.977	73.110	.000	.497
	Huynh-Feldt	2556.977	1.000	2556.977	73.110	.000	.497
	Lower-bound	2556.977	1.000	2556.977	73.110	.000	.497
PREPOST * TXCOND	Sphericity Assumed	1449.740	1	1449.740	41.452	.000	.359
	Greenhouse-Geisser	1449.740	1.000	1449.740	41.452	.000	.359
	Huynh-Feldt	1449.740	1.000	1449.740	41.452	.000	.359
	Lower-bound	1449.740	1.000	1449.740	41.452	.000	.359
Error (PREPOST)	Sphericity Assumed	2588.102	74	34.974			
	Greenhouse-Geisser	2588.102	74.000	34.974			
	Huynh-Feldt	2588.102	74.000	34.974			
	Lower-bound	2588.102	74.000	34.974			

Tests of individual contrasts are shown in Table 9. The default contrasts are orthogonal polynomial (linear, quadratic, cubic, etc.) contrasts. Because there is only 1 df for these effects only one contrast is shown, the linear effect. These linear effects are the same effects that are shown in the "Tests of Within-Subject Effects" table. There are no corrections for sphericity in the tests of within-subjects contrasts because each contrast is a single degree of freedom test.

Table 9. Tests of Within-Subjects Contrasts

Measure: IES_AVD

Source	PREPOST	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
PREPOST	Linear	2556.977	1	2556.977	73.110	.000	.497

PREPOST * TXCOND	Linear	1449.740	1	1449.740	41.452	.000	.359
Error (PREPOST)	Linear	2588.102	74	34.974			

The results of the tests of the within-subject effects indicate significant effects for both the prepost main effect, $F(1, 74) = 73.11, p < .0005$, and the prepost by treatment condition interaction, $F(1, 74) = 41.45, p < .001$. For this study, the interaction was critical. It was hypothesized that avoidance symptoms would decrease from pretreatment to posttreatment for the participants in the immediate treatment condition and that there would no change for participants in the delayed treatment condition because they have not yet received treatment.

Univariate tests of the main effect of treatment condition (a between-subjects effect) are shown in Table 10. Recall that the between-subjects effects are tested by the transformed variable called *average*.

Table 10. Tests of Between-Subjects Effects

Measure: IES_AVD

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Intercept	35439.889	1	35439.889	278.554	.000	.790
TXCOND	315.600	1	315.600	2.481	.120	.032
Error	9414.874	74	127.228			

The **intercept** tests whether the transformed variable *average* is different from zero. It is not of interest to us.

The main effect for treatment condition is not significant, $F(1, 74) = 2.48, p = .12$.

Multivariate Tests

Multivariate tests of the within-subject effects are also displayed by the GLM procedure, see Table 11. When the within-subjects tests are single degree of freedom tests, as in this example, then the multivariate *F*s will be identical to the univariate *F*s.

Table 11. Multivariate Tests^b

Effect		Value	F	Hypothesis df	Error df	Sig.	Eta Squared
PREPOST	Pillai's Trace	.497	73.110 ^a	1.000	74.000	.000	.497
	Wilks' Lambda	.503	73.110 ^a	1.000	74.000	.000	.497
	Hotelling's Trace	.988	73.110 ^a	1.000	74.000	.000	.497
	Roy's Largest Root	.988	73.110 ^a	1.000	74.000	.000	.497
PREPOST * TXCOND	Pillai's Trace	.359	41.452 ^a	1.000	74.000	.000	.359
	Wilks' Lambda	.641	41.452 ^a	1.000	74.000	.000	.359
	Hotelling's Trace	.560	41.452 ^a	1.000	74.000	.000	.359
	Roy's Largest Root	.560	41.452 ^a	1.000	74.000	.000	.359

a Exact statistic
b Design: Intercept+TXCOND Within Subjects Design: PREPOST

top 

6. Interpreting Significant Effects: Displaying the Means

There are three different ways to display the means:

- (a) display the descriptive statistics;
- (b) display the estimated marginal means, and
- (c) display profile plots or plots with error bars.

Descriptive Statistics

Table 12. Descriptive Statistics

	TREATMENT GROUP	Mean	Std. Deviation	N
IES Avoidance, Pretreatment	immediate treatment	21.03	9.64	39
	delayed treatment	17.73	8.50	37
	Total	19.42	9.19	76
IES Avoidance, Posttreatment/ Postdelay	immediate treatment	6.64	9.25	39
	delayed treatment	15.70	8.53	37
	Total	11.05	9.95	76

Descriptive statistics are displayed by checking the **Display Descriptives** box in the **Options...** dialog box.

The descriptive statistics for the two variables (iesa_pre, iesa_pos) are shown in Table 12.

The means in the rows labeled "Total" are **weighted** means.

Estimated Marginal Means

Estimated marginal means are displayed from the **Estimated Marginal Means** window of the **Options...** dialog box. To display the means for the significant interaction move the *txcond*prepost* interaction from the **Factor(s) and Factor Interactions** window to the **Display Means for:** window and press **Continue**. The main effect means could also be displayed, but if there is a significant interaction then main effects are not normally discussed or interpreted.

The estimated marginal means for the interaction are shown in Table 13. The estimated marginal means are **unweighted** means.

Inspection of the 95% confidence intervals indicates that there is a significant decrease in

Table 13. Estimates
Measure: IES_AVD

		Mean	Std. Error	95% Confidence Interval	
TREATMENT GROUP	PREPOST			Lower Bound	Upper Bound
immediate treatment	1	21.026	1.458	18.121	23.930
	2	6.641	1.426	3.799	9.483
delayed treatment	1	17.730	1.497	14.748	20.712
	2	15.703	1.464	12.785	18.620

scores from pretest to posttest for the immediate treatment group, but no change for the delayed treatment group.

Profile Plots

To display a profile plot go to the **Plots...** dialog box. The dependent variable is always on the y-axis. In a repeated measures design the within-subjects factor is typically placed on the horizontal axis, move *prepost* to the **horizontal axis** window. Then move the between subjects factor, *txcond*, to the **separate lines** window

The resulting profile plot is shown in Figure 1.

The plots show a large decrease in the IES avoidance scores for the immediate treatment group from pretreatment to posttreatment. The change for the delayed treatment group from pretreatment to postdelay is small.

A more useful plot would include confidence intervals as error bars, see Figure 2. The pre- and posttest scores for the immediate treatment group are shown on the left and those for the delayed (untreated) group are shown in the right.

The plot was created from the normal error bar chart (not the interactive error bar chart). The following options were chosen at the opening dialog box:
Clustered and **Data in Chart Area**

Figure 1. Estimated Marginal Means of IES_AVD

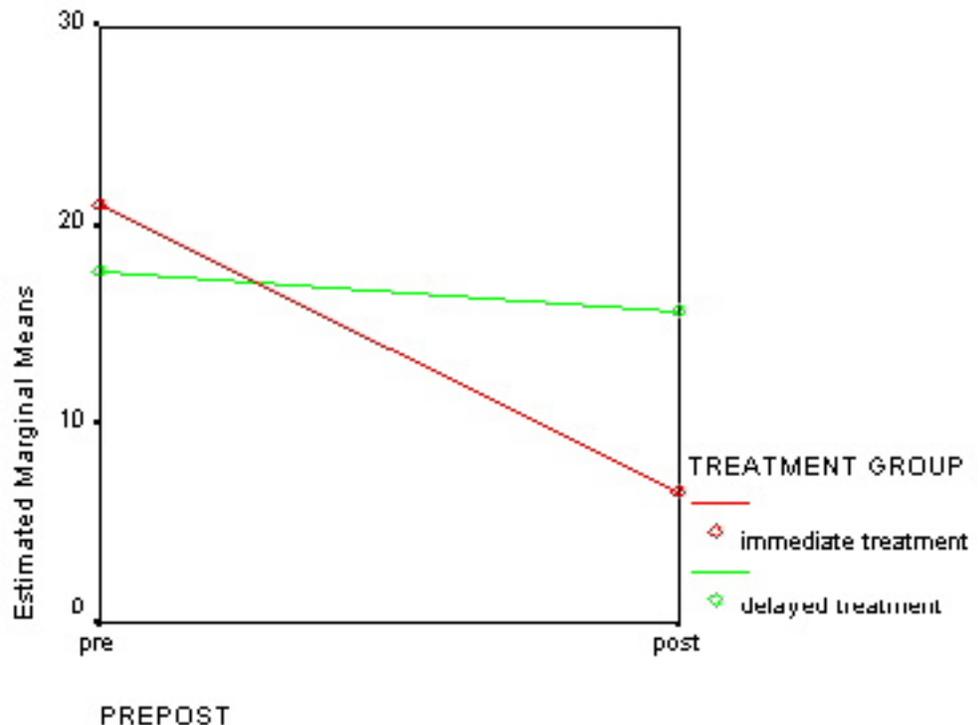
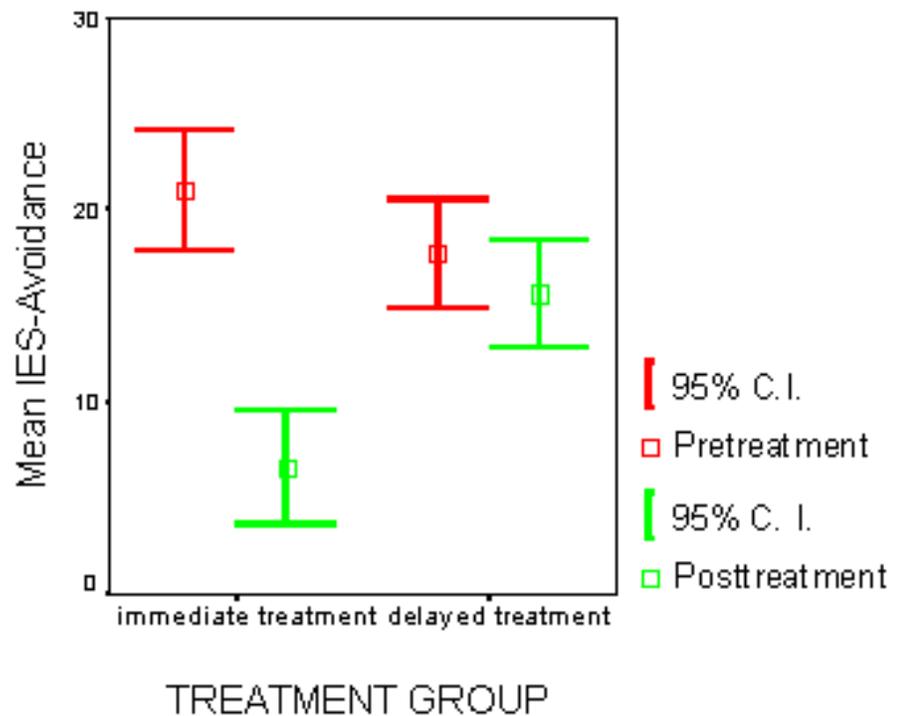


Figure 2. Plots with error bars (95% C.I.)

Summaries of separate variables. The IES pretest and posttest scores, in that order, were selected as the **Variables**: and treatment group was selected as the **Category Axis**. The chart was edited to: increase the size of the text and the width of the error bar lines, have a more descriptive label for the y-axis, and a more descriptive legend titles.

In have been unable to get a similar plot from the interactive error bar graphics program without creating two separate plots. That interactive graphics program does not have the Clustered option.



Inspection of this error bar plot indicates that there is a significant decrease in scores from pretest to posttest for the immediate treatment group, but no change for the delayed treatment group. The error bars represent the 95% confidence interval based on the standard error of each individual mean. Hence the error bars are appropriate for between-subject comparisons (e.g., comparisons between the two treatment groups). They are too conservative for within-subject comparisons (e.g., pretest vs. posttest within a treatment group).

top 

7. Interpreting Significant Effects: Tests for Main Effects

If there are no significant interactions, then it is appropriate to interpret any significant main effects. Main effect tests are available from the **Compare main effects** option in the **Estimated Marginal Means** section of the **Options...** dialog box and from the **Post Hoc** dialog box.

Estimated Marginal Means

The **Compare main effects** option provides pairwise main-effect tests for either within- or between-subjects factors. You can select uncorrected (LSD) confidence intervals for the pairwise comparisons or you could correct the confidence intervals using either the Bonferroni or Sidak adjustments.

The estimated marginal means are the unweighted means. These are the appropriate means for designs with either equal *ns* or unequal *ns*.

In other words, the tests provided by the **Compare main effects** option is a general solution to making pairwise comparisons that is appropriate for every situation.

Post Hoc Dialog Box

The **Post Hoc** dialog box provides a wide array of post hoc tests for between subjects main effects. It will not provide post hoc tests for within-subjects factors.

In our example you could run post-hoc tests on the *txcond* between-subjects main effect.

These post hoc tests use the **unweighted** means so the tests are not appropriate if you have more than one between subjects factor and you have unequal *ns*.

So when is it appropriate to use these tests?

Table 14. When are the Post Hoc... Tests Appropriate?

1. When there is only one between subjects factor.
2. When there are more than one between-subjects factors and you have equal *ns* in each cell of the design.

As a general rule of thumb I would stick with the more general set of comparisons that are available from the **Compare main effects** option within the **Estimated Marginal Means** section of the **Options...** dialog box.

8. Interpreting Significant Interactions: Simple Main Effects

The hypotheses for this study were that EMDR treatment would result in a significant decrease in PTSD symptoms, and that there would be no change in PTSD symptoms for participants in the delayed treatment condition prior to treatment. This pair of hypotheses imply that the interaction would be significant, which it is. But the hypotheses state a specific form of the interaction. These two specific

hypotheses can be tested as simple main effects the change from pretreatment to posttreatment within each treatment condition.

Simple main effects for a mixed design such as this (i.e., a design with both within- and between-subjects factors) are computed in the same way as for a between-subjects design (see the notes for [GLM: Simple Main Effects](#)). Go to the Estimated Marginal Means section of the **Options...** dialog box. Make sure that the interaction term is displayed in the **Display Means for:** window and press **Continue**. The next step is to edit the syntax for the **/EMMEANS** subcommand. To do that press the **Paste** button in the GLM - Repeated Measures dialog box, the Syntax Editor window will open.

Find the **/EMMEANS** subcommand in the syntax editor window It should look like the **/EMMEANS** command at the right. The **/EMMEANS** command will display the means for the *txcond* by *prepost* interaction.

```
/EMMEANS =
TABLES
(txcond*prepost)
```

Add "COMPARE (prepost) ADJ(sidak)" to the interaction **/EMMEANS** subcommand, as shown at the right. The **COMPARE (txcond)** syntax will display the simple prepost main effects within each treatment condition.

```
/EMMEANS =
TABLES
(txcond*prepost)
COMPARE
(prepost) ADJ
(SIDAK)
```

Then **Run** the syntax commands.

Two types of output are displayed, the ANOVA for the two simple main effects, see Table 15, and pairwise comparisons, see Table 16. The ANOVA table presents a set of multivariate tests of the simple main effects. Normally people would report the *F* value rather than any of the multivariate test values. In this example the IES avoidance score at posttreatment ($M = 6.64$, $SE = 1.426$) was significantly lower than the pretreatment score ($M = 21.03$, $SE = 1.458$), $F(1, 74) = 115.37$, $p < .0005$. There was no difference between the postdelay ($M = 15.70$, $SE = 1.464$) and predelay scores ($M = 17.73$, $SE = 1.497$) for participants who had not received EMDR treatment, $F(1,74) = 2.17$, $p = .145$.

Table 15. Multivariate Tests

TREATMENT GROUP		Value	F	Hypothesis df	Error df	Sig.	Eta Squared
immediate treatment	Pillai's trace	.609	115.367 (b)	1.000	74.000	.000	.609
	Wilks' lambda	.391	115.367 (b)	1.000	74.000	.000	.609
	Hotelling's trace	1.559	115.367 (b)	1.000	74.000	.000	.609
	Roy's largest root	1.559	115.367 (b)	1.000	74.000	.000	.609
delayed treatment	Pillai's trace	.029	2.173(b)	1.000	74.000	.145	.029
	Wilks' lambda	.971	2.173(b)	1.000	74.000	.145	.029
	Hotelling's trace	.029	2.173(b)	1.000	74.000	.145	.029
	Roy's largest root	.029	2.173(b)	1.000	74.000	.145	.029

Each F tests the multivariate simple effects of PREPOST within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a Computed using alpha = .05

b Exact statistic

You come to the same conclusion by examining the pairwise comparisons for the simple main effects. Even the significance levels are the same because only two means are being compared for each simple main effect.

Table 16. Pairwise Comparisons
Measure: IES_AV D

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
TREATMENT GROUP	(I) PREPOST	(J) PREPOST				Lower Bound	Upper Bound
immediate treatment	1	2	14.38(*)	1.339	.000	11.716	17.053
	2	1	-14.38(*)	1.339	.000	17.053	-11.716
delayed treatment	1	2	2.03	1.375	.145	-.713	4.767
	2	1	-2.03	1.375	.145	4.767	-.713

Based on estimated marginal means

* The mean difference is significant at the .05 level.

Statistical Note: Simple main effects for 2-way interactions in a two-way, mixed design always compare either (a) between-subject means (e.g. simple main effects of treatment group at pretreatment and at posttreatment) or (b) within-subject means (e.g., simple main effects of pre- and posttreatment measures within each treatment condition). When between subject means are compared, the comparisons are between subject *t* tests. The comparisons are made on the individual means using the standard errors of each mean. When within subject means are compared, the comparisons are paired *t* tests. The the comparisons are made on the difference scores using the standard error of the difference score.

Writing up the Results of the Simple Main Effects Analysis

The description of the results would begin by describing the overall analysis of variance. The description should include the dependent variable(s) analyzed and each of the factors in the design. When the design is mixed then each factor should be described as either within-subjects or between subjects. followed by the significant effects from that analysis. The first time you use the term "analysis of variance" you should write out the whole phrase and put the abbreviation (ANOVA) in parentheses. Subsequently you should just use the abbreviation, ANOVA.

The IES avoidance scores were analyzed in a 2 x 2 analysis of variance (ANOVA) with prepost (pretreatment vs. posttreatment) as a within-subject factor and treatment group (EMDR vs. delayed treatment) as a between subjects factor.

Next the significant effects from that analysis would be described.

The IES avoidance scores were analyzed in a 2 x 2 analysis of variance with prepost (pretreatment vs. posttreatment) as a within-subject factor and treatment group (EMDR vs. delayed treatment) as a between subjects factor. **The ANOVA yielded a significant pre-post main effect, $F(1, 74) = 73.11, p < .0005$, and a significant interaction between the pre-post measures and the treatment condition, $F(1, 74) = 41.45, p < .0005$. The pre-post main effect needs to be interpreted in light of the significant interaction and will not be described further.**

Finally, you would describe the simple main effects analysis and the results from that analysis. Write out your description of the simple main effects analysis in the box that follows and then press SUBMIT. Please press the Enter key at the end of each line of text. We will discuss your (anonymous) contributions in class.

Click here: [View Descriptions](#) to read your descriptions of the simple main effects analysis.

top 

9. Interpreting Significant Interactions: Tukey's HSD

Pairwise comparisons for all between subjects effects or all within subjects effects

Tukey's HSD test could be used to make all pairwise comparisons for significant main effects.

$$\psi(HSD) = q_{\alpha, p, v} \sqrt{\frac{MS_{error}}{n}}$$

The accepted rule is use the MS_{error} term that was used to test the main effect.

If you want to use the HSD to make all pairwise comparisons in a 2-way interaction where the factors are both within subjects or both between subjects, then use the MS_{error} that was used to test the interaction.

Pairwise comparisons for mixed between- and within-subjects interactions

If you want to use the HSD to make all pairwise comparisons in a 2-way interaction where one of the factors is a between subjects and the other factor is within subjects then things become more complicated. One might expect that the MS_{error} that was used to test the interaction would be appropriate. Dr. Klebe recommends using that approach. Here are the values of the components of the HSD formula --

MS_{error} = 34.974,
 v (df for the appropriate error term) = 74,
 α = .05,
 p (the number of means to be compared) = 4,
 n = the harmonic mean of the $n = 37.97$.

$$N_h = 4 / (1/39 + 1/37 + 1/39 + 1/37) = 37.97$$

q (from the Studentized Range Statistic table) with $\alpha = .05$, p (the number of means) = 4, and v (df for error) = 74 is 3.39.

The critical value for Tukey's HSD using the MS_{error} for the within subject effect is --

$$HSD = 3.39 * \text{SQRT}(34.974/37.97) = 3.25$$

If the difference between any pair of means is greater than or equal to 3.25 they are significantly different using this Tukey's HSD procedure.

However that is not always the recommended procedure. Shavelson (1996) recommends using a pooled MS_{error}. Pooling is accomplished by adding the Sums of Squares for the two error terms and dividing by the sum of the degrees of freedom for the two error terms --

$$MS_{PW} = (SS_{\text{error.between}} + SS_{\text{error.within}}) / (df_{\text{error.between}} + df_{\text{error.within}})$$

In this example --

$$\begin{aligned} MS_{PW} &= (2588.102 + 9414.874) / (74 + 74) \\ &= 12,002.976 / 148 \\ &= 81.101 \end{aligned}$$

The harmonic mean of the $n_s = 37.97$

Look up q with $\alpha = .05$, p (the number of means) = 4, and v (df for error) = $74 + 74 = 148$ in the Studentized Range table --

$$q = 3.68$$

Putting it all together the critical value for Tukey's HSD is --

$$HSD = 3.68 * \text{SQRT}(81.101/37.97) = 5.37.$$

If the difference between any pair of means is greater than or equal to 5.37 they are significantly different using Tukey's HSD procedure. The pooling approach typically is much more conservative than the using the MSerror from the within subjects error term.

I suggest using a simple main effects analysis rather than making all pairwise comparisons using the Tukey's HSD.

top 

10. References

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of event scale: A measure of subjective distress. *Psychosomatic Medicine*, *41*, 209-218.

Shavelson, R. J. (1996). *Statistical reasoning for the behavior sciences* (3rd ed.). Boston: Allyn and Bacon.

Wilson, S. A., Becker, L. A., & Tinker, R. A. (1995). Eye movement desensitization and reprocessing

(EMDR) treatment for psychologically traumatized individuals. *Journal of Consulting and Clinical Psychology*, 63, 928-937.

Wilson, S. A., Becker, L. A., & Tinker, R. A. (1997). Fifteen-month follow-up of eye movement desensitization and reprocessing (EMDR) treatment for posttraumatic stress disorder and psychological stress. *Journal on Consulting and Clinical Psychology*, 65, 1047-1056.

top 

©Lee A. Becker, 1997-1999 -revised 12/01/99