# Complete Example (Repeated Measures With Interactions)

New to this section:

* While it is very tempting to just use it as one IV, you would be ignoring the interaction component of the two IVs. So, make sure you treat them as two separate variables in JASP.
* Main effects: interpreting each IV on its own, ignoring the effect of the other IVs.
* Interaction: interpreting the IVs together, seeing if the conditions are significantly different OR if the pattern of data across levels is different for the other IV.

Chart of ANOVA Analysis:

|  |  |  |  |
| --- | --- | --- | --- |
|  | ANOVA | | |
|  | Main Effect | Main Effect | Interaction |
| If levels > 2  And significant | Dependent t-test  Bonferroni/Tukey correction | Dependent t-test  Bonferroni/Tukey correction | Dependent t-test  Bonferroni/Tukey correction  (split ~ by looking at the columns) |
| If levels = 2 | Interpret means | Interpret means |

If the interaction is significant, often people ignore any analyses with the main effects:

* This procedure reduces Type 1 error because you are running less post hoc tests.
* You are interested in the interaction anyway, so why only interpret one variable at a time?

In this experiment people were given word pairs to rate based on their “relatedness”. How many people out of a 100 would put LOST-FOUND together?

**Datafile:** rm 2 anova.csv

**IVs:**

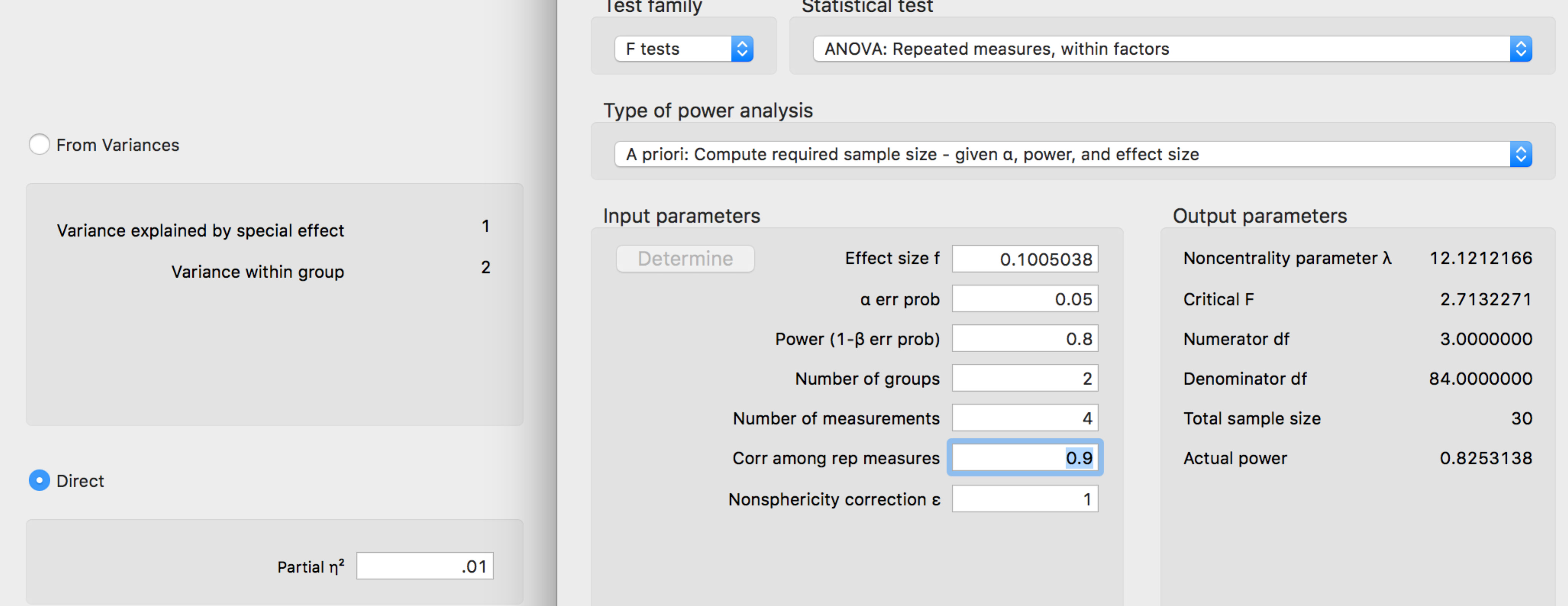
* FSG – how strong the relationship was between LOST-FOUND
  + FSG – two levels (low versus high)
* BSG – how strong the relationship was between FOUND-LOST
  + BSG – two levels (low versus high)
* Creates 4 conditions (lo lo, lo hi, hi lo, hi hi).
* Everybody got all four types of word pairs.

**DV:**

* Rating of each word pair type.

**Power:**

1. Open Gpower!
   1. Test family: F-test
   2. Statistical Test: ANOVA repeated measures, within factors
   3. Estimate an effect size: click determine 🡪 click direct 🡪 use eta square sizes you think might be accurate, remember small, medium, and large estimates from the notes.
   4. Alpha = .05
   5. Power (1-beta .20) = .80
   6. Number of groups = number of IVs
   7. Number of measurements = number of levels or conditions
   8. Corr among rep measures = correlation between levels or conditions
      1. You can estimate from previous research.
      2. Look at the correlations in a pilot study, go with the lowest one you find.
      3. .5-.7 is a good estimate if you are giving them the same test a couple times.
   9. Nonsphericity correction = epsilon
      1. You will not really know this number before you start a study. More useful if you have some participants to estimate from (see below on how to get that number).
2. Let’s estimate the following:
   1. Small effect size
   2. Two IVs
   3. Conditions from our current study
   4. Correlation = .9
   5. Epsilon = 1
3. Says we needed to run 30 people to find a significant effect with a small effect size.



**Example write up:**

**Results**

Participants were given pairs of words and asked to rate them on how often they thought 100 people would give the second word if shown the first word. The strength of the word pairs was manipulated through the actual rating (forward strength: FSG) and the strength of the reverse rating (backward strength: BSG). Data was screened for assumptions and outliers. One outlier was found using *z*-score distance as a criterion and was excluded in the analysis.

Figure 1 shows the mean ratings for each combination of low and high forward and backward strength. FSG condition had a significant effect on ratings, *F*(1, 155) = 980.70, *p* < .001, *ηp²* = .86, which indicated that participants could correctly rate low related pairs (*M* = 53.17, *SD* = 12.95) lower than high related pairs (*M* = 71.14, *SD* = 13.96). However, BSG condition had a significant effect on ratings, *F*(1, 155) = 372.99*, p* < .001*, ηp²* = .71. Participants again were able to rate low related pairs (*M* = 57.27, *SD* = 14.42) lower than high related pairs (*M* = 67.04, *SD* = 16.39). Finally, the interaction between FSG and BSG was significant, *F*(1, 155) = 71.58, *p* < .001, *ηp²* = .32. Dependent *t*-tests were used for the post hoc analysis of the interaction between FSG and BSG with a Bonferroni correction. BSG had an approximately six point effect on ratings for low FSG pairs, *t*(155) = 11.93, *p*<.001, *ddiff* = -0.96. However, for high related FSG pairs, BSG increase from low to high also increased ratings about 18 points, *t*(155) = 16.12, *p*<.001, *ddiff* = -1.29. See Figure 1 for the interaction.

(note: I got the M and SD for the main effect from descriptives – be sure you have the F values as the most important thing).

*Figure 1.* The interaction between forward and backward strength in estimating the relationship between word pairs. Error bars denote one standard deviation around the mean.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | BSG | |
|  |  | Low | High |
| FSG | Low | Low Low | Low High |
| High | High Low | High High |

How to split:

* Go across or down.
* Go with your hypothesis.
* Split on the high number of levels, which will give you the smaller number of comparisons (square design doesn’t matter).

**Results**

**Repeated Measures ANOVA**

| **Within Subjects Effects** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Sum of Squares** | | **df** | | **Mean Square** | | **F** | | **p** | | **η²** | |
| FSG |  | 51270 |  | 1 |  | 51269.86 |  | 980.70 |  | < .001 |  | 0.864 |  |
| Residual |  | 8103 |  | 155 |  | 52.28 |  |  |  |  |  |  |  |
| BSG |  | 14573 |  | 1 |  | 14572.50 |  | 372.99 |  | < .001 |  | 0.706 |  |
| Residual |  | 6056 |  | 155 |  | 39.07 |  |  |  |  |  |  |  |
| FSG ✻ BSG |  | 2475 |  | 1 |  | 2475.49 |  | 71.58 |  | < .001 |  | 0.316 |  |
| Residual |  | 5360 |  | 155 |  | 34.58 |  |  |  |  |  |  |  |
|  | | | | | | | | | | | | | |
| *Note.*  Type III Sum of Squares | | | | | | | | | | | | | |

FSG, *F*(1, 155) = 980.70, *p* < .001, *ηp²* = .86

BSG, *F*(1, 155) = 372.99, *p* < .001, *ηp²* = .71

Interaction, *F*(1, 155) = 71.58, *p* < .001, *ηp²* = .32

| **Between Subjects Effects** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Sum of Squares** | | **df** | | **Mean Square** | | **F** | | **p** | | **η²** | |
| Residual |  | 8103 |  | 155 |  | 52.28 |  |  |  |  |  |  |  |
|  | | | | | | | | | | | | | |
| *Note.*  Type III Sum of Squares | | | | | | | | | | | | | |

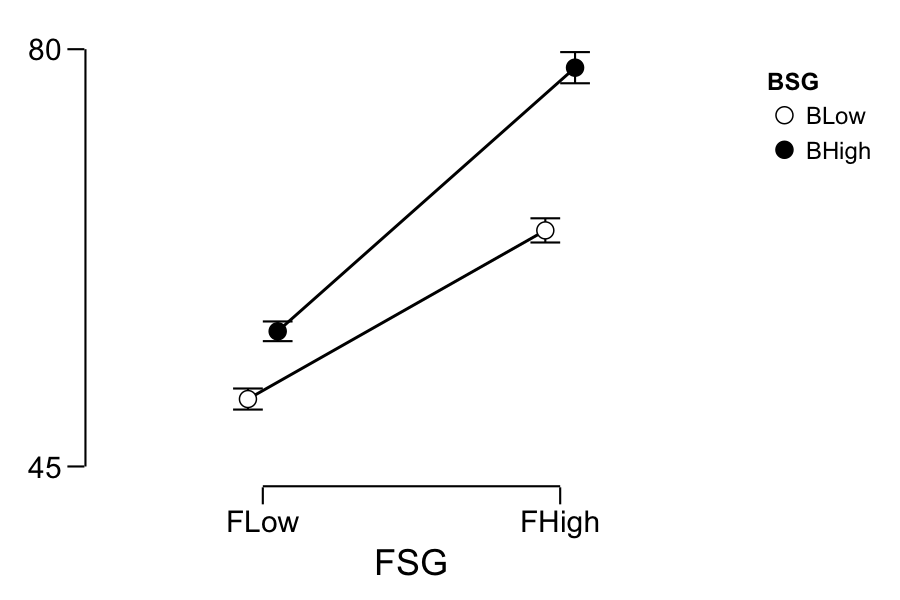
**Assumption Checks**

| **Test of Sphericity** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Mauchly's W** | | **p** | | **Greenhouse-Geisser ε** | | **Huynh-Feldt ε** | |
| FSG |  | 1.000 | ᵃ | NaN | ᵃ | 1.000 | ᵃ | 1.000 | ᵃ |
| BSG |  | 1.000 | ᵃ | NaN | ᵃ | 1.000 | ᵃ | 1.000 | ᵃ |
| FSG ✻ BSG |  | 1.000 | ᵃ | NaN | ᵃ | 1.000 | ᵃ | 1.000 | ᵃ |
|  | | | | | | | | | |
| ᵃ The repeated measure has only two levels. When the repeated measure has two levels, the assumption of sphericity is always met. | | | | | | | | | |

**Descriptives**

| **Descriptives** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **FSG** | | **BSG** | | **Mean** | | **SD** | | **N** | |
| FLow |  | BLow |  | 50.66 |  | 12.30 |  | 156 |  |
|  |  | BHigh |  | 56.34 |  | 12.29 |  | 156 |  |
| FHigh |  | BLow |  | 64.80 |  | 11.88 |  | 156 |  |
|  |  | BHigh |  | 78.45 |  | 10.89 |  | 156 |  |
|  | | | | | | | | | |

**Descriptives Plot**



**T-Test**

| **Paired Samples T-Test** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | |  | | **t** | | **df** | | **p** | | **Cohen's d** | |
| fsglobsglo |  | - |  | fsglobsghi |  | -11.93 |  | 155 |  | < .001 |  | -0.955 |  |
| fsghibsglo |  | - |  | fsghibsghi |  | -16.12 |  | 155 |  | < .001 |  | -1.290 |  |
|  | | | | | | | | | | | | | |
| *Note.*  Student's T-Test. | | | | | | | | | | | | | |

| **Descriptives** | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **FSG** | | **BSG** | | **Mean** | | **SD** | | **N** |
| FLow |  | BLow |  | 50.66 |  | 12.30 |  | 156 |
|  |  | BHigh |  | 56.34 |  | 12.29 |  | 156 |
| FHigh |  | BLow |  | 64.80 |  | 11.88 |  | 156 |
|  |  | BHigh |  | 78.45 |  | 10.89 |  | 156 |

Bonferroni alpha / comparisons

* .05/2 = .025
* compare my p value to .025

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| IV1 | Group 1 | Group 2 | *t* | Effect Size | Interpretation |
| FSG Low | BSG Low  M = 50.66  SD = 12.30 | BSG High  M = 56.34  SD = 12.29 | *t*(155) = -11.93, *p* < .001 | *ddiff* = -0.96  *d*RM | Significant |
| FSG High | BSG Low  M = 64.80  SD = 11.88 | BSG High  M = 78.45  SD = 10.89 | *t*(155) = -16.12, *p* < .001 | *ddiff* = -1.29 | Significant |

While there are new options for effect size in JASP, those work for one-way tests or main effects only. You could still calculate Cohen’s d as described here.