# Complete Example (Repeated Measures – 1 IV only)

A note on repeated measures designs:

* People sometimes find it difficult to separate IV and DV in these designs because they are related. Always make sure you can differentiate between the levels of the IV and what was measured at each of those levels (i.e. the DV).
* These designs are more powerful (need less people!) because you are measuring the same people a couple times. That means that you reduce the error variance (within subjects) because you can control for the fact that everyone is slightly different and take that out of the error variance. However, the drawback is that there is an interaction with the study sometimes (i.e. if you take the same test over and over, of course you’ll get better … or worse because you are bored.). So, you have to balance these effects (carry over effects, fatigue) with the ability to control for participant wackiness.

Participants were tested over several days to measure variations in their pulse given different types of stimuli. One stimulus was a neutral picture (like a toaster), while other stimuli were cute/happy pictures (puppies, babies), and negative stimuli (mutilated faces, pictures of war). Where there differences in pulse for each participant across the stimuli?

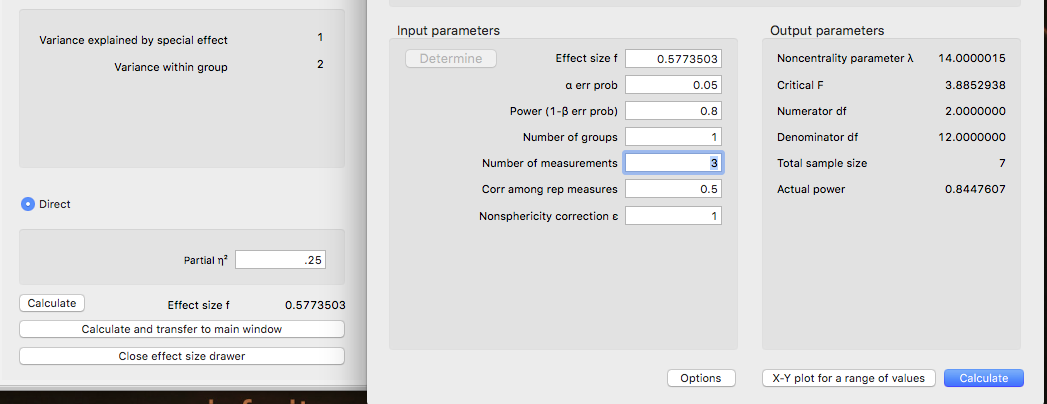
**Datafile:** rm 1 anova.csv

**IV:** Stimuli type – neutral, negative, positive

**DV:** Pulse/heart rate

**Power:**

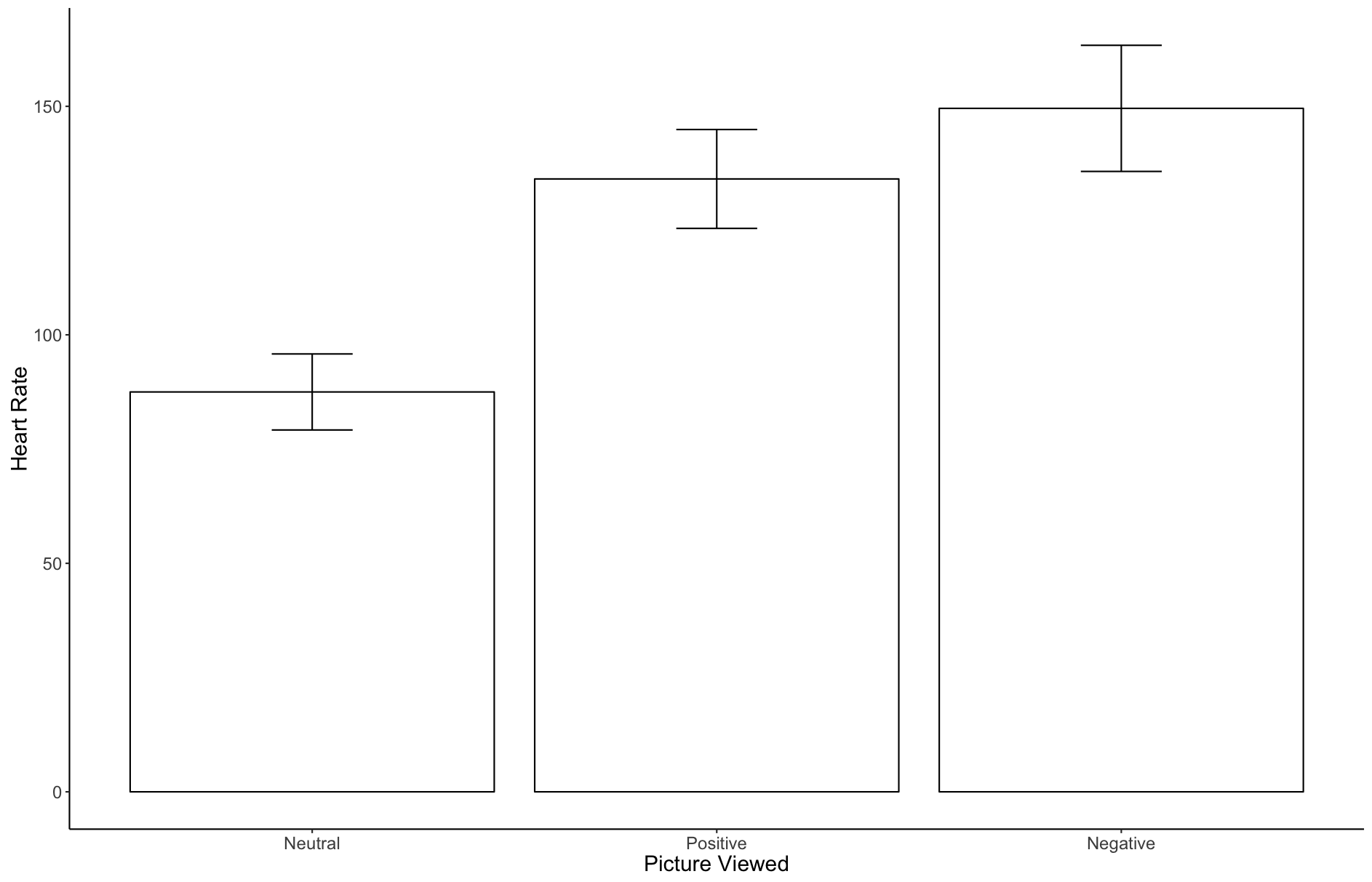
1. Open Gpower!
   1. Test family: F-test
   2. Statistical Test: ANOVA repeated measures, within factors
   3. Estimate an effect size: click determine 🡪 click direct 🡪 use eta square sizes you think might be accurate, remember small, medium, and large estimates from the notes.
   4. Alpha = .05
   5. Power (1-beta .20) = .80
   6. Number of groups = number of IVs
   7. Number of measurements = number of levels
   8. Corr among rep measures = correlation between levels
      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 … 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. Large effect size
   2. One IV
   3. Levels from our current study
   4. Correlation = .5
   5. Epsilon = 1
3. Says we needed to run 7 people to find a significant effect with a large effect size!

****

1. Check your learning:
   1. See if you can estimate a small effect size with a high correlation between measures (*r* = .80). You should get that you need 66 people.

**Results**

Participants were tested on three types of stimuli (neutral, positive, negative) for pulse rate. Data was screened for errors, missing data (none), outliers (none found with *z*-score distance), and assumptions. Normality, linearity, homogeneity, and Mauchly’s test (*p* = .002) indicated that assumptions were met. Using a repeated measures ANOVA, different stimuli were found have different heart rates, *F*(2, 34) = 253.30*, p* < .001*, η²* = .94. Post hoc comparisons were analyzed using dependent *t*-tests with a Tukey correction. Neutral stimuli were found to have a significantly lower mean than Positive stimuli (*t*(17) = -16.24, *p* < .001, *ddiff* = -5.03), as well as Negative stimuli (*t*(17) = -21.62, *p* < .001, *ddiff* = -3.86). Positive stimuli showed lower heart rates than Negative stimuli (*t*(17) = -5.38, *p* < .001, *ddiff* = -1.54). Therefore, heart rates were lowest at Neutral stimuli, followed by Positive stimuli, and highest with Negative Stimuli. Figure 1 displays the average heart rates along with confidence intervals.



*Figure 1.* Average heart rates with 95% confidence intervals for each stimuli type.

*Figure 1*. Average pulse rates with one SD error bars for each stimuli type.

# Results

## Repeated Measures ANOVA

| **Within Subjects Effects** | | | | | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Sphericity Correction** | | | | | **Sum of Squares** | | | | | | | | **df** | | **Mean Square** | | **F** | | **p** | | **η²** | |
| Stimuli |  | None | | | |  | 37572 | | | | | | ᵃ | | 2.000 | ᵃ | 18786.06 | ᵃ | 253.3 | ᵃ | < .001 | ᵃ | 0.937 |  |
|  |  | Greenhouse-Geisser | | | |  | 37572 | | | | | | ᵃ | | 1.286 | ᵃ | 29210.66 | ᵃ | 253.3 | ᵃ | < .001 | ᵃ | 0.937 |  |
|  |  | Huynh-Feldt | | | |  | 37572 | | | | | | ᵃ | | 1.346 | ᵃ | 27911.63 | ᵃ | 253.3 | ᵃ | < .001 | ᵃ | 0.937 |  |
| Residual |  | None | | | |  | 2522 | | | | | |  | | 34.000 |  | 74.17 |  |  |  |  |  |  |  |
|  |  | Greenhouse-Geisser | | | |  | 2522 | | | | | |  | | 21.866 |  | 115.33 |  |  |  |  |  |  |  |
|  |  | Huynh-Feldt | | | |  | 2522 | | | | | |  | | 22.884 |  | 110.20 |  |  |  |  |  |  |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | |
| Note.  Type III Sum of Squares | | | | | | | | | | | | | | | | | | | | | | | | |
| ᵃ Mauchly's test of sphericity indicates that the assumption of sphericity is violated (p < .05). | | | | | | | | | | | | | | | | | | | | | | | | |
| **Between Subjects Effects** | | | | | | | | | | | | | | | |
|  | | **Sum of Squares** | | **df** | | **Mean Square** | | | **F** | | **p** | | | **η²** | |
| Residual |  | 2522 |  | 34 |  | 74.17 | |  |  |  |  |  | |  |  |
|  | | | | | | | | | | | | | | | |
| Note.  Type III Sum of Squares | | | | | | | | | | | | | | | |

### *F*(2, 34) = 253.30, *p* < .001, η² = .94

### Assumption Checks

| **Test of Sphericity** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Mauchly's W** | | **p** | | **Greenhouse-Geisser ε** | | **Huynh-Feldt ε** | |
| Stimuli |  | 0.445 |  | 0.002 |  | 0.643 |  | 0.673 |  |
|  | | | | | | | | | |

### Mauchly’s test was not significant, *p* = .002.

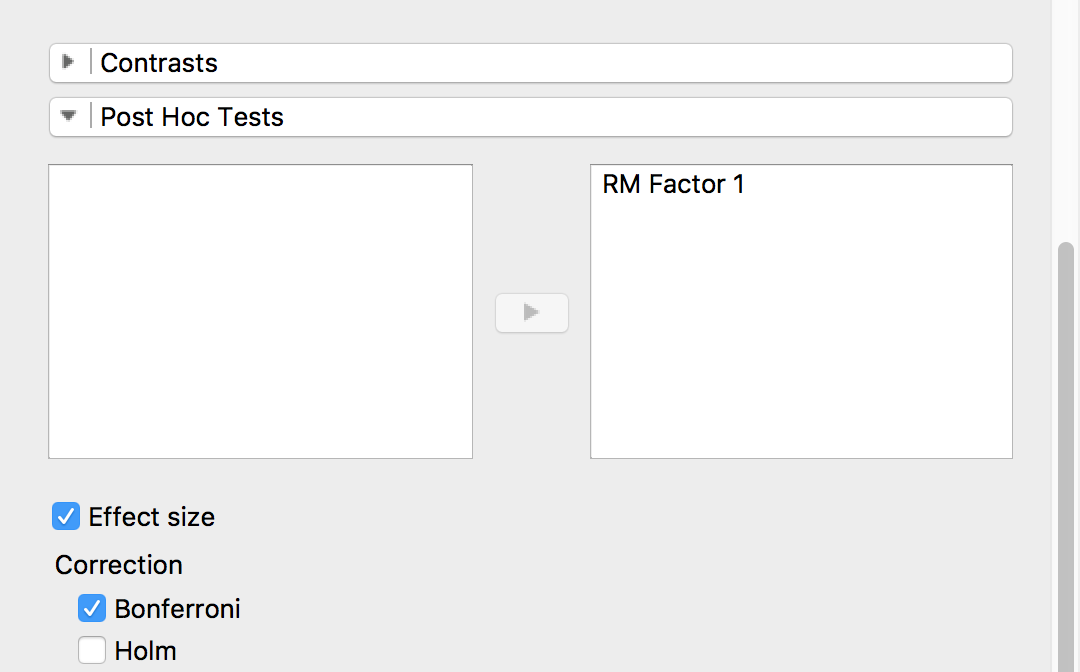
|  |  |  |  |
| --- | --- | --- | --- |
| Group 1 | Group 2 | Comparison | *ddiff* |
| Neutral  M = 87.50  SD = 16.73 | Positive  M = 134.11  SD = 21.75 | *t*(17) = -16.24, *p* < .001 | -5.03 |
| Neutral  M = 87.50  SD = 16.73 | Negative  M = 149.56  SD = 27.75 | *t*(17) = -21.62, *p* < .001 | -3.86 |
| Positive  M = 134.11  SD = 21.75 | Negative  M = 149.56  SD = 27.75 | *t*(17) = -5.38, *p* < .001 | -1.54 |

NEW NEW NEW

### Post Hoc Tests

| **Post Hoc Comparisons - Stimuli** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | **Mean Difference** | | **SE** | | **t** | | **p tukey** | |
| Neutral |  | Positive |  | -46.61 |  | 2.871 |  | -16.236 |  | < .001 |  |
|  |  | Negative |  | -62.06 |  | 2.871 |  | -21.616 |  | < .001 |  |
| Positive |  | Negative |  | -15.44 |  | 2.871 |  | -5.380 |  | < .001 |  |
|  | | | | | | | | | | | |

### You will not be able to use Tukey for repeated measures tests, as they have disabled it. You should choose Bonferroni:



There’s also a new Effect size button that gives you Cohen’s d right there for the post hoc test.

### Post Hoc Tests

| **Post Hoc Comparisons - RM Factor 1** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | **Mean Difference** | | **SE** | | **t** | | **Cohen's d** | | **p bonf** | |
| Level 1 |  | Level 2 |  | -46.61 |  | 2.185 |  | -21.337 |  | -5.029 |  | < .001 |  |
|  |  | Level 3 |  | -62.06 |  | 3.790 |  | -16.374 |  | -3.859 |  | < .001 |  |
| Level 2 |  | Level 3 |  | -15.44 |  | 2.364 |  | -6.532 |  | -1.540 |  | < .001 |  |
|  | | | | | | | | | | | | | |
| *Note.*  Cohen's d does not correct for multiple comparisons. | | | | | | | | | | | | | |

### Descriptives

| **Descriptives** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Stimuli** | | **Mean** | | **SD** | | **N** | |
| Neutral |  | 87.50 |  | 16.73 |  | 18 |  |
| Positive |  | 134.11 |  | 21.75 |  | 18 |  |
| Negative |  | 149.56 |  | 27.75 |  | 18 |  |
|  | | | | | | | |

## T-Test

| **Paired Samples T-Test** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | |  | | **t** | | **df** | | **p** | | **Cohen's d** | |
| neutral |  | - |  | positive |  | -21.337 |  | 17 |  | < .001 |  | -5.029 |  |
| neutral |  | - |  | negative |  | -16.374 |  | 17 |  | < .001 |  | -3.859 |  |
| positive |  | - |  | negative |  | -6.532 |  | 17 |  | < .001 |  | -1.540 |  |
|  | | | | | | | | | | | | | |
| Note.  Student's T-Test. | | | | | | | | | | | | | |