Repeated Measures ANOVA (single factor)

1. Within subjects design
   1. Normally each person gets one treatment – and there are several treatments, necessitating different groups of people – between subjects design
   2. All the variance within groups (error variance, bad variance, SSs/a) was considered a problem because it was the individualness of the person.
   3. Now we can control for that variance, by testing the same people repeatedly (within subject design aka paired aka repeated measures).
2. Why would you use them?
   1. Control
   2. Experimental design (for example, studying test scores across the semester)
   3. Limited resources
   4. More power
3. (you can ignore the section on mixed designs page 348-349, as we are doing factorial anova next)
4. First, the SPSS set up:
   1. Remember that each person gets their own row – so now each person will have multiple columns – one for each time they were measured, or for each question on a quiz, or each piece of information about them.
5. The ANOVA
   1. In SPSS you will only get two variance numbers
      1. SSa for the repeated factor – each treatment (not group now, but individual measure) minus the grand mean
         1. Same as before
      2. SSs/a – error – (axs in the book) – measurement of how different each person is from the treatments AND the other subjects
         1. Before this measure was how different was I from the group (Y – A), now it’s how different am I from the group and from MY normal scores (measures consistency of person)
         2. This is where you get extra power – you subtract out the average score for each person
   2. Other numbers not in SPSS
      1. Total – each person minus the grand mean (exactly the same)
      2. Subject variance – each subjects average minus the grand mean

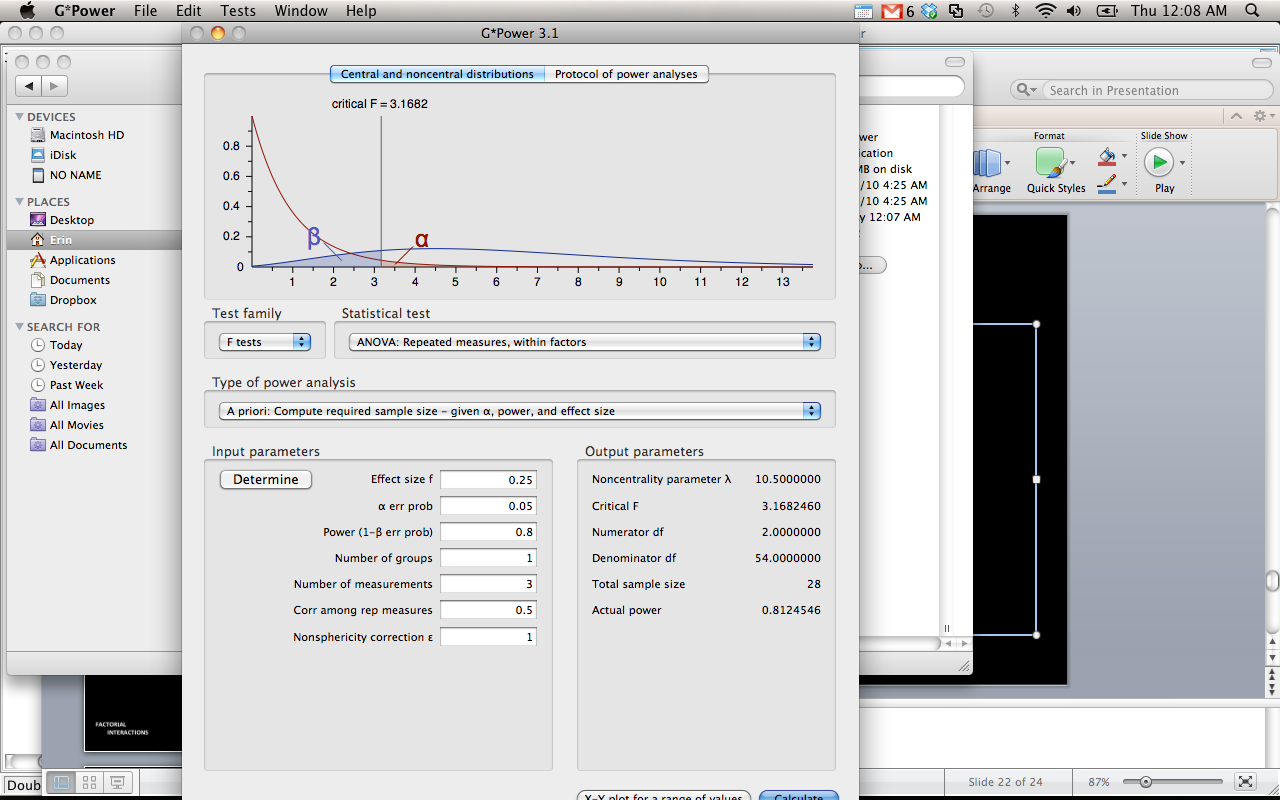
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tests of Within-Subjects Effects** | | | | | | | |
| Measure: MEASURE\_1 | | | | | | | |
| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
| factor1 | Sphericity Assumed | .853 | 2 | .426 | 1.034 | .357 | .005 |
| Greenhouse-Geisser | .853 | 1.991 | .428 | 1.034 | .356 | .005 |
| Huynh-Feldt | .853 | 2.000 | .426 | 1.034 | .357 | .005 |
| Lower-bound | .853 | 1.000 | .853 | 1.034 | .310 | .005 |
| Error(factor1) | Sphericity Assumed | 155.814 | 378 | .412 |  |  |  |
| Greenhouse-Geisser | 155.814 | 376.297 | .414 |  |  |  |
| Huynh-Feldt | 155.814 | 378.000 | .412 |  |  |  |
| Lower-bound | 155.814 | 189.000 | .824 |  |  |  |

* 1. The actual ANOVA table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | SS | Df | MS | F |
| A – treatments | each treatment minus grand mean | A – 1  A means treatments here | SSa / dfa | MSa / MSs/a |
| S/A Error | Each person minus the treatment meant minus the subject average plus grand mean | (n – 1)(a-1)  here little n and big N are the same | SSs/a / dfs/a |  |
| Total (not on SPSS) | Each person minus the grand mean | An - 1 |  |  |

* 1. Post hoc tests
     1. Dependent t-tests are the most commonly used post hoc test (dependent because these means are PAIRED).
        1. However that doesn’t help with the error control business when you use a lot of post hoc tests
     2. Bonferroni, Sidak-Bonferroni, and Scheffe corrections still apply.
        1. Scheffe corrected F = ((dfa X dfs/a) / (dfs/a – dfa + 1)) times F critical (dfa, dfs/a-dfa + 1) – page 361 for better written out
        2. There are options in SPSS (hidden)
        3. Under options > move mean over > click the button under means
     3. Some say you can use the same corrections for fisher-hayter, tukey, etc. calculating the smallest difference – using treatments as groups (not as popular as dependent t)

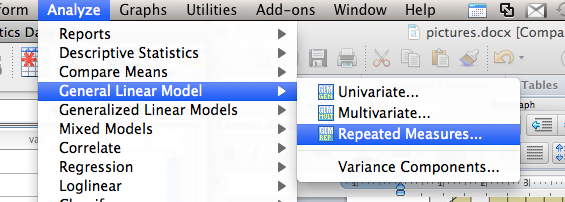
1. Effect size
   1. R2 and n2 are still the most popular = SSa / SSa + SSs/a (since total isn’t on the SPSS output) – you can still get these values when selecting effect size under options
2. Power Gpower
   1. F-test, ANOVA repeated measures, within factors
   2. Alpha = .05
   3. Beta = .80
   4. Number of groups = number of Ivs
   5. Number of measurements = number of times

****

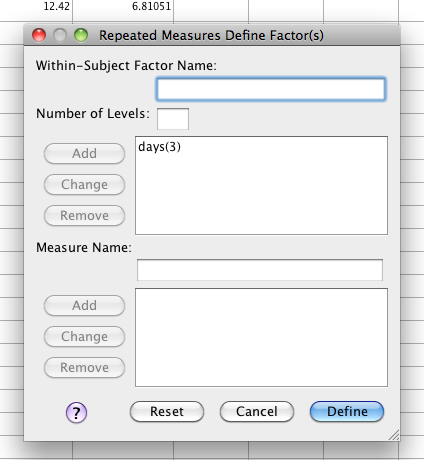
1. New fancy assumption (page 376 – chapter 17)
   1. Spherecity – test of homogeneity for repeated measures (i.e. if you try to get Levene’s from SPSS it will give you an error).
   2. Compound symmetry
      1. Idea that measures should be of the same type and measured in the same way – so they get roughly equal variances
      2. And then that you are testing the same subjects over and over, so the correlations between time measurements should be roughly the same
      3. Mauchley’s will be in SPSS ouput (unless you only have two time measurements)
      4. You will almost NEVER meet this assumption (ACK! WHY?!)
   3. What to do?
      1. Corrections built into ANOVA
      2. Have a lot of people

**Running the analysis:**

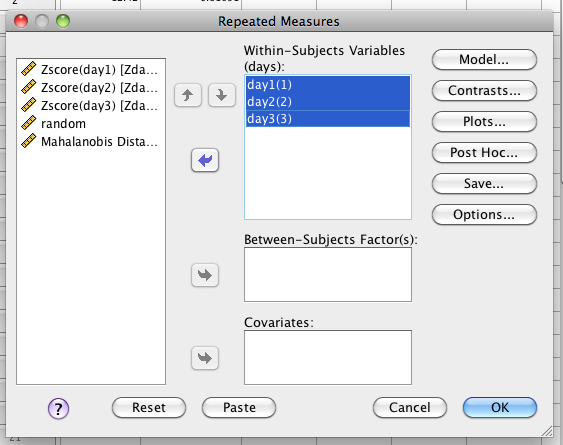
1. Analyze > general linear model > repeated measures (any analysis with a repeated measures variable must use repeated measures GLM).



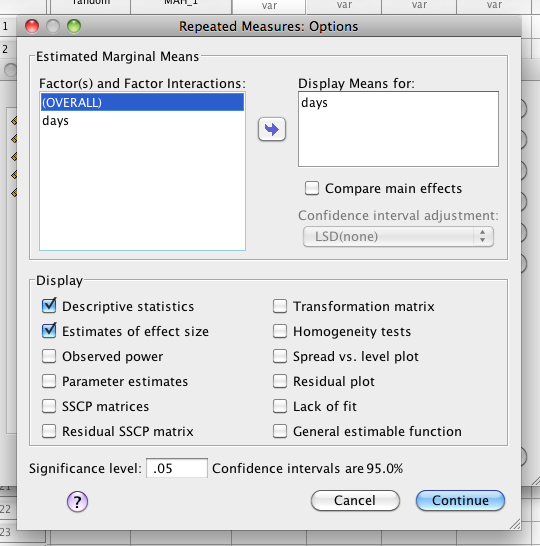
1. When the define box comes up:
   1. Make up a name for the variable (so you can find it later).
   2. List how many times that variable was measured.
   3. Hit add.
   4. Hit define.



1. Move over the time measurements (in the right order!) into the ?\_\_(1) spots.

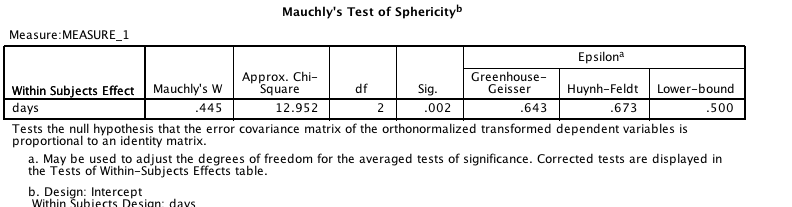


1. Hit options
   1. Move over the mean to the right.
   2. Hit descriptive statistics, effect size
   3. Hit continue and ok.

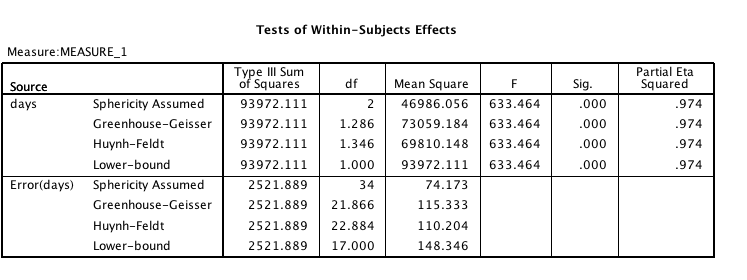


**Translating the output:**

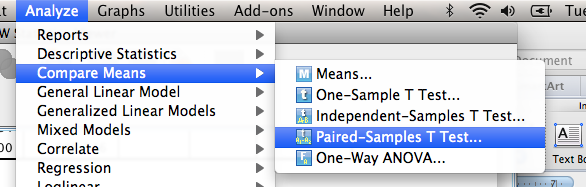
1. Ignore the multivariate box, contrasts and between subject boxes.
2. Check out the Spherecity box – you want this test to be p > .001.
   1. If it is p<.001, then you’ll want to use a correction
      1. Greenhouse-geisser is the most common.
      2. Sphericity is akin to homogeneity for repeated measures.
   2. However, this test is not super popular because you don’t necessarily expect homogeneity…

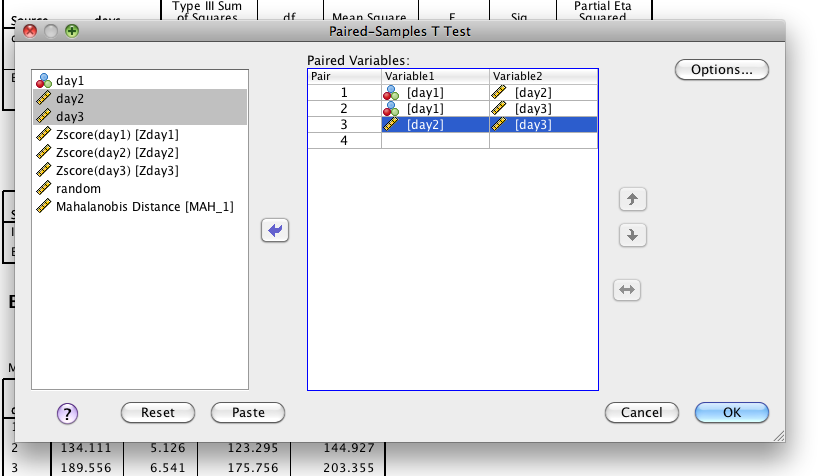


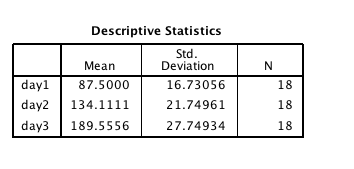
1. The ANOVA box – Tests of within-subjects effects:
   1. You will use only the correction line you choose (most commonly sphericity assumed).
   2. You’ll get the df, F value and Sig - you want Sig to be p<.05.
   3. *F*(2, 34) = 633.46, *p*<.001, *n*2 = .97

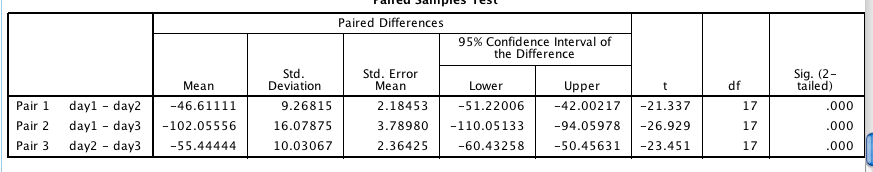


1. **If the ANOVA is significant Post Hocs:**
   1. The matching test for repeated measures is a dependent t-test.
   2. You will want to run the combinations of dependent t’s you hypothesized about (most people have a specific question OR will run all of them).
      1. If you run ALL of them and have a lot, then you will want to correct.
      2. The easiest correction is a Bonferroni – divide your p-value (.05) by the number of tests and look for that new p-value in the sig column. So instead of looking for Sig < .05, you might be looking for Sig < .017.
   3. Analyze > compare means > paired samples
      1. Create all combinations of the time measurements in the paired variables box and hit ok.
   4. Again, a chart really helps:

****

****

****

****

Results

Participants were tested on three successive days for pulse rate. On Day 1, they were told to sit at rest. Day 2 measured pulse rates after a slow jog, and Day 3 measured pulse rates after a cardio workout. Using a repeated measures ANOVA, different days were found have different heart rates, *F*(2,34)=633.46, *p*<.001, η2 = .97. Post hoc comparisons were analyzed using dependent samples t-tests. Day 1 was found to have a significantly lower mean than Day 2 (Mdifference = -46.11, *t*(17) = -21.34, *p*<.001), as well as Day 3 (Mdifference = -102.06, *t*(17) = -26.93, *p*<.001). Day 2 showed lower pulse rates than Day 3 (Mdifference = -55.44, *t*(17) = -23.45, *p*<.001). Therefore, pulse rates were lowest at rest, followed by slow jogs, and highest with more intensive workouts. Figure 1 shows the average pulse rates for each day.

*Figure 1.*