

Power Simulation

Jose Ossandon

2/12/2018

Paired comparisons

To evaluate how many subjects and trials per would be necessary to have a powered experiment replication, I first evaluated the experiment's paired-comparisons between the main conditions. These correspond to the crossing of the factors response mode (2 levels, external 'Re' and anatomical), crossing of the legs (2 levels, uncrossed 'Lu' and crossed 'Lc') and crossing of the hands (2 levels, uncrossed 'Hu' and crossed 'Hc').

p-values

```
##
## Pairwise comparisons using paired t tests
##
## data:  auxres$RT and auxres$cond
##
##      ReLuHu ReLuHc ReLcHu ReLcHc RaLuHu RaLuHc RaLcHu
## ReLuHc 9.6e-07 -      -      -      -      -      -
## ReLcHu 5.4e-11 2.9e-09 -      -      -      -      -
## ReLcHc 1.2e-09 3.5e-08 8.5e-06 -      -      -      -
## RaLuHu 1.00000 4.2e-06 2.3e-10 3.1e-09 -      -      -
## RaLuHc 9.6e-08 0.00043 0.00029 0.77783 1.1e-09 -      -
## RaLcHu 8.0e-09 2.6e-05 2.9e-05 0.31408 2.5e-10 1.00000 -
## RaLcHc 4.4e-09 6.1e-07 0.00050 0.23700 1.5e-09 0.02555 0.01323
##
## P value adjustment method: holm
```

Significant comparison (alpha<.05 corrected)

```
##      ReLuHu ReLuHc ReLcHu ReLcHc RaLuHu RaLuHc RaLcHu
## ReLuHc  TRUE    NA     NA     NA     NA     NA     NA
## ReLcHu  TRUE    TRUE    NA     NA     NA     NA     NA
## ReLcHc  TRUE    TRUE    TRUE    NA     NA     NA     NA
## RaLuHu  FALSE    TRUE    TRUE    TRUE    NA     NA     NA
## RaLuHc  TRUE    TRUE    TRUE    FALSE  TRUE    NA     NA
## RaLcHu  TRUE    TRUE    TRUE    FALSE  TRUE    FALSE    NA
## RaLcHc  TRUE    TRUE    TRUE    FALSE  TRUE    TRUE     TRUE
```

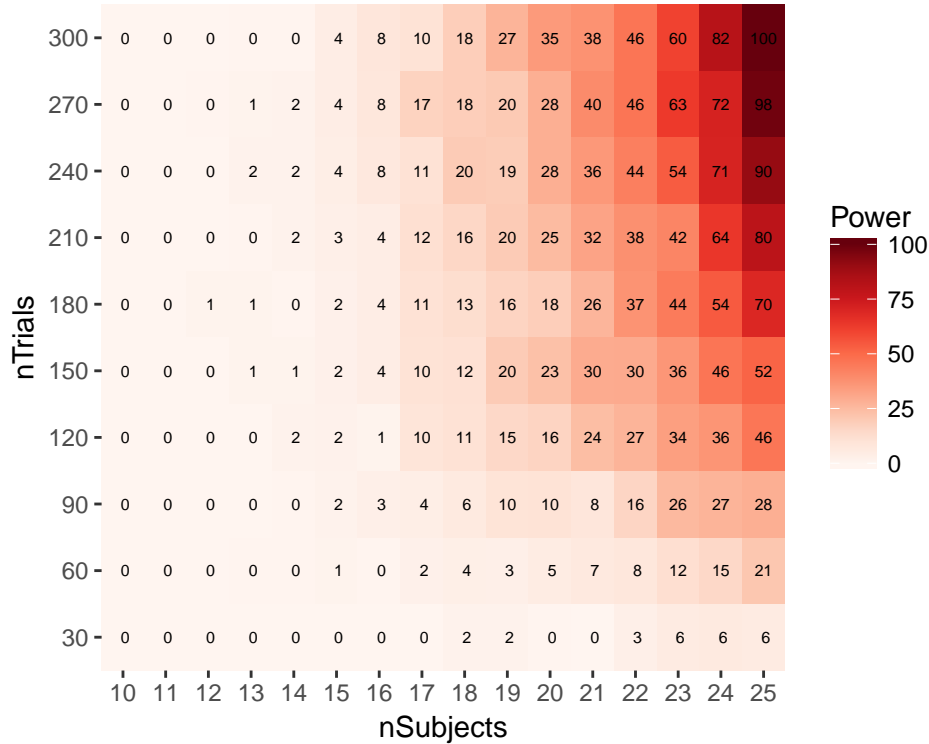
As it can be seen from this result, most of the comparisons between conditions are significant and with very low p-values, reflecting the fact that they were large and strongly consistent across subjects.

Simulations

Each experiment simulation was performed by taking a random sample of subjects (total # of subjects: 25) and trials per condition (total # of trials per condition: 300). For every combination of #subjects and #trials 200 repetitions were done. The result of each simulation's significant-comparison table was compared to the actual result of the experiment (the table above).

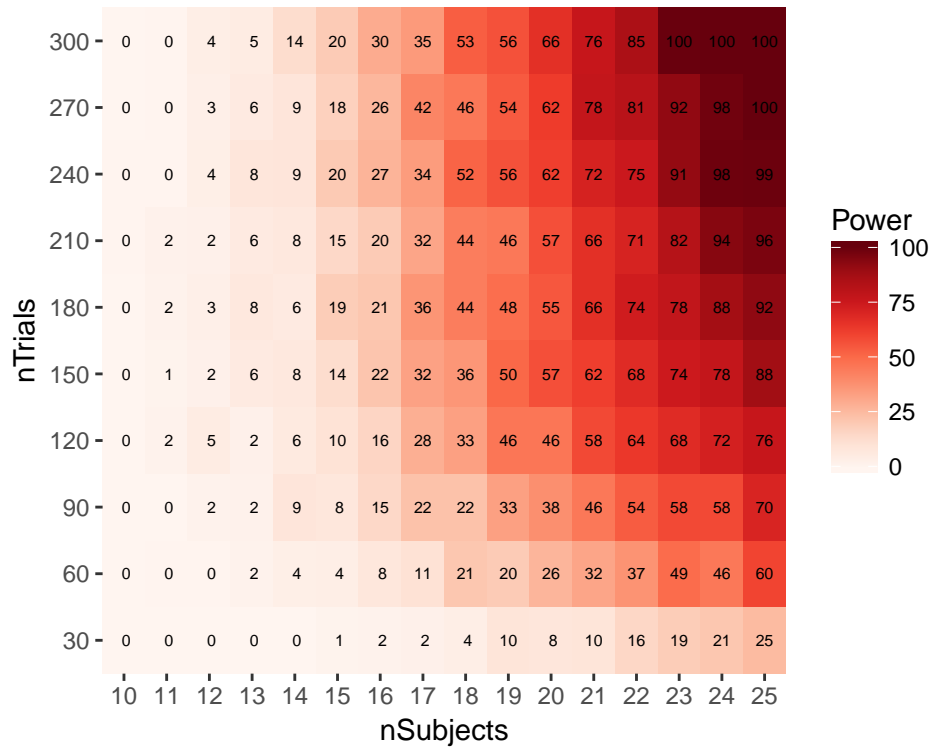
For the following plot, a simulation was considered a 'successful' experiment replication when the comparison-tables were exactly equal (i.e., the result of being significant or not of the 28 possible paired-comparisons was exactly the same between the actual results and the result in the simulated experiment). The 'power' value indicates the % of the 200 simulations that showed exactly the same result.

Simulations exactly equal

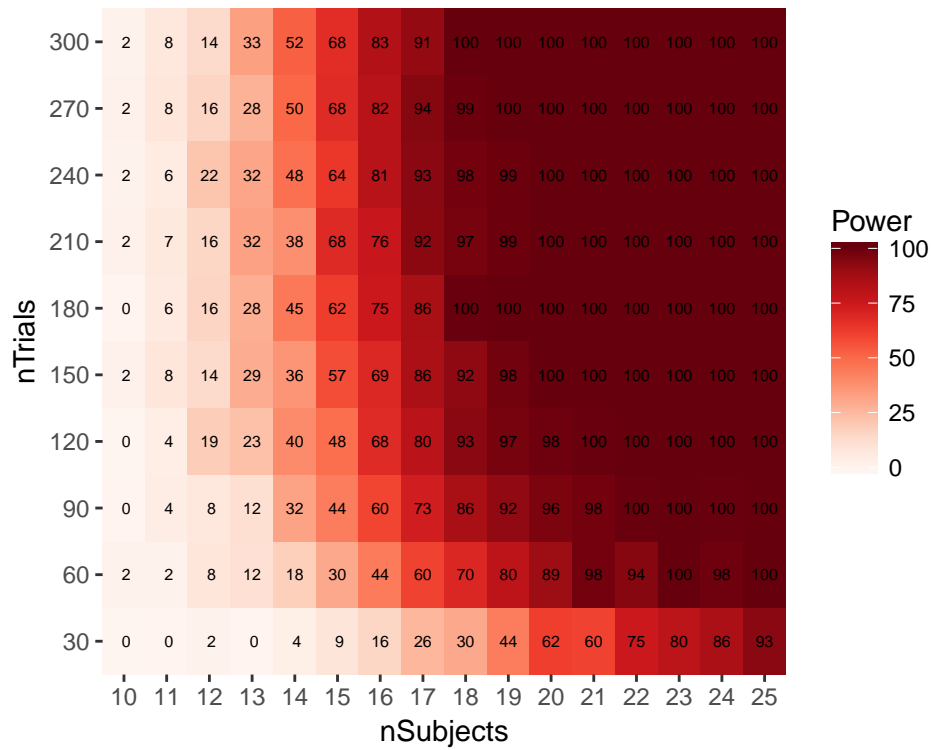


Considering a replication successful only when all the comparisons have the same result might be too strict, specially when considering the number of comparisons (28), that most comparisons were significant in the original experiment (23/28) but two of them had p-values that were only between 0.01 and 0.05 (adjusted). Furthermore these two barely significant comparisons represent differences that are numerically small (20-40 ms) when compared to the other mayor effects we see. Therefore the following three plots show the simulated power when considered a replication succesful when there is only one, two, or three comparisons that are different.

Different result in only one comparison



Different result in two comparisons



Different result in three comparisons

