# Power Project – Data Extraction Steps

## Extraction Spreadsheet

1. Design
   1. The design used in the power calculation. If they described the design used in the power calculation we enter it here.
   2. No clear description of the design used: unknown
   3. No design mentioned: unknown
2. Alpha
   1. The alpha level used (e.g., 0.05) and whether it was one or two-tailed. If they do not report the tails, we can just enter the level and tail\_not\_reported.
3. Power
   1. The power they used in the power calculation (e.g., 80% of power).
4. Effect Size Type
   1. The effect size they used in the power calculation (e.g., partial eta squared).
   2. If they calculate using the pn2 (and they say that explicitly), then we should go with the pn2.
   3. If they say pn2 was used to calculated f, then go with pn2.
5. Effect Size Value
   1. The value of the effect size they used.
6. Software
   1. The software they used to run the power analysis.
   2. No mentioned: unknown.
7. Reproducible with information provided:
   1. Yes – only if we have all the information needed to reproduce the power calculation.
   2. No – If we do not have enough information
8. Eta square setting
   1. If they mention the effect size they used was eta squared, then choose from a drop-down list.
   2. If they did not use eta square, choose: not\_app
9. Conditionally reproducible
   1. Yes - If we have to make assumptions to reproduce the power calculation
   2. No – Even when making assumptions we are not able to reproduce
   3. Not applicable – if we were able to reproduce without assumptions
10. Conditions Assumed
    1. Write down the assumptions necessary to reproduce the power calculation
       1. Take notes on correlation and nonsphericity (GPower) only when they deviate from the default (0.5 and 1, respectively).
11. Appropriate software
    1. Yes – if the design they used was anything but (see below)
    2. No – If the design they used was a mixed factors anova with 3 or more within subject levels , because in this case they would need superpower. Superpower makes the case that only SAS can handle 3 or more within subject levels. It has to do with the Greenhouse–Geisser correction (non-sphericity correction). In G power it says: all the within-subject variables are correlated with one another at whatever correlation value is specified as the default but this is unlikely to be the case in a real experiment (the sphericity is violated especially when we have more than 2 within subjects variables). Superpower can handle that because it can fit a matrix of correlations.
    3. It the design is unknown AND you were not able to conditionally reproduce: NOT\_APP
    4. If the design is unknown AND you WERE able to conditionally, check to see if the software was appropriate. If it was, YES. If not, NO.
    5. If the design is known, but you were not able to reproduce it: NOT\_APP
    6. If the design is known, but the app is not mentioned: NOT\_APP
12. Hypothesis:
    1. Quote from the article. If they do not use words such as “We hypothesized”  
        or We predicted, go with “The purpose of the study was…”
13. Matches primary hypothesis
    1. Match the information in the hypothesis column with the analysis they used to run the power calculation. Possible scenarios include: yes, no, match one, match none.
    2. Match\_one: matches at least one of the hypotheses in articles that present more than no.
14. Match design
    1. Check to see if the design used in the power calculation matches the design used in the statistical analyses.
15. Reproduce Software
    1. Record the software used to reproduce the power calculation
16. Notes
    1. Anything worth taking notes of.

## Primary outcome information

Scenarios with no explicit primary outcome:

1. One DV measured at one time point. Simple! One primary outcome.
2. One DV measured at *k* time points. More difficult. Two sub-scenarios:
   1. Multiple time points are analyzed separately (i.e., acquisition, retention, transfer). I suggest each time point is an outcome.
   2. Multiple time points are analyzed together in same analysis. I suggest each level of the within-subject time factor is an outcome. (\*note that acquisition would be 1 time point if analyzed as described above but could be *k* time points if each block is included in the same analysis. I think this – in theory - reflects the researcher’s opinion about which time scale is most relevant).
3. *L* DVs measured at one time point. *L* primary outcomes.
4. *L* DVs measured at *k* time points. *L* X *k* primary outcomes. I had been thinking this would be the point we would simply say “primary\_not\_specified.” An alternative approach would be to add a column that codes whether a primary is made explicit. If yes, then the primary outcome column records how many there are. If no, then the primary outcome column treats all outcomes as primary and calculates the total. This might be better, what are your thoughts?
5. Within-subject designs: We focus only on separate DVs included in the analyses. This could get a bit tricky because the DV that gets analyzed might include a time component (i.e., change in a given measure; difference score between measures taken at different times/conditions).

If the article explicitly states which DVs are primary, but doesn’t specify which time points, we follow the logic above but focusing only on the primary DVs. If the article actually specifies which DV and time point is primary then we use that.