Bayesian mixture model / estimating the effect of publciaiton bias on effect sizes accounting for nulls

/

Maybe,

For effect sizes and test statistics such as Cohen's d, correlations, t statistics, values were converted into correlation coefficents and standard errors were calculated following {Open Science Collaboration, 2015 #611}. However, F values with a df1 of > 1, χ2 values, standard errors have been converted to correlations and Cohen's d values, but methods of obtaining standard errors cannot be computed directly. As a simplifying assumption, the following analyses were run estimating the precision of each study as as [] equation standard error ^ 2 []. In the

#### Conversion

All effect sizes were converted into correlation coefficients, or extracted as such. Only effect sizes which could be converted into correlation coefficients are used in the current analysis.

This was done for two primary reasons. Most of the large scale replication projects have primarily reported their results in Pearson's r (i.e., all but one of the projects included in the current research). Secondarily, r values are broadly well understood and commonly used by psychologists.

One of the benefits of this approach is that standard errors can be developed for the Cohen's z transform of the correlation coefficients, and for differences between z-transformed correlation coefficients, if values were developed from test statistics of r, t, or F(1,df2) following (Open Science Collaboration, 2015). However, standard errors developed this way are not valid for F statistics with a $df\_1$ of > 1 or chi square statistics. These studies have been excluded from analyses which require these values [...!...].

Effects which were non-significant were included following the inclusion rules used by studies included - if the original authors reported them as significant. This means that some of the p values for included studies were in fact over .05. Results which were not reported as significant were not included.

Because in the loopr study, some measures used shorter form version of the original results, all results presented have been disattenuated using the Spearman-Brown prediction formula and Spearman disattenuation formula to estimate the trait-outcome associations that would be expected if our outcome measure had used the same number of items as the original study (Lord & Novick, 1968). Following the other large scale replication studies, the signs of negative original correlations were set to positive (and the sign of the replication sample were switched too).

xPhi included 2 origianl studies which were non-significant, these were removed from analysis.

IT MAY BE WORTH EXLUDING THOSE NON-sig few from the begining

Many labs 2 - CIs and P values were recalculated from summary statitsics (i.e., from Cohen's d) and convereted to rs and z values, assuming statistically significant main result. Two studies were removed because effect sizes could not be derrived (the original and replication studies examined a differnece in effect sizes, no easy method of summarising these differeces), and two additional were removed because their effect sizes were only avalible in Cohen's q.