**Supplementary Material for Harrison *et al*. 2020**

*Calculating I2*

We extracted *I*2 from our meta-analytic intercept-only models (see Supplementary Table S1) using the following equation:

Where is the total variance, is the phylogenetic variance, is the between-study variance, is the species-specific variance, is the study-specific variance (observation-level random effect), and is the remaining within-study sampling variance (random effects) (Nakagawa & Santos 2012).

Further, we can then partition *I*2 to calculate study-level *I*2 and species-level *I*2 (*I*2s and *I*2u, respectively) (Nakagawa & Santos 2012):

*I*2s = / ,

*I*2u = /

Finally, we can determine the strength of phylogenetic variance …

Where = 0 there is no phylogenetic signal, but when = 1 there is a strong effect of phylogeny on heterogeneity (Nakagawa & Santos 2012).

*Can talk about categorising SSD in here … include dataset that had my references in it… also for the phylogenies*

*Keyword search terms and exclusion*

Move table of keywords into here, then expand on exclusion terms, what was excluded (i.e. papers that were missing sample sizes, error, means were 0 so we couldn’t use them in our models etc). lots of ns for transparency…

*Exploratory analysis*

What we did and why it was just exploratory analysis … maybe cut and paste from the methods section in ms

*Results - Mating system, age, population, study environment and study type can influence personality*

Monogamous and multiple mating systems were not significantly different from each other for means or variability for any of the taxonomic groups (Supplementary Table S7). Mean personality effect sizes for adults were marginally significantly different to juveniles for invertebrates (intercept: =0.24, 95% CIs: -0.03, 0.51, *t* = 1.74, *p*=0.08; juvenile: =-0.03, 95% CIs: -0.34, 0.28, *t* = -0.18, *p*=0.86), but not any other taxonomic group, and not for variability (Supplementary Table S8). Fish from the wild had greater differences in variability than fish from lab populations (intercept: =-0.09, 95% CIs: -0.18, -0.01, *t* = -2.11, *p*=0.04; lab: =0.08, 95% CIs: -0.02, 0.19, *t* = 1.55, *p*=0.12), but not for mean personality differences (Supplementary Table S9). Studies conducted in the lab were significantly different to field studies for mammals (intercept: =0.24, 95% CIs: -0.09, 0.57, *t* = 1.44, *p*=0.15; lab: = -0.31, 95% CIs: -0.56, -0.05, *t* = -2.34, *p*=0.02), but not for variability, and not for any other taxonomic group (Supplementary Table S10). Finally, effect sizes from observational studies were significantly different from experimental studies for mammals (intercept: =0.00, 95% CIs: -0.21, 0.22, *t* = 0.03, *p*=0.98; observation: = 0.39, 95% CIs: 0.13, 0.64, *t* = 2.39, *p*=0.004), but not for variability, and not for any other taxonomic group (Supplementary Table S11).