Methods

Data Collection

We defined aggression as any variable that measured antagonistic behaviors (i.e. biting, chasing, or social rank) toward a conspecific (same sex and age) or mirror image. We searched Web of Science using the search terms color AND aggression OR aggressive. This yielded 751 papers as of XXX. We, then, filtered by removing papers with human subjects, no measure of coloration, or no measure of individual aggression. This yielded 449 papers. These papers were then subject to increased scrutiny and we removed any papers that did not measure individual color and aggression. This yielded 230 papers. From each paper, we collected the species name, color, what coloration pattern was measured (i.e. total body coloration or eye color), classification of color (melanic, xanthophore, carotenoid, or auxiliary) the measure of aggression (means/standard deviations or test statistics with p-values and degrees of freedom), the sample sizes for each measurement, the unit of aggression (i.e. time, number of acts, or distance), age, sex, location of study (captivity or in the wild), and season of coloration (Supplemental Table 1).

Effect Size Calculations

We used correlation coefficients that were converted to Fisher’s Z statistics for the meta-analysis for a standardized effect size statistic. The correlation coefficients (Pearson’s r, Spearman’s r, and XXX) and sample size were recorded for each study. When studies did not include correlation coefficients, we used the statistics (*t, F, X2*, or means and standard deviations) reported in the study to estimate a correlation coefficient using methods described in Nakagawa et al. (2007).

Phylogenetic Tree

We constructed a phylogenetic tree of all the species to limit any non-independence of our data due to evolutionary history and the relationship between species (Figure XX). We used our phylogenetic tree as a factor in our meta-analysis. We assembled a phylogenetic tree using clade specific phylogenies and removing any species that were not in our analysis (ADD IN CITATIONS). We, then, used the Time Tree, a database of published trees, to connect our clades (CITATION).

Meta-Analysis

We used the R package MCMCglmm to analyze our data. We used a random effects model with the random effects study, species, color, age, sex, location of study, and season with the standardized effects size as our predictor variable. We incorporated our phylogeny using the nodes argument. We repeated this analysis for each color classification (melanin, carotenoid, xanthophore, and auxiliary) and limited our phylogeny accordingly. Each MCMC was run for XX iterations and XX steps were removed as our burnin. We chose this number of iterations because our MCMC converged and produced biologically relevant values. All credible intervals reported are of the last XX iterations of our MCMC.

Since we do not have any fixed effects other than the intercept, we used a residual sum of squares for our measure of heterogeneity, *QREML*. *QREML* was then tested against a *X2* distribution with *k-1* degrees of freedom (Nakagawa et al., 2007). All analyses were done in R version XXX (R Core Team, 2020). Code and data to preform analyses can be viewed on GitHub, <https://github.com/sruckman/meta-analysis>.

References

Nakagawa, S., Ockendon, N., Gillespie, D.O., Hatchwell, B.J., and Burke, T. (2007). Assessing the function of house sparrows bib size using a flexible meta-analysis method. Behav. Ecol. *18*, 831.

R Core Team (2020). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).