Ungeneralizable generalizations? A meta-meta-analysis of the influence of taxonomic bias on the study of behavior.

Ned A. Dochtermann [a,*], M.A. Sekhar [a], S. Nakagawa [b]

[a] Department of Biological Sciences, North Dakota State University

[b] Department of Biological Sciences, University of Alberta

[*] corresponding author: ned.dochtermann@gmail.com

Abstract

Meta-analysis is a powerful tool for synthesizing behavioral research and identifying general patterns. However, are the conclusions we draw from these analyses truly representative across animal groups? Alternatively, are our conclusions shaped by taxonomic biases in the underlying research? For example, in animal behavior, vertebrates are overrepresented in the research we conduct. This taxonomic imbalance raises concerns about the validity of generalizations drawn in the field. To examine this issue, we examined the meta-analyses published in Animal Behaviour, Behavioral Ecology, and Behavioral Ecology and Sociobiology from 2000 – 2024. We then conducted a "meta-meta-analysis" to calculate the degree to which overall effects in prior meta-analytical results may have been mis-estimated due to taxonomic bias. We found that taxonomic biases in the primary research strongly influence effect size estimates in meta-analyses and can lead to improper inferences and generalizations. On average, meta-analytical averages are mis-estimated by ~35% ($p << 0.01$) and significance changes in about 25% of instances when sampling is taxonomically representative. Because meta-analyses aggregate data, they propagate the biases present in an area of research, leading to potentially incorrect generalizations. Addressing this taxonomic bias is critical to generalizations that describe the true richness of animal behavior.

Introduction

A major goal of animal behavior research is to understand why animals behave the way they do from both proximate and ultimate perspectives (Tinbergen 1963). Drawing general conclusions to answer such questions across species can be challenging and meta-analyses have emerged as the primary tool for doing so (Spake et al. 2022). However, the inferences drawn from any analysis, including meta-analyses, are reliant on whether the data being used are an appropriate sample (Gurevitch and Hedges 1999, Jennions et al. 2013, Konno et al. 2020). Put another way, if the data going into behavioral meta-analyses are not representative of animal behavior broadly, then the resulting inferences may be incorrect.

Biases in the data used in meta-analyses is a topic that has been extensively explored previously (e.g. Dickersin 2005, Rothstein et al. 2005). However, most of this discussion has focused on issues like the well-known "file drawer problem" and time-lag bias. The file drawer problem concerns publication bias wherein "statistically significant" findings are more likely to be published (reviewed by Dickersin 2005). This bias leads to the absolute magnitude of effect sizes being overestimated. Time-lag bias refers to the general observation that effect sizes tend to decrease over time (Trikalinos and Ioannidis 2005), as has been found in ecology (Jennions and Møller 2002), but which may not be particularly strong (Costello and Fox 2022). Time-lag bias leads to the absolute magnitude of effect sizes being overestimated early in a field's development. While not trivial, both types of biases can be at least partially addressed statistically within meta-analyses (Nakagawa and Santos 2012, Nakagawa et al. 2022).

More generally, and potentially more importantly, we do not have a good idea of how taxonomically biased the data going into meta-analyses might be (Gurevitch and Hedges 1999). However, for animal behavior research, we do know that individual research projects are conducted in a highly taxonomically biased manner. Rosenthal et al. (2017) examined papers published in *Animal Behaviour* between 1953 and 2015 and found that vertebrates were strongly over-represented (Figure 1A). At the coarsest level, vertebrates represent only 5% of animal species but were the focus of study in 71% of surveyed studies. Even within vertebrates there was considerable bias in what animals were studied (Rosenthal et al. 2017): endotherms were far more frequently studied than expected based on their taxonomic representation (Figure 1B) and well over 50% of all vertebrate studies were on birds or mammals (Figure 1B). Within invertebrates, most behavioral research was conducted in a single order (Hymenoptera, Rosenthal et al. 2017). Interestingly, this bias in favor of vertebrates is reduced in sexual selection and sexual conflict research. Zuk et al. (2014), found that roughly 30 percent of sexual selection and 50 percent of sexual conflict research is conducted with insects. However, even in these areas of study, vertebrates remain over-represented and over 20 percent of the work done on insects was from a single genus (i.e. Drosophila, Zuk et al. 2014). Subsequent work examining meta-analyses of sexual selection research found similar bias, with most meta-analyses focused on bird species and with insects still poorly represented (Pollo et al. 2024).

This potential for biased inferences also has impacts beyond our basic understanding of behavior. From an applied perspective, as discussed by Rosenthal et al. (2017), taxonomic bias in the study of behavior can affect conservation and management efforts, understanding population dynamics, and understanding zoonotic disease risk. Consequently, evaluating the presence of taxonomic bias in animal behavior meta-analyses and considering how such bias may affect the inferences we draw is of considerable importance. Specifically, it is necessary to know whether the conclusions we draw from meta-analyses are representative and generalizable.

We sought to address this concern by asking two questions: First, is the taxonomic bias identified by Rosenthal et al. (2017) also present in the data used in meta-analyses?

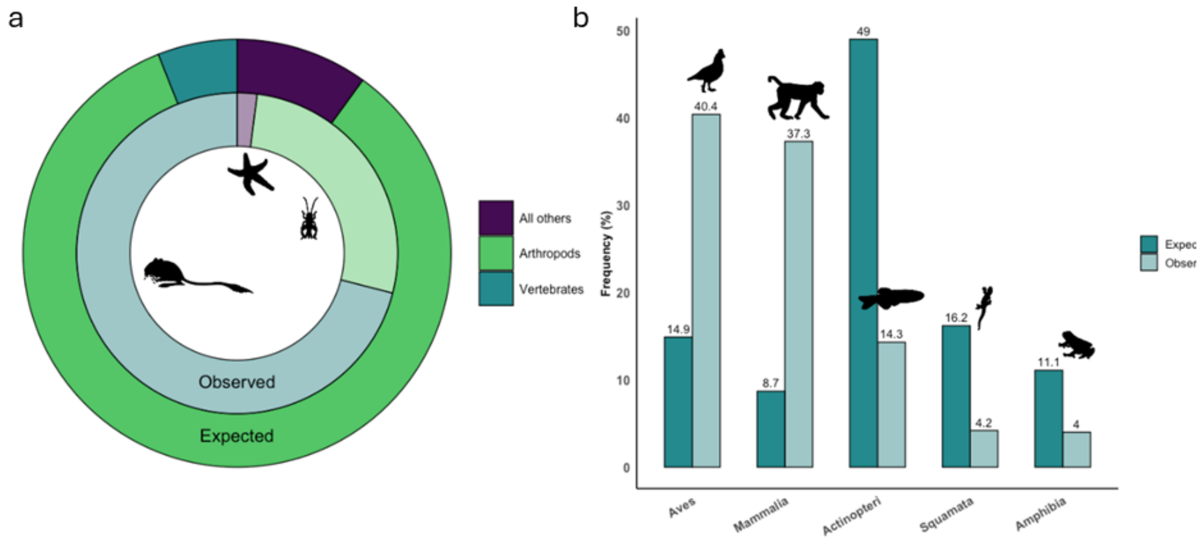

Figure 1. Taxonomic bias in the animals used in the study of animal behavior as identified by Rosenthal et al. (2017) both in general (a) and just within vertebrates (b).

Second, does taxonomic bias lead to misestimation and incorrect inferences in meta-analytical results? We answered this second question by reanalyzing data and then conducting a meta-meta-analysis to allow estimation of effects under taxonomically representative sampling.

Methods

To answer these two questions, we identified the meta-analyses published in three leading behavioral journals. We then determined the taxonomic representation of the constituent studies used in these meta-analyses and how estimates from these meta-analyses would change if the data were sampled in a taxonomically representative manner.

Identifying meta-analyses

To identify meta-analyses in animal behavior, we searched the journals Animal Behaviour, Behavioral Ecology and Sociobiology, and Behavioral Ecology for meta-analyses published in 2000 – 2024 using Web of Science (Core Collection, Science Citation Index Expanded). While behavioral meta-analyses are published elsewhere, we assumed that meta-analyses published in these journals would be generally representative of the field, if not broader in scope. We used the topic terms meta-anal*, meta, meta anal*. This search returned 118 articles and was conducted on 4 February 2025. We repeated this search on 23 September 2025 and added the topic terms meta-analysis, metaanal*, metaregres*, meta-regres*, quantitativ* review*, quantitative* synthe*, global* synthe*, and quantitativ* evidence synthe *, yielding a total of 136 articles. The titles and abstracts of these articles were then screened based on two search criteria: 1) the study had to have been identified by the authors as a meta-analysis, and 2) the study had to have been focused on non-human animals. We defined a meta-analysis as any analysis of effect sizes, even if standardized effect sizes were not analyzed. Based on these search criteria, we identified a total of 75 articles for secondary (full-text) screening (sensu Foo et al. 2021).

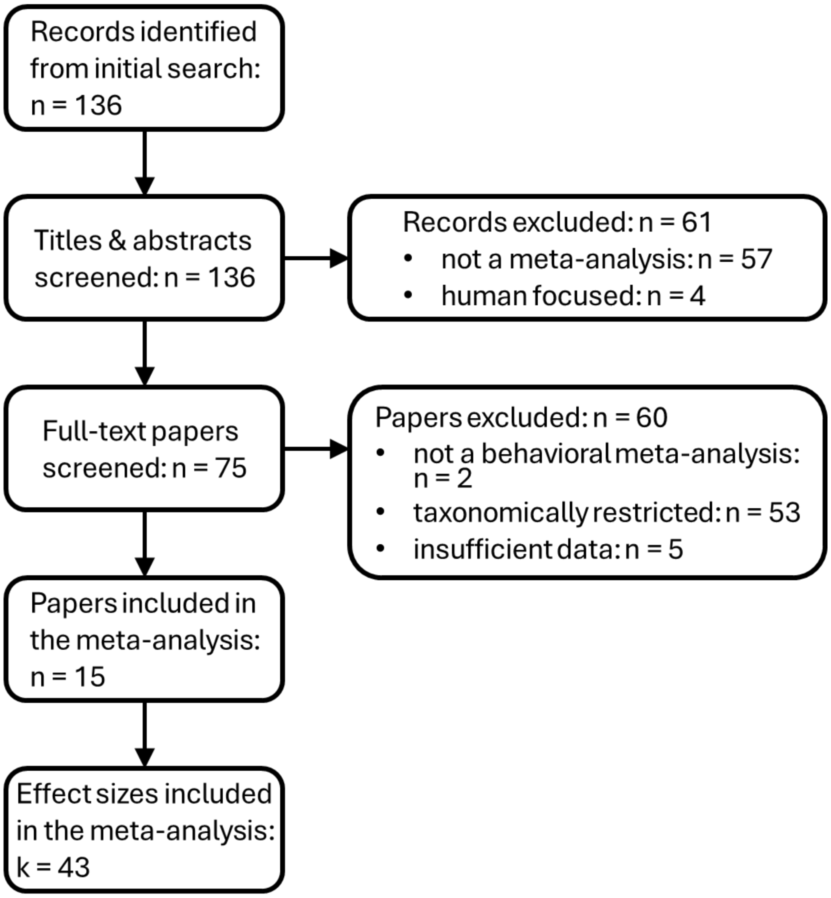

Figure 2. PRISMA diagram of included studies and included effect size estimates in the final analysis.

During the secondary screening we double-checked whether the 75 articles met our inclusion criteria. Next, the additional inclusion criteria at this stage were that the meta-analyses were not taxonomically restricted, defined here as including both vertebrates and invertebrates, with at least one invertebrate estimate, and that available data were sufficient to estimate effect size means for vertebrates and invertebrates. During secondary screening we also identified two articles that were meta-analyses but of research practices: reviews of publication biases and statistical power (Jennions and Møller 2003, Moller et al. 2005). These were excluded as they were not addressing questions of animal behavior per se. When data were not available from online sources, we requested the data directly from the authors. This ultimately led to a sample size of 15 articles and 43 estimates (Figure 2, Table 1). Title and abstract screening and secondary screening was done by NAD and MAS.

The 15 included articles covered a wide-range of behavioral topics. These topics included ornamentation patterns (Kraaijeveld et al. 2007), selection on “personality” and behavioral syndromes (Smith and Blumstein 2008), to winner-loser effects (Yan et al. 2024). All included articles are listed in Table 1.

Identifying taxonomic representation

For each published meta-analysis, we examined the data used to determine whether specific effect sizes were from invertebrates or vertebrates. From this, we determined the relative representation of invertebrates to vertebrates in the meta-analysis. We then compared this representation of estimates to that expected based on known animal diversity. Because of the expected lack of non-vertebrate estimates (Figure 1a) and the large number of non-vertebrate phyla, we compared the representation of species in published meta-analysis to the expected representation just at the level of vertebrates and invertebrates. We compared this representation based on taxonomic diversity without accounting for differences in abundance.

Estimating taxonomic misestimation of meta-analytical results

To determine the degree to which taxonomic representation could lead to misestimation of meta-analytical results, we compared meta-analytic grand means to estimates of what those means would be if estimates were drawn proportionally from the diversity of Animalia. We did so by calculating the marginalized means for each included meta-analysis under the assumption of taxonomically representative sampling (sensu Nakagawa et al. 2023).

For example, consider a meta-analysis based on 100 estimates with 75 estimates coming from vertebrates and 25 from invertebrates. This degree of taxonomic mis-representation is consistent with the findings of Rosenthal et al. (2017). Assuming estimates for vertebrates and invertebrates are equally precise despite different numbers of estimates, the grand mean effect size in this meta-analysis would be primarily driven by the mean for vertebrates: if the mean effect size for vertebrates were 0.5 and the mean for invertebrates was -0.5, the overall “grand mean” would be 0.25. However, this overall mean is taxonomically biased as vertebrates only represent around 5% of all animals (vertebrates: 4.74%; Catalogue of Life, 2025). If the estimates in this hypothetical example were drawn proportionally from the animal kingdom, the grand mean would instead have been estimated as -0.45. This grand mean is calculated by weighting by taxonomic representation rather than data representation.

Here, we first calculated the overall meta-analytical mean for a particular meta-analysis. Many of the constituent studies conducted multiple meta-analyses and we attempted to replicate as many of those as possible. In all cases we attempted to replicate the original analyses either using code originally provided by the authors or de novo code based on the described methods and for the effect sizes used or described by the authors. When possible, this included the inclusion of phylogenetic error structure. This provided an overall meta-analytical mean for the observed data (“data-proportional”). For each of these analyses we next added a moderator contrasting vertebrates and invertebrates. This provided separate meta-analytical estimates for vertebrates and invertebrates that incorporated the estimation uncertainty and sampling variance among studies. We also recorded the significance of this effect.

Based on the vertebrate and invertebrate meta-analytical estimates, we then estimated the marginalized mean under taxonomic representative sampling ("taxonomic-representative") using the metafor package in R (Viechtbauer 2010, Viechtbauer and López-López 2022, Nakagawa et al. 2023). This process is also known as "poststrafication" and is often used in recalibrating survey results (Gelman et al. 2021). Using the predict function of metafor, we estimated the marginalized taxonomic-representative value with an assumption of 5% of estimates being from vertebrates and 95% of estimates being from invertebrates. This representation is based on the known distribution of vertebrates and invertebrates (Catalogue of Life, 2025). This also allowed us to estimate the uncertainty around the taxonomic-representative mean (Viechtbauer and López-López 2022).

Marginalized means are calculated as described in the above example of weighted averaging but based instead on the meta-analytical group estimates. Specifically, the estimated vertebrate and invertebrate means, and accompanying uncertainties, were reweighted by the expected representation of invertebrates and vertebrates to calculate a taxonomic-representative mean. Because uncertainty around this new mean can be estimated, its statistical significance can also be determined.

For each meta-analysis we estimated the data-proportional (μ_{dp}) and taxonomic-representative (μ_{tr}) overall mean estimates and their uncertainties. We also determined the statistical significances ($\alpha < 0.05$) of μ_{dp} and μ_{tr}. Besides estimating magnitude of biasing, this allowed us to determine if taxonomic bias may lead to different inferential conclusions.

Data analysis

For each meta-analytical estimate pair from a study, we compared the data-proportional versus taxonomic-representative estimates to quantify the relative, proportional effect of taxonomic misrepresentation on overall effects in meta-analysis. To do so, we calculated the absolute value of the log ratio (Hedges et al. 1999):

$$\left| log\,\frac{\mu_{tr}}{\mu_{dp}} \right|$$

we used the absolute value as this allowed comparison across studies regardless of whether data-proportional or taxonomic-representative estimates were larger, which might vary by biological question. This also allowed us to compare estimates across meta-analyses even when those original meta-analyses used different effect sizes. The two limitations of this approach are that it does not allow analysis in cases where the signs of estimates change and the use of absolute values can result in over-estimation of magnitudes (Morrissey 2016, see below). Fortunately, sign changes only occurred for 3 of the 43 pairs of estimates (see results).

As an additional analysis, we also calculated the magnitude difference between data-proportional and taxonomic-representative estimates. Because the constituent meta-analyses used different effect sizes themselves (i.e. Zr, log-odds ratios, and various standardized mean differences), we first transformed all effect sizes to Zr following equations in Borenstein et al. (2009). We then calculated the raw value of the differences between data-proportional and taxonomic-representative Zr scores.

For both the relative and magnitude differences between pairs, the standard deviation of the differences was estimated as:

$$\sqrt{sd_{dp}^2 + sd_{tr}^2 - 2sd_{dp} \times sd_{tr} \times r}$$

where sd_{dp} and sd_{tr} are the estimate standard deviations of the overall meta-analytical means (reported in the metafor outputs as standard errors) and r represents the correlation between the uncertainties (because the uncertainties are based on the same datasets). For the differences in Zr scores, we used the estimated uncertainties in the above equation. For the absolute value of log ratios, sd_{dp} and sd_{tr} were calculated as $sd_{dp}/\ \mu_{dp}$ and $sd_{tr}/\ \mu_{tr}$, respectively. r is not analytically known but is between 0 and 1 and should be approaching one and so was set to 0.8. This did not substantively affect our conclusions or estimation (Supplemental Materials).

To estimate the overall effect of taxonomically biased sampling on meta-analytical inferences, we next fit a random effects meta-analysis with the study an estimate was drawn from as a random effect to the estimates of relative differences. We also included individual estimate identity as a random effect because each estimate was based on a different sample size and often addressing a different question. The overall effect size was estimated while weighting by the inverse of the standardized mean difference's sampling variance (Borenstein et al. 2009). We evaluated the magnitude and uncertainty of the grand mean (x_{gm}) from these analyses as indicative of taxonomic bias in meta-analytical inferences. Consequently, a significant grand mean would indicate significant taxonomic bias in meta-analytical estimates.

Importantly, calculating the grand mean of absolute values results in positive bias (Morrissey 2016). Therefore, we used an "analyze-then-transform" approach following Morrissey (2016). To do so, based on the above analysis, we transformed the meta-analytic grand means of the magnitude of taxonomic bias (x_{gm}) and its uncertainty ($\mathrm{se}_{\mathrm{gm}}$) according to a folded normal distribution as:

$$x_{folded} = se_{gm}\sqrt{\frac{2}{\pi}} \times e^{\left(\frac{-x_{gm}^2}{2se_{gm}^2}\right)} + x_{gm} \times \mathbf{erf}\left(\frac{x_{gm}}{\sqrt{2se_{gm}^2}}\right)$$

and

$$sd_{folded} = x_{gm}^2 + se_{gm}^2 + x_{folded}^2$$

where x_{folded} is the transformed grand mean and sd_{folded} is its standard deviation. **erf** is the Gauss error function. Uncertainties of parameters from linear models, like those provided by the metafor package, are reported as standard errors but these values specifically refer to the standard deviation of the parameter's sampling distribution and so were used here in lieu of standard deviations. We calculated the folded values for both the relative and absolute analyses. To aid interpretation, these values were then converted to percent differences as:

$$(e^{x_{folded}} - 1) \times 100$$

We next conducted a heterogeneity analyses to determine the contributors to variability in taxonomic bias. We determined whether there was significant heterogeneity among estimates based on Cochran's Q and its significance. We then calculated the proportion of variation (I^2) attributable to sampling error, to study differences, and to estimate differences within studies (Nakagawa and Santos 2012).

Because we were analyzing differences between marginal means, as opposed to original findings, there are not clear expectations of publication bias or approaches to testing for such bias. Moreover, because our estimates were constrained to be positive, typical funnel plots and trim-and-fill analyses were not appropriate. Given the lack of both a priori expectations regarding basis or methods, we did not conduct publication bias tests.

In addition, we conducted a series of post hoc analyses and comparisons. First, we conducted a meta-regression of taxonomic bias versus the number of studies in the original meta-analyses that were for invertebrates. A significant negative slope for this moderator would suggest that apparent taxonomic bias is at least partially driven by poor sampling of invertebrates. We then conducted a second post hoc meta-regression using the proportion of studies in a meta-analysis that were for invertebrates. A significant negative slope would be consistent with taxonomically biased sampling leading to misestimation (bias should decrease toward zero with increasing invertebrate representation). In these analyses, the number and proportion of studies were centered when used as covariates. These post hoc analyses therefore allowed us to estimate bias at average sampling levels. As a final post hoc addition, we compared the significance of the contrast between vertebrates and invertebrates as a moderator in underlying meta-analyses to whether significance of the overall means differed when estimated in a data-proportional versus taxonomic-representative manner.

All analyses were conducted in the R statistical language (v4.5.0, R_Core_Team 2025).

Data availability

All data and analysis code are available at: **link**.

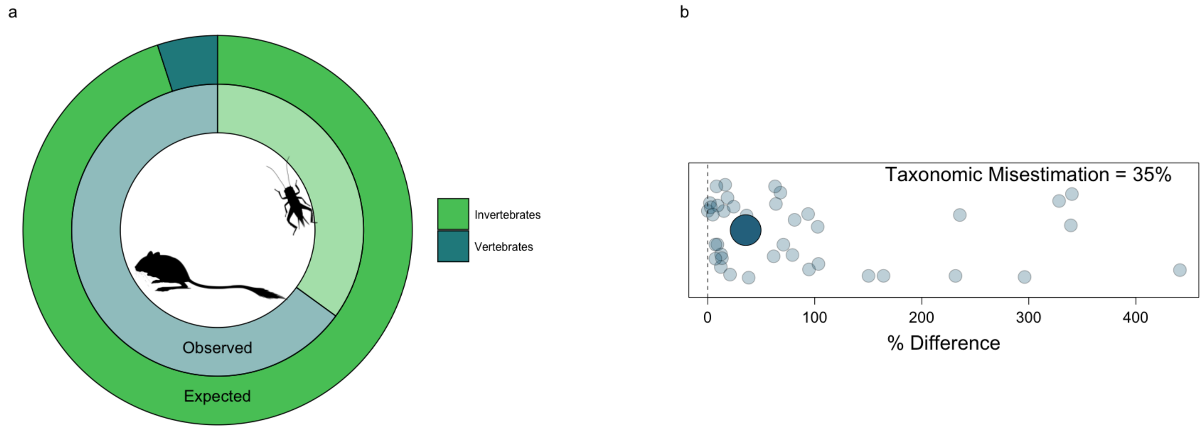

Figure 3. a) Data included in meta-analyses by taxa. Vertebrates are indicated in the same color as in Figure 1a while Arthropods and other invertebrates are now pooled under a single invertebrate category. Invertebrates are indicated in the same color as Arthropods in Figure 1a as Arthropod species dominated this category. b) Overall (large dot) and individual estimates (small dots) of the percentage misestimation due to taxonomic misrepresentation (uncertainty around this misestimation is smaller than the point).

Results

Taxonomic Representation of Data in Meta-Analyses

We found that the data used in behavioral meta-analyses was highly taxonomically biased in favor of vertebrates (Figure 3a, Table 1). 65% of the estimates used in meta-analyses were from vertebrates, slightly less than the bias observed more broadly in behavioral research (Rosenthal et al. 2017), but far, far greater than the 5% expected according to taxonomic representation.

Relative Effect of Taxonomic Misrepresentation on Meta-Analytical Inferences

The meta-analytical mean, after transformation due to the folded distribution, of absolute log ratios was 0.30 (s.e. = 0.08, $p << 0.01$; Figure 4). Put another way, meta-analytical means are, on average, misestimated by 35% (Figure 3b). Significance changed between the data-proportional and taxonomic-representative estimates for 10 of 43 estimate pairs (Figure 4), though the signs of effects only changed for 3 pairs. Of the significance changes, 8 instances changed from significant to non-significant and two changed from non-significant to significance. Moreover, while the vertebrate versus invertebrate contrast was significant in 19 of 43 constituent meta-analyses, it was significant for only 4 of the instances when overall mean significance changed.

There was also significant heterogeneity among estimates ($Q_{df:39} = 102.15$, $p << 0.01$). Of this heterogeneity, most was attributable to sampling variance ($I^2 = 0.69$). But there was also considerable heterogeneity among studies ($I^2 = 0.13$) and among estimates within studies ($I^2 = 0.19$).

Absolute Effect of Taxonomic Misrepresentation on Meta-Analytical Inferences

The meta-analytical mean of magnitude differences between Zr scores was 0.15 (s.e. = 0.06, $p = 0.01$). However, one estimate pair differed by around 3 times more than the next largest difference (Figure S1). When this estimate pair was excluded, the difference dropped to 0.098 (se = 0.026) and remained significant ($p < 0.01$). Whether with or without this extreme data point, these effects translate to substantive differences in more conventional effect sizes: a difference between Zr scores of 0.15 corresponds to a difference of 0.29 between correlations and a difference of 0.091 converts to a difference in correlations of 0.20. There was significant heterogeneity among estimates ($Q_{df:41} = 2.1\times 10^5$, $p << 0.01$) but 99% of this heterogeneity was attributable to sampling variance with no appreciable heterogeneity among studies or estimates ($I^2 < 0.01$).

Meta-regression for the effect of sampling on estimation

Both the number and proportion of invertebrate studies in a constituent meta-analysis had a significant negative effect on the magnitude of taxonomic bias ($p = 0.03$ and <0.01, respectively). The significant effect of number of studies suggests that some proportion of the taxonomic bias is influenced by poor estimation at small sample sizes but the magnitude of this effect was small ($\beta = -0.0012$). Interestingly, the estimated magnitude of taxonomic bias was higher at the average invertebrate sample size than estimated in our main analysis (39%). The magnitude of the effect of proportional invertebrate representation was larger ($\beta = -0.684$), consistent with taxonomic bias in sampling biasing meta-analytical conclusions.

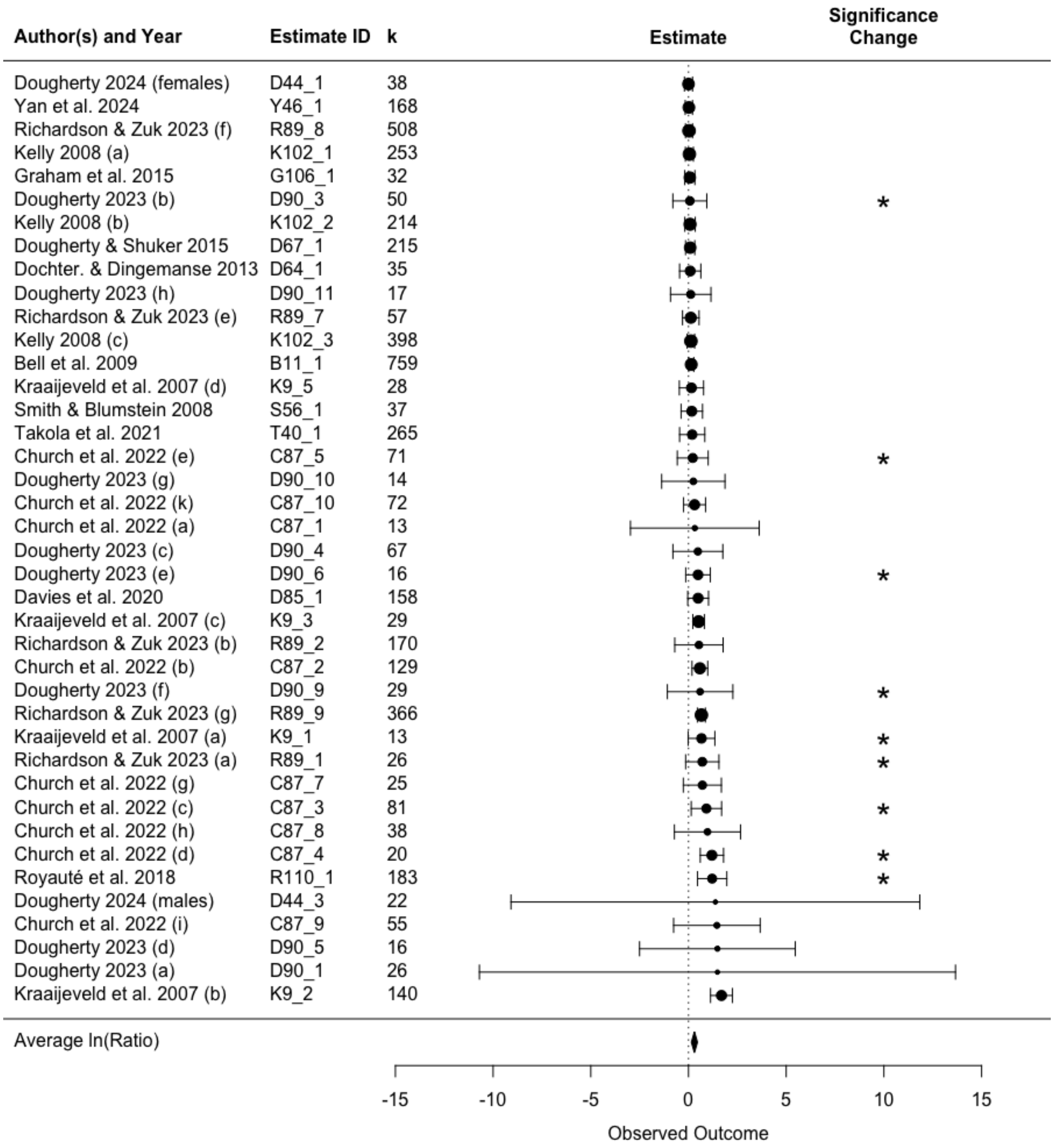

Figure 4. Within-study estimates and standard errors of the relative magnitude of taxonomic bias on meta-analytical estimates.

Discussion

We found that the impact of taxonomic bias on meta-analytical estimates in the study of animal behavior was surprisingly large (Figure 3b). Specifically, our results demonstrate that meta-analytical means from behavioral meta-analysis are misestimated by around 35%. The observed magnitude of misestimation also led to changes in significance for around 25% of estimates. Interestingly, the significance of a moderator allowing for differences between invertebrates and vertebrates within underlying meta-analyses was not particularly informative: In more than half of the instances when significance changed under taxonomic representative estimation, a moderator for taxonomic grouping was not significant. This estimated magnitude of misestimation, 35%, and its impact on direction or significance of effects can lead to incorrect inferences being drawn.

Our analysis of the magnitude of taxonomic bias on the raw effect of meta-analytical estimates demonstrates a somewhat smaller impact. However, this is not entirely unexpected because effect sizes estimated in meta-analyses are typically quite small (Møller and Jennions 2002, Low-Décarie et al. 2014) and so differences in magnitudes will necessarily be constrained to also be small. Even in terms of magnitude differences, the observed bias still translates into substantial difference in the magnitudes of effects, changing correlations by an average of 0.3.

As one example, consider the findings of Royauté et al. (2018). We use this example as one of us (NAD) was the senior author of that study and so as to not single out other authors. In Royauté et al. (2018), the authors concluded that correlations between behaviors and physiological or life-history traits were not in the direction expected according to the "pace-of-life" syndrome (POLS) hypothesis proposed by Réale et al. (2010). This was a novel and surprising finding given the intuitive predictions of POLS and its dramatic impact on discussions of how behavior might be integrated with physiology and life-history. The conclusion of Royauté et al. (2018) was based on the meta-analytical mean, estimated as r = 0.06 (95% CI: -.01 : 0.14), which was predicted to have been positive according to POLS. However, as reported by Royauté et al. (2018), there was a significant difference between vertebrates and invertebrates regrading support for POLS. There also was substantial taxonomic bias in the data that went into Royauté et al.'s (2018) analysis. Here, via the estimation of marginal means adjusting for taxonomic bias, we see that if the sampling of vertebrates and invertebrates was taxonomically proportional, the meta-analytical mean would have been estimated as r = 0.22 and would have been significantly different than zero. While this is a relatively modest absolute difference in estimates (0.16), it is a large relative difference (~2.4 times, Figure 4). This ends up resulting in a substantive reinterpretation of the evidence for POLS: If the data going into the original analysis were consistent with actual taxonomic representation, and the estimates of the original analysis hold, this would have been taken as meta-analytical support for POLS. Many of the included meta-analyses also drew conclusions based on overall mean effect sizes and so similar issues will arise across our dataset. Consequently, major changes in inferences would likely be drawn from other meta-analyses if taxonomic bias were to be similarly addressed, as demonstrated by the observed changes in significance (Figure 4). Likewise, changes in the sign of three mean effect sizes would also lead to inferential changes.

This example, and our overall finding, confirms concerns raised by others. In particular, Gurevitch and Hedges (1999) discussed how researcher biases and taxonomic preferences in the generation of original data have the potential to affect subsequent meta-analytical inferences. However, the degree to which this was an actual problem was not clear until our analysis here. This is a concerning problem but not a new one: as discussed by Rosenthal et al. (2017), taxonomic bias in the study of animal behavior is pervasive and pronounced (their data are recreated in Figures 1a & 1b).

Importantly, changes in interpretation as discussed above are dependent on how robust our marginal mean estimates are for invertebrates. This is not currently clear and, across the included meta-analyses, the poor representation of invertebrates (Figure 3a, Table 1) necessarily means that these estimates have greater uncertainties than do estimates for vertebrates. Our post hoc meta-regressions suggest that this likely has only a minor effect on our interpretation, as demonstrated by the magnitude of the regression coefficient relating the number of invertebrate studies to the magnitude of bias. Moreover, this post hoc analysis suggests that our primary analysis may be underestimating the magnitude of taxonomic bias. Also, because of the poor representation of invertebrates in behavioral studies, the diversity *within* invertebrates could not be captured in our analysis. Because invertebrates are also not proportionally sampled (Zuk et al. 2014), our field's meta-analytical conclusions may be even more biased than could be revealed here.

While changes in interpretation due to taxonomic bias might be discouraging, they also are of intrinsic interest. Sub-group analysis and stratification (Borenstein et al. 2009), long-standing parts of meta-analyses in other fields, would allow the evaluation of when a hypothesis like POLS holds in some taxonomic groups and not others. For example, it is perhaps more interesting to ask why POLS holds in invertebrates but not vertebrates rather than simply asking whether the hypothesis is supported in general. Such analyses may allow for greater insight than currently provided by most behavioral meta-analyses. Indeed, researchers in areas of ecology and evolution are often interested specifically in when and why groups, like vertebrates and invertebrates, differ (Nakagawa et al. 2017, Yang et al. 2025). Disentangling contributors to heterogeneity in effect sizes will often be more informative than simple estimations of overall means.

It is also worth emphasizing that relatively few studies could be included in our analyses and, as a result, the effects of taxonomic bias might therefore be even more severe. Specifically, 53 meta-analyses published in Animal Behaviour, Behavioral Ecology, and Behavioral Ecology and Sociobiology were taxonomically restricted (Figure 2). Put another way, the taxonomic bias in the 15 studies included here necessarily underestimates the bias present in the 53 studies for which bias could not even be estimated. Indeed, in examining the taxonomically restricted meta-analyses, over 75% were restricted to vertebrates (Table S1). Echoing—and exaggerating—the biases within vertebrates identified by Rosenthal et al. (2017 and Figure 1b), 87% of the constituent studies used in vertebrate restricted meta-analyses were of birds (Table S1).

In some cases, taxonomic restriction may be because a specific hypothesis applies to, for example, mammals but not reptiles. In such cases, the type of taxonomic restriction we have identified is a requirement to ask a question and is to be expected. However, it is also possible that this represents a secondary opportunity for researchers to express taxonomic preferences (Gurevitch and Hedges 1999). This is, potentially, reflected in the taxonomic representation seen in the restricted studies. As mentioned above, most of these focused on birds and mammals. Even if taxonomic restriction is justified, that most restricted meta-analyses focus on birds and mammals results in research effort that does not reflect the broader taxonomic representation. This, again, biases our understanding of animal behavior. Most generally, taxonomically restricted meta-analyses, when not necessary, will reduce the generality of conclusions.

If a broad goal of animal behavior research is to increase our general understanding of behavior, our results demonstrate that taxonomic bias should be a major concern. If, instead, we are primarily interested in examining personally interesting examples of behavior, the bias is of less importance. Regardless, a primary aim of meta-analyses *is* to draw generalizable inferences (Spake et al. 2022). Our results suggest that the taxonomic bias inherent in the behavioral literature makes this generalization difficult. This may profoundly affect how right and wrong we are about our understanding of behavior. Addressing this bias will require individual researchers to consider whether their study system choices will increase the generality of our field's understanding.

Acknowledgements

We thank the authors of behavioral meta-analyses for providing data either via public repositories or upon request. MAS was supported by the Department of Biological Sciences and the Environmental and Conservation Sciences Program at North Dakota State University. NAD was supported during this work by the National Science Foundation (IOS-2222929). SN was supported by the Canada Excellence Research Chair Program (CERC-2022-00074).

References

Bell, A. M., S. J. Hankson, and K. L. Laskowski. 2009. The repeatability of behaviour: a meta-analysis. Animal Behaviour **77**:771-783.
Borenstein, M., L. V. Hedges, J. P. Higgins, and H. R. Rothstein. 2009. Introduction to meta-analysis. John wiley & sons.
Church, K. D., J.-M. Matte, and J. W. Grant. 2022. Territoriality modifies the effects of habitat complexity on animal behavior: a meta-analysis. Behavioral Ecology **33**:455-466.
Costello, L., and J. W. Fox. 2022. Decline effects are rare in ecology. Ecology **103**:e3680.
Davies, A. D., Z. Lewis, and L. R. Dougherty. 2020. A meta-analysis of factors influencing the strength of mate-choice copying in animals. Behavioral Ecology **31**:1279-1290.

Dickersin, K. 2005. Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. Publication bias in meta‐analysis: Prevention, assessment and adjustments:9-33.

Dochtermann, N. A., and N. J. Dingemanse. 2013. Behavioral syndromes as evolutionary constraints. Behavioral Ecology **24**:806-811.

Dougherty, L. R. 2023. The effect of individual state on the strength of mate choice in females and males. Behavioral Ecology **34**:197-209.

Dougherty, L. R. 2024. Mating reduces responsiveness to sexual stimuli in females but not in males. Animal Behaviour **214**:87-94.

Dougherty, L. R., and D. M. Shuker. 2015. The effect of experimental design on the measurement of mate choice: a meta-analysis. Behavioral Ecology **26**:311-319.

Foo, Y. Z., R. E. O'Dea, J. Koricheva, S. Nakagawa, and M. Lagisz. 2021. A practical guide to question formation, systematic searching and study screening for literature reviews in ecology and evolution. Methods in Ecology and Evolution **12**:1705-1720.

Gelman, A., J. Hill, and A. Vehtari. 2021. Regression and other stories. Cambridge University Press.

Graham, S., E. Chapuis, S. Meconcelli, N. Bonel, K. Sartori, A. Christophe, P. Alda, P. David, and T. Janicke. 2015. Size-assortative mating in simultaneous hermaphrodites: an experimental test and a meta-analysis. Behavioral Ecology and Sociobiology **69**:1867-1878.

Gurevitch, J., and L. V. Hedges. 1999. Statistical issues in ecological meta‐analyses. Ecology **80**:1142-1149.

Hedges, L. V., J. Gurevitch, and P. S. Curtis. 1999. The meta‐analysis of response ratios in experimental ecology. Ecology **80**:1150-1156.

Jennions, M. D., C. J. Lortie, M. S. Rosenberg, and H. R. Rothstein. 2013. Publication and related biases. Handbook of meta-analysis in ecology and evolution:207-236.

Jennions, M. D., and A. P. Møller. 2002. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. Proceedings of the Royal Society of London. Series B: Biological Sciences **269**:43-48.

Jennions, M. D., and A. P. Møller. 2003. A survey of the statistical power of research in behavioral ecology and animal behavior. Behavioral Ecology **14**:438-445.

Kelly, C. D. 2008. The interrelationships between resource-holding potential, resource-value and reproductive success in territorial males: how much variation can we explain? Behavioral Ecology and Sociobiology **62**:855-871.

Konno, K., M. Akasaka, C. Koshida, N. Katayama, N. Osada, R. Spake, and T. Amano. 2020. Ignoring non‐English‐language studies may bias ecological meta‐analyses. Ecology and Evolution **10**:6373-6384.

Kraaijeveld, K., F. J. Kraaijeveld-Smit, and J. Komdeur. 2007. The evolution of mutual ornamentation. Animal Behaviour **74**:657-677.

Low-Décarie, E., C. Chivers, and M. Granados. 2014. Rising complexity and falling explanatory power in ecology. Frontiers in Ecology and the Environment **12**:412-418.

Møller, A., and M. D. Jennions. 2002. How much variance can be explained by ecologists and evolutionary biologists? Oecologia **132**:492-500.

Moller, A. P., R. Thornhill, and S. W. Gangestad. 2005. Direct and indirect tests for publication bias: asymmetry and sexual selection. Animal Behaviour **70**:497-506.

Morrissey, M. B. 2016. Meta ‐ analysis of magnitudes, differences and variation in evolutionary parameters. Journal of Evolutionary Biology **29**:1882-1904.

Nakagawa, S., M. Lagisz, M. D. Jennions, J. Koricheva, D. W. Noble, T. H. Parker, A. Sánchez ‐ Tójar, Y. Yang, and R. E. O'Dea. 2022. Methods for testing publication bias in ecological and evolutionary meta ‐ analyses. Methods in Ecology and Evolution **13**:4-21.

Nakagawa, S., M. Lagisz, R. E. O'Dea, P. Pottier, J. Rutkowska, A. M. Senior, Y. Yang, and D. W. Noble. 2023. orchaRd 2.0: An R package for visualising meta-analyses with orchard plots.

Nakagawa, S., D. W. Noble, A. M. Senior, and M. Lagisz. 2017. Meta-evaluation of meta-analysis: ten appraisal questions for biologists. Bmc Biology **15**:18.

Nakagawa, S., and E. S. Santos. 2012. Methodological issues and advances in biological meta-analysis. Evolutionary Ecology **26**:1253-1274.

Pollo, P., M. Lagisz, Y. Yang, A. Culina, and S. Nakagawa. 2024. Synthesis of sexual selection: a systematic map of meta ‐ analyses with bibliometric analysis. Biological Reviews **99**:2134-2175.

R_Core_Team. 2025. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Réale, D., D. Garant, M. M. Humphries, P. Bergeron, V. Careau, and P. O. Montiglio. 2010. Personality and the emergence of the pace-of-life syndrome concept at the population level. Philosophical Transactions of the Royal Society B-Biological Sciences **365**:4051-4063.

Richardson, J., and M. Zuk. 2023. Unlike a virgin: a meta-analytical review of female mating status in studies of female mate choice. Behavioral Ecology **34**:165-182.

Rosenthal, M. F., M. Gertler, A. D. Hamilton, S. Prasad, and M. C. Andrade. 2017. Taxonomic bias in animal behaviour publications. Animal Behaviour **127**:83-89.

Rothstein, H. R., A. J. Sutton, and M. Borenstein. 2005. Publication bias in meta ‐ analysis. Publication bias in meta ‐ analysis: Prevention, assessment and adjustments:1-7.

Royauté, R., M. A. Berdal, C. R. Garrison, and N. A. Dochtermann. 2018. Paceless life? A meta-analysis of the pace-of-life syndrome hypothesis. Behavioral Ecology and Sociobiology **72**:64.

Smith, B. R., and D. T. Blumstein. 2008. Fitness consequences of personality: a meta-analysis. Behavioral Ecology **19**:448-455.

Spake, R., R. E. O'dea, S. Nakagawa, C. P. Doncaster, M. Ryo, C. T. Callaghan, and J. M. Bullock. 2022. Improving quantitative synthesis to achieve generality in ecology. Nature ecology & evolution **6**:1818-1828.

Takola, E., E. T. Krause, C. Müller, and H. Schielzeth. 2021. Novelty at second glance: a critical appraisal of the novel object paradigm based on meta-analysis. Animal Behaviour **180**:123-142.

Tinbergen, N. 1963. On aims and methods of ethology. Zeitschrift fur Tierpsychologie **20**:410-433.

Trikalinos, T. A., and J. P. Ioannidis. 2005. Assessing the evolution of effect sizes over time. Publication bias in meta‐analysis: Prevention, assessment and adjustments:241-259.

Viechtbauer, W. 2010. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software **36**.

Viechtbauer, W., and J. A. López‐López. 2022. Location‐scale models for meta‐analysis. Research Synthesis Methods **13**:697-715.

Yan, J. L., N. M. Smith, D. C. Filice, and R. Dukas. 2024. Winner and loser effects: a meta-analysis. Animal Behaviour **216**:15-22.

Yang, Y., D. W. Noble, R. Spake, A. M. Senior, M. Lagisz, and S. Nakagawa. 2025. A pluralistic framework for measuring, interpreting and decomposing heterogeneity in meta‐analysis. Methods in Ecology and Evolution.

Zuk, M., F. Garcia-Gonzalez, M. E. Herberstein, and L. W. Simmons. 2014. Model systems, taxonomic bias, and sexual selection: beyond Drosophila. Annual Review of Entomology **59**:321-338.

Table 1. Behavioral meta-analyses included in the current analyses

Study ID	Number of included analyses	Number of estimates in each analysis	Number of estimates for vertebrates	Number of estimates for invertebrates	Author(s) (Year)
K9	4	13	4	9	Kraaijeveld et al. (2007)
		140	135	5	
		29	26	3	
		28	23	5	
B11	1	759	493	266	Bell et al. (2009)
T40	1	265	261	4	Takola et al. (2021)
D44	3	38	2	36	Dougherty (2024)
		22	4	18	
Y46	1	168	47	121	Yan et al. (2024)
S56	1	37	31	6	Smith and Blumstein (2008)
D64	1	35	30	5	Dochtermann and Dingemanse (2013)
D67	1	215	92	123	Dougherty and Shuker (2015)
D85	1	158	125	33	Davies et al. (2020)
C87	10	13	12	1	Church et al. (2022)
		129	85	44	
		81	74	7	
		20	19	1	
		71	56	15	
		26	25	1	
		25	22	3	
		38	35	3	
		55	49	6	
		72	29	43	
R89	7	26	9	17	Richardson and Zuk (2023)
		170	61	109	
		52	24	28	
		8	4	4	
		57	43	14	
		508	92	416	
		366	325	41	
D90	8	26	8	18	Dougherty (2023)
		50	27	23	
		67	41	26	
		16	6	10	
		13	13	3	
		29	22	7	

		14	5	9	
		17	6	11	
		253	187	66	
K102	3	214	168	46	Kelly (2008)
		398	339	59	
G106	1	32	3	29	Graham et al. (2015)
R110	1	183	145	38	Royauté et al. (2018)

Supplementary Materials

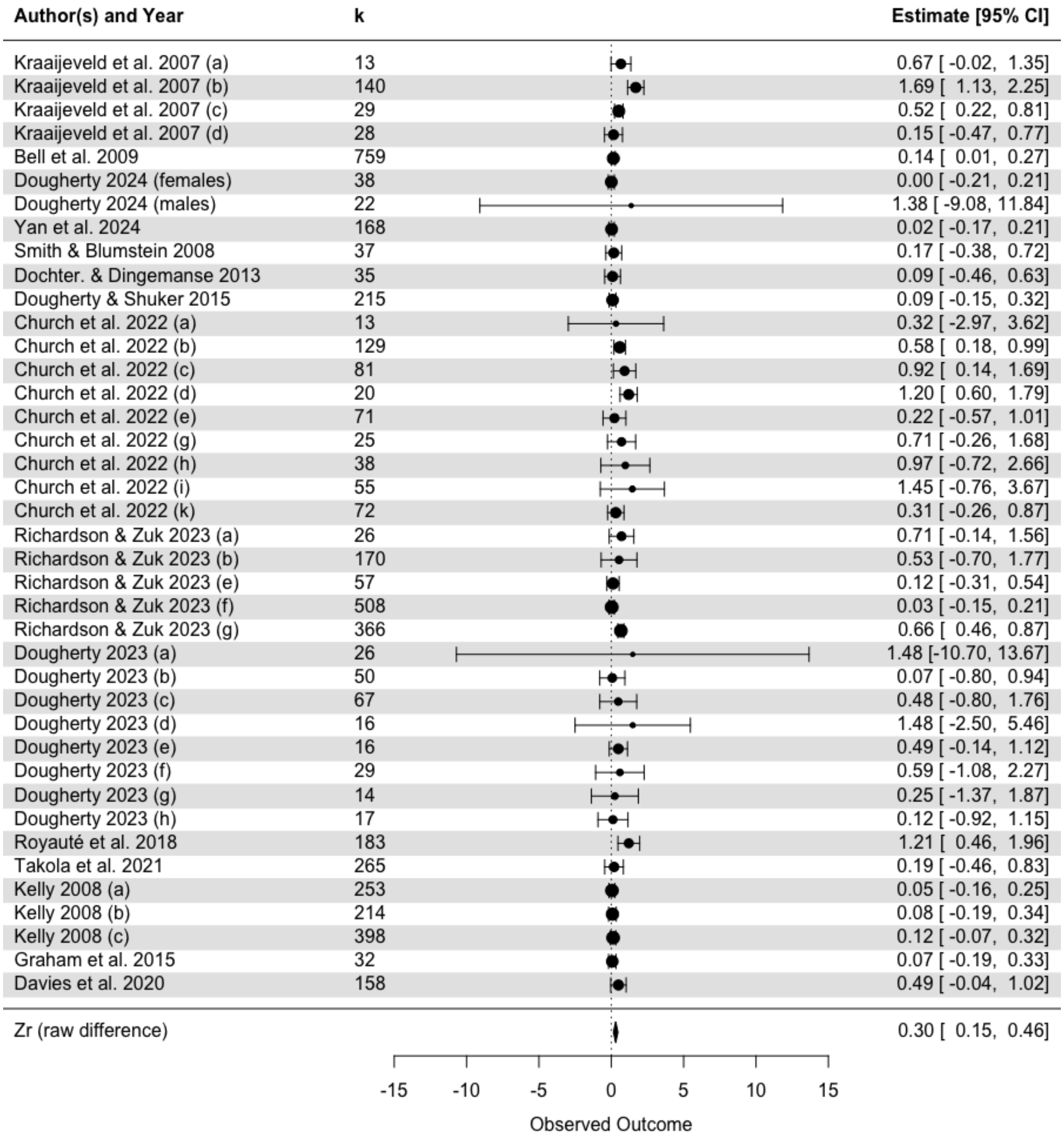

Figure S1. Within-study estimates and standard errors of the magnitude of taxonomic bias (magnitude differences) on meta-analytical estimates.

Table S1. Taxa included in taxonomically restricted meta-analyses.

Article ID	Phylum	Class	Number of included studies	Reference
1	Arthropoda	Insecta	122	Arnqvist & Nilsson (2000) Anim. Behav.
2	Chordata	Mammalia	27	Schino (2001) Anim. Behav.
3	Chordata	Mammalia	5	Roberts et al. (2004) Anim. Behav.
3	Chordata	Reptilia	6	Roberts et al. (2004) Anim. Behav.
3	Chordata	Aves	25	Roberts et al. (2004) Anim. Behav.
4	Arthropoda	Insecta	7	Shuker et al. (2004) Anim. Behav.
6	Chordata	Reptilia	5	Hirschenhauser & Oliveira (2006) Anim. Behav.
6	Chordata	Amphibia	9	Hirschenhauser & Oliveira (2006) Anim. Behav.
6	Chordata	Pisces	26	Hirschenhauser & Oliveira (2006) Anim. Behav.
6	Chordata	Mammalia	38	Hirschenhauser & Oliveira (2006) Anim. Behav.
6	Chordata	Aves	65	Hirschenhauser & Oliveira (2006) Anim. Behav.
7	Chordata	Aves	12	Griffith et al. (2006) Anim. Behav.
7	Chordata	Aves	27	Griffith et al. (2006) Anim. Behav.
10	Chordata	Mammalia	54	Majolo et al. (2008) Anim. Behav.
10	Chordata	Mammalia	86	Majolo et al. (2008) Anim. Behav.
12	Chordata	Aves	24	Sridhar et al. (2009) Anim. Behav.
12	Chordata	Aves	27	Sridhar et al. (2009) Anim. Behav.
12	Chordata	Aves	66	Sridhar et al. (2009) Anim. Behav.
15	Chordata	Aves	127	Santos et al. (2011) Anim. Behav.
16	Chordata	Aves	11	Hasselquist & Nilsson (2012) Anim. Behav.
16	Chordata	Aves	14	Hasselquist & Nilsson (2012) Anim. Behav.
20	Arthropoda	Insecta	1	Paterson et al. (2013) Anim. Behav.
20	Arthropoda	Malacostraca	6	Paterson et al. (2013) Anim. Behav.
20	Arthropoda	Insecta	6	Paterson et al. (2013) Anim. Behav.
20	Mollusca	Gastropoda	12	Paterson et al. (2013) Anim. Behav.
20	Arthropoda	Insecta	14	Paterson et al. (2013) Anim. Behav.
24	Chordata	Mammalia	7	Street et al. (2016) Anim. Behav.
24	Chordata	Mammalia	26	Street et al. (2016) Anim. Behav.
26	Chordata	Aves	14	Kriengwatana et al. (2016) Anim. Behav.
28	Chordata	Aves	555	Wood et al. (2017) Anim. Behav.
32	Chordata	Aves	379	Parker et al. (2018) Anim. Behav.
34	Chordata	Mammalia	14	Amici et al. (2019) Anim. Behav.
34	Chordata	Aves	11	Amici et al. (2019) Anim. Behav.
34	Chordata	Actinopterygii	1	Amici et al. (2019) Anim. Behav.
37	Chordata	Reptilia	1	Penndorf & Aplin (2020) Anim. Behav.
37	Chordata	Aves	5	Penndorf & Aplin (2020) Anim. Behav.
37	Chordata	Mammalia	10	Penndorf & Aplin (2020) Anim. Behav.
43	Chordata	Aves	142	Shuai et al. (2024) Anim. Behav.
45	Chordata	Aves	1001	Petalas et al. (2024) Anim. Behav.

49	Chordata	Aves	90	Garamszegi & Moller (2004) Behav. Ecol.
49	Chordata	Aves	121	Garamszegi & Moller (2004) Behav. Ecol.
49	Chordata	Aves	133	Garamszegi & Moller (2004) Behav. Ecol.
49	Chordata	Aves	396	Garamszegi & Moller (2004) Behav. Ecol.
50	Chordata	Mammalia	30	Schino (2004) Behav. Ecol.
52	Chordata	Aves	89	Parker et al. (2006) Behav. Ecol.
53	Chordata	Mammalia	31	Schino (2007) Behav. Ecol.
53	Chordata	Mammalia	36	Schino (2007) Behav. Ecol.
54	Chordata	Aves	149	Nakagawa et al. (2007) Behav. Ecol.
54	Chordata	Aves	182	Nakagawa et al. (2007) Behav. Ecol.
54	Chordata	Aves	216	Nakagawa et al. (2007) Behav. Ecol.
54	Chordata	Aves	246	Nakagawa et al. (2007) Behav. Ecol.
54	Chordata	Aves	331	Nakagawa et al. (2007) Behav. Ecol.
54	Chordata	Aves	473	Nakagawa et al. (2007) Behav. Ecol.
54	Chordata	Aves	602	Nakagawa et al. (2007) Behav. Ecol.
54	Chordata	Aves	1236	Nakagawa et al. (2007) Behav. Ecol.
55	Arthropoda	Insecta	7	Meunier et al. (2008) Behav. Ecol.
55	Arthropoda	Insecta	10	Meunier et al. (2008) Behav. Ecol.
55	Arthropoda	Insecta	15	Meunier et al. (2008) Behav. Ecol.
57	Chordata	Aves	11	Beauchamp (2008) Behav. Ecol.
57	Chordata	Aves	18	Beauchamp (2008) Behav. Ecol.
57	Chordata	Aves	20	Beauchamp (2008) Behav. Ecol.
57	Chordata	Aves	25	Beauchamp (2008) Behav. Ecol.
57	Chordata	Aves	29	Beauchamp (2008) Behav. Ecol.
57	Chordata	Aves	35	Beauchamp (2008) Behav. Ecol.
57	Chordata	Aves	43	Beauchamp (2008) Behav. Ecol.
60	Chordata	Aves	75	Soma & Garamszegi (2011) Behav. Ecol.
62	Chordata	Aves	8	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	10	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	11	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	14	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	19	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	21	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	24	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	24	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	25	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	26	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	33	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	48	Simons & Verhulst (2011) Behav. Ecol.
62	Chordata	Aves	77	Simons & Verhulst (2011) Behav. Ecol.
65	Chordata	Aves	32	Garamszegi et al. (2013) Behav. Ecol.
65	Chordata	Actinopterygii	18	Garamszegi et al. (2013) Behav. Ecol.
65	Chordata	Mammalia	14	Garamszegi et al. (2013) Behav. Ecol.
65	Chordata	Reptilia	3	Garamszegi et al. (2013) Behav. Ecol.
66	Chordata	Aves	9	Ihle & Forstmeier (2013) Behav. Ecol.
68	Chordata	Aves	43	Arct et al. (2015) Behav. Ecol.
70	Chordata	Mammalia	8	Moore et al. (2016) Behav. Ecol.
70	Chordata	Amphibia	14	Moore et al. (2016) Behav. Ecol.

70	Chordata	Reptilia	21	Moore et al. (2016) Behav. Ecol.
70	Chordata	Aves	75	Moore et al. (2016) Behav. Ecol.
73	Chordata	Amphibia	22	Roca et al. (2016) Behav. Ecol.
73	Chordata	Aves	139	Roca et al. (2016) Behav. Ecol.
77	Arthropoda	Insecta	117	Holman (2018) Behav. Ecol.
82	Chordata	Aves	8	Santema et al. (2019) Behav. Ecol.
82	Chordata	Aves	8	Santema et al. (2019) Behav. Ecol.
84	Arthropoda	Arachnida	23	Ximenes et al. (2020) Behav. Ecol.
86	Chordata	Mammalia	296	Beauchamp et al. (2021) Behav. Ecol.
88	Chordata	Amphibia	10	Stuber et al. (2022) Behav. Ecol.
88	Chordata	Mammalia	14	Stuber et al. (2022) Behav. Ecol.
88	Chordata	Reptilia	18	Stuber et al. (2022) Behav. Ecol.
88	Chordata	Pisces	50	Stuber et al. (2022) Behav. Ecol.
88	Chordata	Aves	108	Stuber et al. (2022) Behav. Ecol.
93	Chordata	Mammalia	153	Huang et al. (2024) Behav. Ecol.
94	Arthropoda	Insecta	59	Kochensparger et al. (2024) Behav. Ecol.
95	Chordata	Aves	1333	Dubois & Cézilly (2002) Behav. Ecol. Socio.
96	Chordata	Amphibia	62	Rohr et al. (2002) Behav. Ecol. Socio.
97	Arthropoda	Insecta	60	Torres-Vila et al. (2004) Behav. Ecol. Socio.
98	Arthropoda	Insecta	29	Torres-Vila & Jennions (2005) Behav. Ecol. Socio.
99	Chordata	Aves	87	Garamszegi (2005) Behav. Ecol. Socio.
99	Chordata	Aves	102	Garamszegi (2005) Behav. Ecol. Socio.
99	Chordata	Aves	140	Garamszegi (2005) Behav. Ecol. Socio.
99	Chordata	Aves	240	Garamszegi (2005) Behav. Ecol. Socio.
100	Chordata	Aves	3	Verdolin (2006) Behav. Ecol. Socio.
100	Chordata	Mammalia	29	Verdolin (2006) Behav. Ecol. Socio.
101	Chordata	Aves	82	Thorup & Rabol (2007) Behav. Ecol. Socio.
103	Chordata	Aves	19	Meunier et al. (2011) Behav. Ecol. Socio.
104	Chordata	Mammalia	180	Spinka et al. (2011) Behav. Ecol. Socio.
105	Arthropoda	Insecta	68	Wilson-Rankin (2014) Behav. Ecol. Socio.
108	Arthropoda	Insecta	5	Dougherty & Shuker (2016) Behav. Ecol. Socio.
117	Chordata	Amphibia	184	Lee et al. (2022) Behav. Ecol. Socio.
131	Arthropoda	Insecta	31	Nieberding & Holveck (2017) Behav. Ecol. Socio

Supplementary Analysis: Effect of *r* on sampling variance.

Because we estimated the standard deviation of estimates as:

$$\sqrt{sd_{dp}^2 + sd_{tr}^2 - 2sd_{dp} \times sd_{tr} \times r}$$

it was possible that choices about the value of *r* might influence the ultimate meta-meta-analytical results. We know that *r* is much greater than 0 since the same data were being used for the data-proportional and taxonomic-representative estimates. However, we don't know if this value is exactly 1. We *a priori* decided to use a value of 0.8 since the value is necessarily large but also calculated meta-analytical means over a range of values for *r* (Figure S2). A higher *r*, corresponding to a more conservative estimte of the standard deviation, led to an increase in the meta-analytical estimate but did not change significance or our overall interpretation.

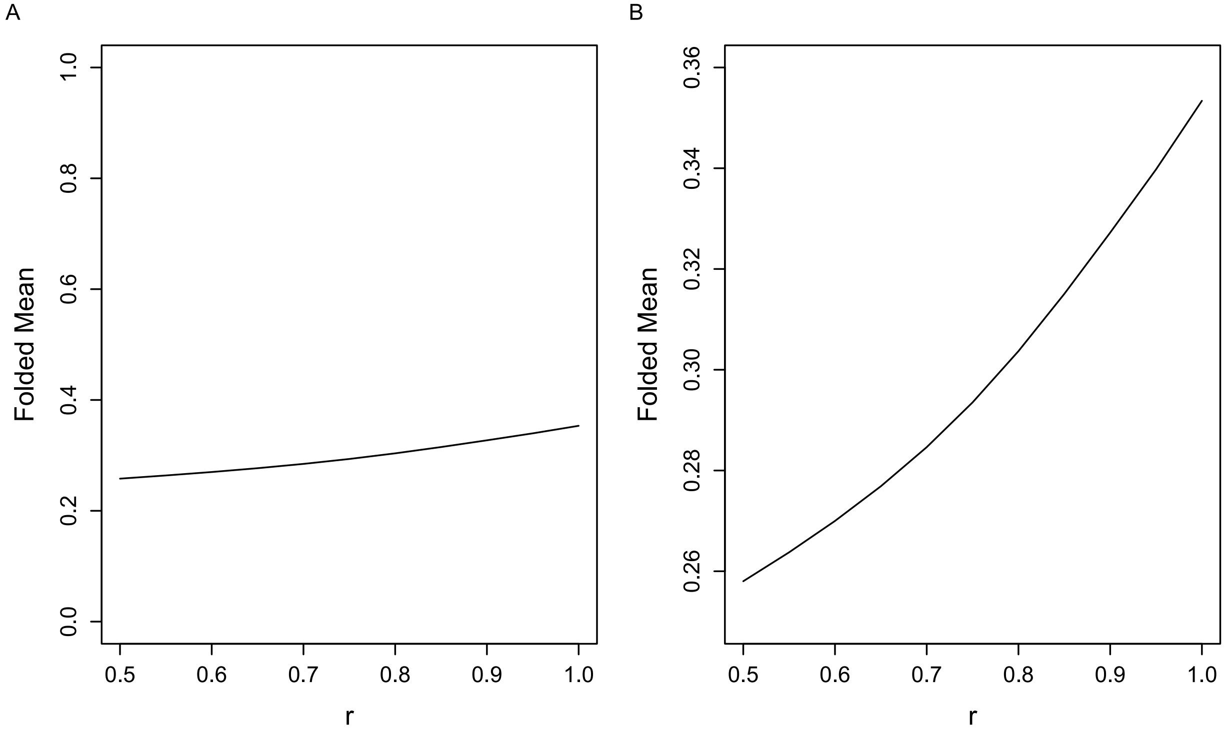

Figure S2. Change in the estimated folded mean for a range of values for *r*. A and B differ only in the scale of the y axes.