Mate choice copying in non-human animals: an update of two meta-analyses

Eduardo S. A. Santos [*1], Aleksandra Milenović [#1], Aneta Arct [#1], Anna Lenz [#1], Ayumi Mizuno [#1], Cassidy Schneider [#1], Christine Sosiak [#1], Erick Lundgren [#1], Hao Qin [#1], Iwo Gross [#1], Jimuel Celeste [#1], Kyle Morrison [#1], Mahi Zakir [#1], Marija Purgar [#1], Santiago Ortega [#1], Sergio Hernandez [#1], Shinichi Nakagawa [‡1], and Malgorzata Lagisz [‡1]

[1]Collaboration for Open Science and Synthesis in Ecology and Evolution (COSSEE), Department of Biological Sciences, Faculty of Science, University of Alberta, Edmonton, Canada

[*]Corresponding author: esantos2@ualberta.ca
[#]Listed alphabetically by first name.
[‡]Co-senior author

Abstract

Mate choice copying, the use of social information from the mating decisions of others, is a taxonomically widespread form of social learning with the potential to shape sexual selection. Two meta-analyses concluded that social information moderately and positively biases mate choice, but both rest on literature searches now at least seven years old. Using a multilingual, grey-literature-inclusive search across seven languages, we added 69 effect sizes from 29 studies to the original datasets and harmonized the two source metrics onto common scales, producing 227 Hedges' *g* effect sizes from 87 studies and 172 log odds ratios from 69 studies across 33 species. Re-analysed with phylogenetically informed multilevel models, the overall effect remained positive and significant but was appreciably smaller than before (Hedges' g = 0.45, 95% CI: 0.25 to 0.66, *vs.* 0.58; odds ratio = 1.81, 95% CI: 1.34 to 2.45, *vs.* 2.71). Heterogeneity was high and located within studies and among species, not across the phylogeny. We detected strong small-study effects on both scales, and correcting for them drew the adjusted mean towards zero. Overall, in some species, individuals do copy the mate choices of conspecifics, but the effect is smaller, more variable, and more sensitive to publication bias than previously estimated.

Keywords: meta-analysis, mate choice copying, nonindependent mate choice, social learning, sexual selection, social information

Summary statement: Updating two meta-analyses with seven more years of evidence confirms that animals copy each other's mate choices, but the average effect is smaller and less certain than previously reported.

1 Introduction

The process of mate choice allows individuals to evaluate and eventually mate with partners who provide benefits that translate into greater fitness [1]. From a social perspective, mate choice events can incidentally convey information that is generally useful about the quality of a mate [2]. When individuals observe mate choice events, acquire social information, and use it to inform their own mating decisions, we call this process mate choice copying [3]. As a form of social learning, mate choice copying is theorized to evolve when the costs of independently sampling mates are high, information is limited, or the quality of a potential mate is difficult to assess [2, 3]. By using social information from conspecifics, individuals can reduce decision times and mitigate the energetic costs or predation risks of assessing competing choices [4].

Mate choice copying is taxonomically widespread and has been experimentally documented in a broad sample of non-human animals, including mammals, birds, fishes, and arthropods [2, 5, 6]. Understanding the mechanisms and occurrence of mate choice copying is important because socially mediated mating decisions can influence the strength and dynamics of sexual selection within a population [7]. Furthermore, when observers form a generalized preference for specific phenotypes rather than isolated individuals, this behaviour can facilitate widespread cultural transmission of mating preferences and ultimately modify the course of evolutionary change [8, 9]. Two independent meta-analyses, by Jones and DuVal [5] and Davies et al. [6], have examined the magnitude, moderators, and evolutionary implications of mate-choice copying using partly different scopes and analytical approaches.

In general, these two meta-analyses [5, 6] found that the presence of social information was moderately and positively associated with an observer's likelihood of choosing that potential mate, compared with no social information. Jones and DuVal [5] found that copying effects were stronger in free-living populations, among sexually inexperienced females, and when social information countered an observer's innate preference for an attractive mate. Expanding upon this, Davies et al. [6] explicitly included tests of male mate choice copying and found little evidence of difference in the strength of this behaviour between the sexes. However, they did detect variations based on the focal taxonomic group and the type of experimental design employed.

Updating the evidence base through replication and quantitative research synthesis is funda-

mental to cumulative science [10]. In ecology and evolution, syntheses and reanalyses can revise conclusions about established effects and reduce their estimated magnitude [10–12]. An updated meta-analysis challenged a canonical example of status signalling [11], while corrections for selective reporting, including publication bias, substantially reduced meta-analytic effect-size estimates [12]. We therefore expected an expanded and bias-corrected synthesis to recover a positive but smaller overall effect of mate-choice copying than the original meta-analyses [5, 6].

In this meta-analysis, we used the datasets from Jones and DuVal [5] and Davies et al. [6] as our starting point. We updated the literature search, added newly identified studies, harmonized the effect-size data, and reanalyzed the combined datasets using phylogenetically informed multilevel models. We also tested for small-study effects and time-lag patterns, applied corrections for publication bias, and conducted sensitivity analyses. Together, these steps allowed us to assess the evidence more rigorously while accounting for potential biases. Our specific research questions were: (Q1) Do the main findings of the two previous meta-analyses remain supported? and (Q2) How much do the magnitudes of the estimated effect sizes change when we add new evidence and update analytical methods?

2 Materials and Methods

We preregistered our methods and planned analyses before data extraction and analysis in Open Science Framework (https://osf.io/vfahk/overview; [13]). Throughout our review and analysis process to guide the reporting, we followed the PRISMA-EcoEvo (Preferred Reporting Items for Systematic reviews and Meta-analyses in Ecology and Evolutionary biology; [14]) guidelines (Supplementary Table S1). We report author contributions to each methodological step using MeRIT (Method Reporting with Initials for Transparency; [15]), giving author initials in-line throughout the Methods. Where initials would be ambiguous, we disambiguate by surname: AMizuno (Ayumi Mizuno) and AMilenovic (Aleksandra Milenovic); CSchneider (Cassidy Schneider) and CSosiak (Christine Sosiak).

2.1 Additions and deviations

We made a few changes to elements of our preregistered protocol. We only implemented the search strategy described under question 2 in the preregistered protocol, because the more

comprehensive, multilingual, grey-literature-inclusive strategy fully encompassed the narrower search, making the additional strategy redundant. We included the same studies as in the earlier meta-analyses (these were automatically retained in our dataset), plus new studies from all periods retrieved in our searches, as retention of the original studies is a key element of an update rather than entirely new synthesis, and allowed for direct reconciliation of old and new estimates on a common scale. We did not extract *de novo* effect sizes from original meta-analyses, but instead cross-checked effect sizes present in both original meta-analyses; cross-checking (rather than re-extracting) lets us flag discrepancies between the two source datasets while avoiding the risk of introducing new extraction error.

2.2 Eligibility criteria and search protocols

We used the PECOS (Population, Exposure, Comparator, Outcomes, and Study design) framework [16, 17] to define a single set of eligibility criteria, applied to both the literature search and the two screening stages, following our pre-registered protocol [13] and the criteria of Jones and DuVal [5] and Davies et al. [6]. **Population**: Non-human, sexually mature female or male individuals, from wild or captive populations of multicellular animals. Studies that experimentally manipulated sexual maturity or prior mating experience were eligible provided the focal individuals were mature at testing. **Exposure**: Tests of the audience effect during mate choice trials; that is, the observation of a social cue (*e.g.*, a female near or copulating with a male). The social cue could be the presence of a model or demonstrator individual of the same sex as the focal individual, a chemical, visual, or auditory cue, or an observed copulation. **Comparator**: A group (sample) of females or males making a choice of mate without social cues or with neutral/conflicting cues. For instance, a no-information control: individuals tested under identical conditions but with no model present, or the model hidden behind an opaque partition—so any preference reflects chance/baseline rather than social information. Pre-demonstration preference comparators, and comparisons with an explicit chance expectation were also considered. **Outcome**: The mating preference of the focal/observer individual. Studies had to report, for each focal individual, either a discrete mate choice or a continuous measure of preference. Examples include preference-zone time, defined as the number of seconds the observer spent within a specific distance of a target individual after the demonstration; a social learning index, calculated to quantify the bias toward the phenotype chosen by the observer; and copulation choice, defined as the phenotype the focal/observer individual chose to mate with when physical

access was allowed. **Study design**: Empirical experimental studies that expose focal/observer individuals to opposite-sex mates (with and without an audience effect), in which one focal or observer individual chooses between two individuals of the opposite sex, one of which—the target—was associated with stimuli from a model or demonstrator.

Additional restrictions. Beyond the PECOS criteria, studies were eligible only if (i) they were reported in English, Japanese, Polish, Portuguese, Russian, Simplified or Traditional Chinese, or Spanish; and (ii) they reported sufficient statistical information to calculate an effect size—group means with standard deviations or standard errors and sample sizes, raw choice counts, or a convertible test statistic. Effects reported only as F-statistics, χ^2 values, or generalised linear mixed-model outputs were not convertible and were excluded (Supplementary Table S3).

We conducted a literature search across multiple databases and platforms, including Scopus, ISI Web of Science, Google Scholar (for non-English studies), Scientific Electronic Library Online (SciELO), and Bielefeld Academic Search Engine (for theses and dissertations; *i.e.*, grey literature). We designed the search strings (see Supplementary Information) based on a pilot search-string strategy to assess sensitivity, feasibility, and transparency using the benchmarking/relative recall approach [18]. We did not restrict the database searches by publication year, language, subject area, or document type. The strings were translated for searches in non-English languages, and search results were assessed by reviewers with expertise in the respective languages: AMizuno for Japanese, ML for Polish and Russian, ESAS for Portuguese, HQ for Simplified/Traditional Chinese, and SO for Spanish. We did not limit the number of records, which were sorted by relevance, in the Google Scholar searches in each language. We conducted Google Scholar searches using the software `Publish or Perish, v8.19` [19].

2.3 Study selection and screening

After the literature search, we deduplicated records and proceeded to our two-stage screening (with participation by all authors), following the procedures described in our preregistered protocol. In cases of disagreement between the reviewers, discrepancies were discussed and resolved to reach a consensus. The screening process and results are shown in the PRISMA flowchart (Figure 1; built with the `PRISMA2020 Shiny App` [20]).

2.4 Data collection

We extracted three types of information from each study. First, we collected citation information, such as title, author name, and publication year. Second, we collected the details of each experiment within studies: the species studied, type of experimental design (before-and-after or no pre-test), type of copying mechanism (generalized: trials presented observers with different target individuals that maintained a certain trait; or individual: trials presented observers with the same target individuals), type of social stimulus (mating, visual association, acoustic, chemical), type of choice outcome (mating, association time), sex of observer individual (female or male), if individual is virgin (yes or no; if not specified then it was coded as NA), and demonstrator individual age (same-age or different age from observer). Across original studies, observers' responses were quantified using a variety of measures, including association time with the target stimuli, the number of responses, and the number of mating bouts. Henceforth, we refer to these measures and responses as 'mate choice.'

Third, we obtained data to calculate effect sizes (*e.g.*, means, standard deviations, or standard errors, and sample sizes for the control and treatment groups; or the number of individuals that chose each target stimulus) from figures, tables, or text. Values reported only in figures were extracted with the assistance of a generative artificial intelligence tool and validated against an established digitisation package (see Use of artificial intelligence tools). If information was ambiguous or missing, the study was excluded from our dataset (see Supplementary Table S3). Data were primarily extracted by ESAS, with 10% of records checked for accuracy by ML.

2.5 Effect size calculation

Following the original meta-analyses of Jones and DuVal [5] and Davies et al. [6], we calculated effect size estimates as Hedges' g and Odds Ratio, so that we could investigate the effect of additional effect sizes on either one of the original meta-analyses. When studies reported means, standard deviations, and sample sizes for the treatment and control groups, we used standard formulas to estimate the standardized mean difference as Hedges' g (a bias-corrected version of Cohen's d). For studies that reported their results as count data, *i.e.*, the number of observer individuals who preferred each stimulus type (control and treatment), we estimated the Odds Ratio using the standard formula. All formulae and their sampling variances follow the standard meta-analytic treatments of Borenstein et al. [21]. Given that we are updating two

meta-analyses that used distinct effect sizes to report their results, to be able to add (combine) our effect sizes to the original datasets, it was necessary to convert the effect sizes from our dataset between the two metrics (Equation 16–Equation 19).

2.5.1 Hedges' g from group means and standard deviations

When a study reported the mean, standard deviation (*SD*) and sample size of a treatment and a control group, we computed the standardised mean difference d using the pooled standard deviation S_p (Equation 1) and applied the small-sample bias correction J (Equation 3) to obtain Hedges' g (Equation 4) and its sampling variance (Equation 5) [21]:

$$S_p = \sqrt{\frac{(n_1 - 1)\,SD_1^2 + (n_2 - 1)\,SD_2^2}{n_1 + n_2 - 2}}, \tag{1}$$

$$d = \frac{\bar{X}_{\text{treat}} - \bar{X}_{\text{ctrl}}}{S_p}, \tag{2}$$

$$J = 1 - \frac{3}{4(n_1 + n_2 - 2) - 1}, \tag{3}$$

$$g = J\,d, \tag{4}$$

$$\text{Var}(g) = \frac{n_1 + n_2}{n_1\,n_2} + \frac{g^2}{2(n_1 + n_2 - 2)}, \tag{5}$$

where n_1 and n_2 are the treatment and control sample sizes. Where only a standard error (*SE*) was reported, the standard deviation was recovered as $SD = SE\sqrt{n}$.

2.5.2 Hedges' g from test statistics

For studies reporting a t-statistic rather than group descriptives, we converted t to d following Borenstein et al. [21], adjusting the degrees of freedom for the experimental design. For a one-sample or paired design ($df = n - 1$), d was obtained from Equation 6 and the variance from Equation 9:

$$d = \frac{t}{\sqrt{n}}, \tag{6}$$

$$J = 1 - \frac{3}{4(n-1)-1}, \tag{7}$$

$$g = J\,d, \tag{8}$$

$$\mathrm{Var}(g) = \frac{1}{n} + \frac{g^2}{2(n-1)}. \tag{9}$$

For an independent two-sample design ($df = n_1+n_2-2$), d was obtained from Equation 10:

$$d = t\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \tag{10}$$

$$J = 1 - \frac{3}{4(n_1+n_2-2)-1}, \tag{11}$$

$$g = J\,d, \tag{12}$$

$$\mathrm{Var}(g) = \frac{n_1+n_2}{n_1\,n_2} + \frac{g^2}{2(n_1+n_2-2)}. \tag{13}$$

When individual group sizes were not reported for an independent-samples design, we assumed an equal split, $n_1 = n_2 = N/2$. Effects reported only as F-statistics, χ^2 values or generalised linear mixed-model outputs were not directly convertible and were excluded from this step.

2.5.3 Odds ratio from a 2×2 table

For studies reporting binary choice frequencies, we summarised the outcome as an odds ratio (*OR*, which was log-transformed for model fitting; Equation 14), with the variance of the log odds ratio obtained by the delta-method approximation (Equation 15) [21]:

$$OR = \frac{A\,D}{B\,C}, \tag{14}$$

$$\mathrm{Var}(\ln OR) = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}, \tag{15}$$

where A and B are the number of treatment-group individuals choosing the option associated with the social information and the alternative, respectively, and C and D the corresponding counts for the control group. When no separate control group existed, we used the null expectation that individuals chose each option with equal probability, setting $C = D = n/2$, as the two original meta-analyses did. This applied to 22 of the 38 binary effect sizes in our new extraction (11 of 17 studies). Where n was odd, the resulting half-integer cell counts were retained rather than rounded, because $\ln(OR)$ and its delta-method variance are defined for non-integer counts; this affected nine effect sizes. No 2×2 table contained a zero cell, so no continuity correction was applied.

2.5.4 Conversion between effect-size metrics

To place all studies on a common scale, effects available in only one metric were converted to the other using the logistic-to-normal approximation [21, 22]. A log odds ratio was converted to a standardised mean difference (Equation 16) and then bias-corrected to g (Equation 18):

$$d = \ln(OR) \times \frac{\sqrt{3}}{\pi}, \tag{16}$$

$$\mathrm{Var}(d) = \mathrm{Var}(\ln OR) \times \frac{3}{\pi^2}, \tag{17}$$

$$g = J\,d, \qquad \mathrm{Var}(g) \approx J^2\,\mathrm{Var}(d), \tag{18}$$

with J evaluated as above ($J \approx 1$ where n was unknown). Conversely, Hedges' g was converted to a log odds ratio using Equation 19:

$$\ln(OR) = g \times \frac{\pi}{\sqrt{3}}, \tag{19}$$

$$\mathrm{Var}(\ln OR) = \mathrm{Var}(g) \times \frac{\pi^2}{3}. \tag{20}$$

These conversions are approximate and assume that the underlying continuous trait follows a logistic distribution [22]; converted values were flagged in the data set to distinguish them from directly calculated effects in sensitivity analyses.

2.6 Phylogeny

Our combined sample (original + new datasets) comprised 33 species spanning three broad taxonomic groups: arthropods (*e.g., Drosophila melanogaster, Enchenopa binotata, Schizocosa ocreata, Uca mjoebergi*), fish (*e.g., Poecilia reticulata, Gasterosteus aculeatus, Oryzias latipes, Syngnathus typhle, Porichthys notatus*), and other vertebrates including birds (*e.g., Coturnix japonica, Taeniopygia guttata, Molothrus ater, Ficedula hypoleuca*) and mammals (*e.g., Dama dama, Mus musculus, Rattus norvegicus*). Several families were represented by multiple species (*e.g.*, Poeciliidae: *P. reticulata, P. latipinna, P. mexicana, P. formosa, Limia nigrofasciata, L. perugiae, Gambusia holbrooki, Brachyrhaphis rhabdophora*; Drosophilidae: *D. melanogaster, D. subobscura, D. simulans, D. serrata; Percidae: Etheostoma flabellare, E. zonale, E. olmstedi*). We therefore controlled for the non-independence arising from shared evolutionary history by incorporating a phylogeny as a correlation structure in our models [23].

Because no single published phylogeny contained all of our species, we assembled a synthetic tree from the Open Tree of Life [24] using the R package `rotl v3.0` [25], accessed through `prepR4pcm` [26] and following the workflow of Cinar et al. [27]. We retrieved the induced subtree for our species, resolved polytomies at random into a strictly bifurcating topology (`ape::multi2di`), and, as reliable branch lengths could not be obtained, computed branch lengths using Grafen's [28] method as implemented in the R package `ape v5` [29]. Species names were first reconciled to full binomials and against the Open Tree taxonomy: subspecies were collapsed to the species level (*e.g., Taeniopygia guttata castanotis* became *Taeniopygia guttata*), species recorded only by epithet were resolved to full binomials from their family, and accepted-name synonyms were relabelled to match our dataset (*e.g.*, the fiddler crab returned as *Austruca mjoebergi* was relabelled *Uca mjoebergi*). After this reconciliation, every species in both datasets was matched to a tip in the Open Tree synthetic tree, so no species required grafting next to a congener.

From the resulting ultrametric tree, we derived a phylogenetic correlation matrix, which we included as a random effect in our multilevel meta-analytic models (see below). We built two trees following identical procedures: one restricted to the 15 species of our own extraction (69 effect sizes) and one for the full combined dataset of 33 species, pooled across our extraction, Davies et al. [6], and Jones and DuVal [5]. The final phylogenies are shown in Figure 2 and Supplementary Fig. S1, for the combined dataset.

2.7 Meta-analysis and meta-regressions

We used the `rma.mv` function from the package `metafor v.5.0-1` [30] in `R v.4.6.0` [31] for our analyses. We started by fitting multilevel mixed-effects meta-analytic models only to the dataset of effect sizes compiled for our study. These meta-analytic models explicitly incorporated random factors: Study ID, Species ID, and Observation ID, which were fitted using the above function [30]. The models also included a variance-covariance matrix to handle dependencies among multiple effect sizes within the same study (assuming $r = 0.5$ for nested measurements). To account for the non-independence arising from phylogenetic relatedness among species, we included a correlation matrix of phylogenetic relatedness among species as a random effect [32].

We used the same logic described above to fit separate meta-analytic models to the original odds ratio data of Jones and DuVal [5] and the Hedges' *d* data of Davies et al. [6]. We combined the Hedges' *d* data of Davies et al. [6] with our continuous data coded as Hedges' *g*. The last step in this chain was to fit separate meta-analytic models (for Hedges' *g* and for odds ratio) using the combined datasets, thus providing the full updated meta-analytic mean effect of mate choice copying experimental evidence.

We also quantified the total I^2 (a measure of heterogeneity not attributed to sampling error) [33] and how much heterogeneity each random factor explained (partial I^2), calculated by the `i2_ml` function from the package `orchaRd v.2.2.0` [34]. We created all result plots using the `orchard_plot` and `bubble_plot` functions from the package `orchaRd` [34].

2.7.1 Model specification and inference

All models were fitted by restricted maximum likelihood (REML), and inference on fixed effects used t-distributed test statistics with containment degrees of freedom (`test = "t"`, `dfs = "contain"` in `rma.mv`), which is preferable to z-based inference given the moderate number of clusters available. For every model we report the variance component of each random effect (σ^2) alongside total and partial I^2, and a 95% prediction interval for the overall mean. We report both because I^2 describes the *proportion* of observed variance not attributable to sampling error rather than its absolute magnitude, and is sensitive to the precision of the contributing studies [35, 36]; the prediction interval expresses the expected range of true effects in a new study on the original scale.

2.7.2 Meta-regression models

We fitted taxonomic group (arthropods, fish, and other vertebrates) as an intercept-coded categorical moderator on each combined dataset, following the heterogeneous-variance workflow of `orchaRd 2.0` [34]. Because residual variance differed markedly among taxonomic groups, we allowed the effect-size–level variance to differ by group, specified as a heteroscedastic compound-symmetric structure (`struct = "HCS"`) on `~ tax_group | identifierEffectSizeID` with the between-group correlation fixed at zero (`rho = 0`), so that one residual variance is estimated per group with no cross-group covariance. A single pooled between-study variance (`~ 1 | identifierStudyId`) and a homoscedastic species variance (`~ 1 | taxonomySpecies`) were retained. As in the intercept-only models, a phylogenetic correlation matrix was included as an additional random effect. As in the main analyses, the sampling variance–covariance matrix was constructed with `metafor::vcalc` assuming a within-study correlation of $\rho = 0.5$, and models were fitted by REML with `test = "t"` and `dfs = "contain"`. Both models converged at the first attempt using the `nlminb` optimiser without warnings. We report the omnibus F-test of the moderator together with group-specific marginalised means and pairwise contrasts obtained with `orchaRd::mod_results` (via `emmeans`), so that inference does not depend on an arbitrary reference level.

Four further moderators extracted during data collection—the copying mechanism (generalized *vs.* individual), the modality of the social cue (visual, chemical, or acoustic), the mating status of the observer, and the relative age of the demonstrator—were not part of our pre-registered analyses, and several of their levels are sparsely sampled. We therefore fitted them as single-moderator multilevel models on the new extraction (both scales) and, where the coding could be harmonized with an original dataset, on the combined data, and treat all four as exploratory throughout. The relative age of the demonstrator showed no variation in our extraction (all "same-age") and could not be analyzed.

2.8 Small-study effects, time-lag patterns, and sensitivity analyses

We assessed small-study and time-lag effects with three approaches, applied to both combined datasets (Hedges' *g* and log odds ratio). First, we visually assessed funnel-plot asymmetry—the expected footprint of selective reporting—by plotting the residuals of a meta-analytic model containing all our random factors against effect-size precision ($1/SE$; funnel plots in Supple-

mentary Information). Second, as a regression-based test we fitted a multilevel analogue of Egger's regression, using the standard error ($SE_i = \sqrt{v_i}$) as the precision moderator and, where its slope was significant, re-fitting with the sampling variance (v_i) as the moderator, whose intercept estimates the bias-adjusted mean [37–39]; because the sampling variance of standardised mean differences and log odds ratios depends partly on the estimate itself, we treat this bias-corrected mean as a conservative bound rather than a point estimate (Supplementary Methods). Third, we examined time-lag bias by fitting publication year as a single moderator, and then a multi-moderator model containing both publication year and the precision moderator, to test whether either pattern persisted after conditioning on the other.

We estimated the adjusted overall mean under two complementary correction frameworks: the bias-corrected intercept of a multilevel meta-regression on the sampling variance [37], and a bias-robust fixed-effect estimator with cluster-robust (CR2) variance estimation [12]. The two strike different balances between representing the multilevel dependence structure and guarding against selective reporting; we give their full specification and respective limitations in Supplementary Methods, and overlay both corrected means on the orchard plots (`orchaRd::pub_bias_plot()`).

We tested the robustness of the combined estimates in two further ways, each refitting the model with the same structure as the main analysis. First, leave-one-out and leave-one-species-out analyses removed all effect sizes from one study, or one species, in turn. Second, because the two original meta-analyses overlap in their primary studies but report them on different metrics, we cross-validated by comparing, for every shared study, the Hedges' *g* reported by Davies et al. [6] with the *g* implied by converting the odds ratio reported by Jones and DuVal [5] for the same experiment and outcome (Equation 18); this flagged seven studies whose values disagreed by more than 0.5 Hedges' *g* units. We then refitted each combined model under three versions of the data—the original data (*Main*), a conservative version discarding the seven flagged studies (*Leave-out*), and a best-estimate version with their outcomes re-extracted from the primary sources (*Corrected*). Full specifications of the leave-one-out procedure and the exact re-extraction accounting are given in Supplementary Methods.

All data processing, effect size calculations, model specifications, diagnostic checks, bias assessments, and full moderator outputs are archived in a fully reproducible online supplement (https://esantos2ua.github.io/mate_choice_meta/; code repository https://github

.com/esantos2ua/mate_choice_meta, archived at Zenodo, DOI: 10.5281/zenodo.21829677). The archive includes annotated R code, the complete `summary()` output and variance components (σ^2 for study, species, observation, and phylogeny) for every model reported here, and all figures. Every result and interpretation on which our conclusions rest is reported in the main text or in Supplementary Information; the online archive provides reproducibility, not additional inference.

2.9 Use of artificial intelligence tools

Generative artificial intelligence tools were used at three points in this work. In every case their output was checked by the authors, who take full responsibility for the content of this manuscript. First, numerical values reported only in figures were extracted using `Google Gemini 3.1 Pro` : a cropped figure and its caption were supplied to the model with a prompt requesting the point estimates and error-bar limits, returned as a markdown table. Because automated figure extraction can fail silently, ESAS checked every extracted value against the source figure, and independently re-extracted with `metaDigitise` [40] a validation subset of 20 records; the two methods agreed closely (mean signed *z*-score $= -0.10$, 95% CI: -0.27 to 0.06). Second, `Claude Sonnet 5` was used to draft the analytical code and to structure the reproducibility repository. ESAS reviewed and tested all generated code, and publicly archived the complete annotated source (see Data and resource availability), so that every result reported here can be traced to the code that produced it. Third, `Claude Opus 5` was used to review the written manuscript and suggest edits. All suggestions were evaluated by the authors, and ESAS and SN verified every final wording decision. No artificial intelligence tool was used to search the literature, to screen records for eligibility, or to judge study inclusion. No tool made a final analytical or interpretive decision: model specifications, the treatment of contentious effect sizes, and all conclusions drawn from the results were determined by the authors.

3 Results

3.1 Screening outcomes and dataset characteristics

We obtained a total of 69 new effect sizes from 29 studies for our meta-analysis update [41–69]. Thirteen of these 29 studies (45%) were published in or before 2019 and therefore fell within the search window of the original meta-analyses, yet had not been included in them. That our mul-

tilingual, grey-literature-inclusive strategy recovered this much pre-2019 evidence—including six theses (20.7% of the total)—suggests the original searches were substantially less sensitive than their date ranges imply. The screening process and reasons for exclusion at the full-text screening stage are summarised in the PRISMA flowchart (Figure 1), with additional details in the lists of included and excluded studies (available in Supplementary Table S2 and Supplementary Table S3). These studies reported mate choice copying trials for 15 species (Figure 2; see Supplementary Fig. S1 for the phylogeny combining the new and original meta-analyses data). *Drosophila* and *Poecilia* were the most-studied genera in our sample, appearing in 11 and 8 studies, respectively. These two genera accounted for 46 effect sizes (67% of the total).

3.2 What is the overall evidence for mate choice copying?

Jones and DuVal [5] reported a mean posterior Odds Ratio (OR) for mate choice copying of 2.71 (95% credible interval: 1.60 to 4.80) in their original study. Our mean overall OR using the new data is 1.46 (95% confidence interval; 95% CI: 0.97 to 2.20). Combining both datasets yielded an overall mean OR of 1.81 (95% CI: 1.34 to 2.45; Figure 3). All odds ratios are back-transformed from the log scale on which the models were fitted. Against a 50:50 chance expectation, a mean OR of 1.81 corresponds to an observer choosing the socially favoured male on 64% of occasions; the original estimate of 2.71 corresponds to 73%, and our new data alone (OR = 1.46) to 59%.

Davies et al. [6] reported a mean overall Hedges' *d* of 0.58 (95% CI: 0.34 to 0.83), conventionally described as a 'moderate' effect (*sensu* Cohen), although such benchmarks are arbitrary and poorly calibrated to behavioural and evolutionary data [70]. Our new data had a mean overall Hedges' *g* of 0.21 (95% CI: -0.02 to 0.43), a 'small' effect on the same scale. The combined mean overall Hedges' *g* is 0.45 (95% CI: 0.25 to 0.66; Figure 4).

Heterogeneity was substantial in all models. I^2—the proportion of observed variance not attributable to sampling error—exceeded 74%, and the total among-effect-size variance was τ^2 = 0.43 for Hedges' *g* and 0.58 for log odds ratio, giving 95% prediction intervals for a new study of -0.90 to 1.81 and -1.00 to 2.19, respectively (for details see Table 1). We report both because I^2 is a proportion, and so depends on the precision of the contributing studies, whereas τ^2 and the prediction interval express heterogeneity on the scale of the effect size itself. More specifically, study ID (representing the between-study effect) accounted for up to 10.5% across

models, whereas observation ID (representing the within-study effect) accounted for most of the heterogeneity on both scales (69.8% for Hedges' *g*, 45.4% for log odds ratio). Species ID (the non-phylogenetic species effect) contributed substantially on the log odds ratio scale (23.9%) but little on the Hedges' *g* scale (1.9%). Phylogenetic relatedness explained little heterogeneity across models ($I^2 \leq 7.1\%$).

3.3 Uni-moderator analyses of taxonomic group

We conducted uni-moderator analyses of Hedges' *g* and log-odds ratios to examine how mate choice copying varied among taxonomic groups (arthropods, fish, and other vertebrates). All estimated means were positive across the three groups in both models using the combined datasets (Figure 5). On the Hedges' *g* scale, the point estimates increased from arthropods (*g* = 0.25, 95% CI: 0.03 to 0.47), through fish (*g* = 0.50, 95% CI: 0.30 to 0.70), to the other vertebrates, *i.e.* birds and mammals (*g* = 0.70, 95% CI: 0.35 to 1.05). Total heterogeneity was high (94.3%) and the among-group test approached statistical significance ($F_{2,27}$ = 3.00, *p* = 0.066). On the log-odds-ratio scale the ranking differed, with fish highest (log OR = 0.82, 95% CI: 0.37 to 1.26), followed by the other vertebrates (log OR = 0.61, 95% CI: -0.04 to 1.25) and arthropods (log OR = 0.33, 95% CI: -0.13 to 0.80). Total heterogeneity was again high (89.5%), and differences among groups were not significant ($F_{2,25}$ = 1.17, *p* = 0.327). Residual variance was itself group-specific, and roughly four times larger in fish (τ^2 = 0.53 for Hedges' *g*, 1.04 for log odds ratios) than in arthropods (0.12 and 0.15). Because the two scales give different rankings and neither omnibus test is significant, we do not interpret the apparent gradient.

3.4 Exploratory uni-moderator analyses of biological moderators

No moderator explained a detectable share of the heterogeneity on either scale (all F-tests, $p > 0.2$; total I^2 remained above 80%). Two contrasts are nonetheless worth noting. First, the strength of copying did not significantly differ between the two copying mechanisms: generalized and individual copying were essentially identical in the combined Hedges' *g* data (0.43 *vs.* 0.46; $F_{1,225}$ = 0.05, p = 0.82), and in the new extraction the generalized estimate was, if anything, marginally larger (Hedges' *g* = 0.26 *vs.* 0.16; log OR = 0.48 *vs.* 0.30; both $p > 0.5$). Second, whereas Jones and DuVal [5] reported stronger copying among virgin observers, our new extraction showed the opposite ordering (Hedges' *g* = 0.09 for virgin *vs.* 0.36 for non-virgin observers; $F_{1,27}$ = 1.61, p = 0.22), and the significant difference disappeared when the new data

were pooled with the original binary data (OR = 1.75 *vs.* 1.91; p = 0.72; $F_{1,170}$ = 0.13, p = 0.72). Neither contrast was statistically distinct on either scale. Full diagnostics, model summaries, between-level contrasts, and orchard plots are provided in Supplementary Information.

3.5 Small-study effects, time-lag patterns, and sensitivity analyses

We found strong small-study effects on both the Hedges' *g* and odds ratio scales, and mixed evidence of a time-lag decline (present only for the binary outcomes, Figure 6 and Figure 7). Under the more conservative bias correction, the adjusted mean approached zero on both scales, whereas the bias-robust estimator retained a positive but reduced effect (Figure 3 and Figure 4). The pooled estimates were nonetheless robust to leave-one-out, leave-one-species-out, and to how the contentious effect sizes were handled, with the largest shift in any pooled mean being ≤ 0.07 units (full diagnostics and numeric results in Supplementary Information).

4 Discussion

By updating the datasets from the two original meta-analyses and reanalysing them within a unified, phylogenetically-informed multilevel framework—227 Hedges' *g* effect sizes from 87 studies and 172 log odds ratios from 69 studies, spanning 33 species—we can now answer the two questions we set out with. First, the qualitative conclusion of both foundational syntheses persists: across an expanded, multilingual, and grey-literature–inclusive search, observers experimentally exposed to social information were on average more likely to favour the associated mate, and the pooled effect remained positive and statistically significant on both the odds ratio and standardised mean difference scales. Second, the effect is smaller than previously implied—roughly half the original magnitude when the new evidence is considered on its own, with confidence intervals that span zero, and, under the more conservative bias correction, an adjusted mean close to zero. Mate choice copying therefore occurs across a broad range of taxa, but is more modest and more context-dependent than Jones & DuVal [5] and Davies *et al.* [6] implied—a pattern that is itself informative and that we explore below.

4.1 Mate choice copying as social learning, and its consequences for sexual selection

Theory frames mate choice copying as one expression of a general capacity to use social information, alongside socially guided foraging, habitat selection, and predator avoidance [4, 71, 72], with a central prediction that such information supplements, rather than overrides, an observer's own assessment of a prospective mate [2, 73]. A reliably positive yet modest average effect is what that prediction implies, and it reframes the biological question from whether animals copy to when they do so. Theory does not predict indiscriminate copying: social-learning strategies such as "copy when uncertain" and "copy when young or inexperienced" specify the narrow conditions under which relying on others outperforms personal information [72, 74]. Seen this way, the recent, well-powered studies that report weak, null, or context-restricted copying [42, 64, 66] are not evidence against the phenomenon but, possibly, evidence that, as theory predicts, it is expressed selectively—animals discount or reverse the behaviour when public information is unreliable, when demonstrators are sperm-limited, or when individual differences in personality intervene [75–77].

This conditional view also reframes the most prominent feature of our results—large and largely unexplained heterogeneity—as biological signal rather than statistical noise. Variation accumulated within studies and among species, rather than between laboratories or across the phylogeny, which is what we would expect if the expression of copying is tuned to the observer's experience, the reliability and consistency of the social information, and the local ecological context. Jones & DuVal [5] found exactly such conditionality, with stronger copying among inexperienced (virgin) females; our updated data, however, did not reproduce this virginity effect and, in the new studies, tended to reverse it. The candidate moderators we could examine—copying mechanism, cue modality, observer mating status, and demonstrator age—were explored only in *post-hoc*, non-pre-registered analyses and explained no detectable share of the heterogeneity, consistent with the relevant variation residing at finer or more interactive scales than our coding captured. That copying occurs even in animals lacking complex sociality, most notably *Drosophila*, further suggests that it taps domain-general learning mechanisms shared widely across animals rather than lineage-specific cognition [71, 78]—an interpretation consistent with the negligible phylogenetic signal we recovered.

Where that heterogeneity sits differs between the two scales. Species identity absorbed 23.9%

of it on the log odds ratio scale but only 1.9% on the Hedges' *g* scale, while phylogeny absorbed almost none on either (≤ 3.5%; Table 1). That difference reflects dataset composition rather than biology: species accounted for 29% of the heterogeneity in our new extraction and 45.6% in the binary data of Jones and DuVal [5], but for none at all in the continuous data of Davies et al. [6]. Two points follow. Where species variance is detectable it is not organised by relatedness, which points to mating system, ecology, and assay tradition rather than to lineage-specific cognition. But because *Drosophila* and *Poecilia* supply two-thirds of all effect sizes, "among-species" variance is partly variance among the few laboratories that study them; with the taxonomic gradient reversing between scales and neither omnibus test significant, we do not treat it as established.

Much of the remaining variance likely lies in design features our coding could not capture: before-and-after *versus* no-pretest designs, association time *versus* copulation, live *versus* video demonstrators, and—for 22 of the 38 binary effect sizes—whether a control group existed at all or a chance expectation stood in for one. These axes plausibly generate more variance than the biological moderators we tested, which would explain why none of the latter absorbed a detectable share.

Variation also enters through data extraction process itself. The two source syntheses report overlapping studies on different metrics, and for seven of them the values disagreed by more than 0.5 Hedges' *g* units (when converted to the same scale). Re-extraction from the primary sources resolved every one toward the odds ratios of Jones and DuVal: the continuous extraction had variously captured a different construct (association time, body size), doubled paired sample sizes, produced artifacts of degenerate 2×2 tables, and in one case reversed the sign, turning a reported absence of copying into apparent evidence for it. The binary extraction had its own faults, over-counting effect sizes in two studies. Our estimates are robust to how these values are handled (Supplementary Information), but the episode matters: much of what reads as biological disagreement between the two syntheses was extraction disagreement—an argument for cross-validating source data whenever meta-analyses are combined and/or updated.

Finally, one distinction among our exploratory moderators carries particular weight for the evolutionary consequences: whether observers copy in a *generalized* or an *individual* fashion. Only generalized copying—the acquisition of a preference for a male phenotype or trait rather than for one specific individual—can propagate through a population, bias the targets of sexual se-

lection, and seed the cultural transmission of mating preferences, although theory disagrees on how strongly copying ultimately accelerates or constrains sexual selection [9]; a preference attached to one individual, by contrast, ends with the demonstrated male. It is worth separating this evolutionary distinction from the experimental one it is inferred from. Our moderator codes the *design* of a trial, not the decision process of the observer: generalized designs present observers with different target individuals sharing a trait, whereas individual designs re-present the same targets. The two are not equally informative about the evolutionary question—an individual design cannot distinguish a preference generalised to the trait from one attached to the specific individual, or from other latent cues correlated with it—but neither is the more rigorous test, and individual designs are often the only feasible option for a given system. Our data are therefore consistent with, rather than a direct test of, the generalized form being robustly expressed: copying strength was statistically indistinguishable between the two designs and, if anything, marginally stronger under generalized designs in the new extraction. A design that manipulates the target trait explicitly and pre-registers the generalisation test would generate valuable information, and we regard this as a high priority for future work.

4.2 Limitations, knowledge gaps, and future opportunities

Several features of our approach constrain the strength of these inferences. Because the two source meta-analyses reported different metrics, placing all studies on a common scale required approximate conversions that assume an underlying logistic distribution: in our new extraction alone, 38 Hedges' g values were converted from odds ratios and 31 log odds ratios from Hedges' g. Twenty-two of 38 binary effect sizes had no separate control group and relied on an equal-allocation chance expectation, which likely understates their sampling variance. Much of our effect-size data was extracted from published figures; the agreement between our extraction routine and `metaDigitise` on a validation subset was close (mean signed z-score $= -0.10$), but figure extraction remains a source of error. Finally, several moderator levels were sparsely sampled and none of the biological moderator analyses was pre-registered, so those results are exploratory.

Three features of the evidence base temper these conclusions and set the agenda for future work. First, we found strong small-study effects on both scales; under the bias-corrected intercept [37] the adjusted mean collapsed towards zero, whereas the bias-robust estimator [12]

retained a positive but reduced signal. The honest interpretation is that the true effect very likely lies below the naïve pooled estimate, and that a much smaller effect cannot be excluded on present evidence. Second, for the binary outcomes the effect-size magnitude declined with publication year, echoing the familiar pattern in which early, large effects are gradually tempered by more rigorous and better-powered designs [79, 80]; several of the most recent studies in our sample report weak or null copying [42, 44, 64, 66]. Third, primary studies quantify copying in incompatible ways forcing approximate conversions and leaving a handful of studies whose values disagreed materially between datasets. Although our conclusions proved robust to how these contentious values were handled, the broader lesson is methodological: the field would benefit from registered reports, the routine publication of null and small-sample results, and minimum reporting standards—group means, standard deviations, sample sizes, and raw choice counts—so that future updates rest on less-censored, directly comparable data [6].

A further gap concerns taxonomic breadth. Although our combined sample spanned 33 species across arthropods, fish, and other vertebrates, the evidence remains concentrated in two model genera: *Drosophila* and *Poecilia* together contributed roughly two-thirds of all effect sizes. This is precisely the situation in which a sample concentrated in two genera is least informative, and in which broader taxonomic coverage would be most valuable. Reassuringly, the uni-moderator analyses returned positive means in all three taxonomic groups on both scales, with only weak support for among-group differences, so we treat the apparent gradient cautiously. Recalculating the pooled mean from those group means with the three groups weighted equally—an approximation to balanced taxonomic coverage—shifts it by less than 0.04 Hedges' g units and by less than 0.01 log odds ratio units, so the estimates are not an artefact of the sample's concentration in arthropods. Invertebrates, amphibians, reptiles, non-poeciliid fishes, and a broader range of birds and mammals remain effectively absent. Expanding the phylogenetic spread of the evidence would both test the generality of mate choice copying and sharpen phylogenetic comparative tests that, at present, have little signal to work with.

4.3 Conclusion

Mate choice copying is a quantifiable and taxonomically widespread behaviour: after incorporating seven additional years of evidence and analysing it with contemporary phylogenetic multilevel methods, we still recover a positive and significant overall effect on both the odds

ratio and standardised-mean-difference scales. At the same time, the updated picture is more tempered than the original meta-analyses suggested. The pooled effect is moderate at best, is surrounded by substantial heterogeneity that our models could not attribute to study, taxon, or phylogeny, and bears clear fingerprints of small-study and selective-reporting bias that, under the more conservative correction, draw the estimate close to zero. The qualitative claims of Jones & DuVal [5] and Davies *et al.* [6] hold, but their estimated magnitudes were probably optimistic. Realising the promise of mate choice copying as a driver of sexual selection and cultural evolution will require broader taxonomic sampling, mechanism-focused and pre-registered tests of the moderators that potentially influence its considerable heterogeneity, and a publication culture that faithfully reports null findings. Finally, we offer this update as a template for how meta-analyses in ecology and evolution can be revisited—combining new evidence with the original data, harmonising incompatible effect-size metrics, and stress-testing conclusions against publication bias and contentious extractions.

5 Figures

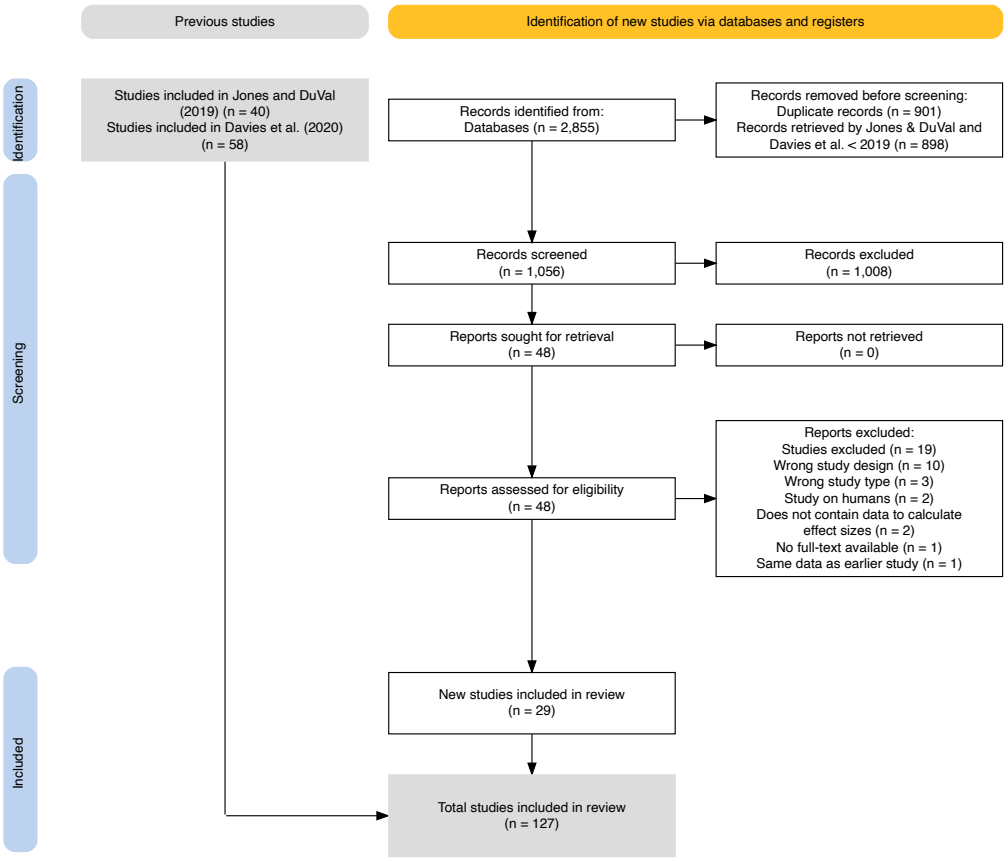

Figure 1: **PRISMA flowchart.** The diagram shows identification, screening, eligibility, and inclusion stages for the update of systematic reviews of the mate choice copying meta-analysis.

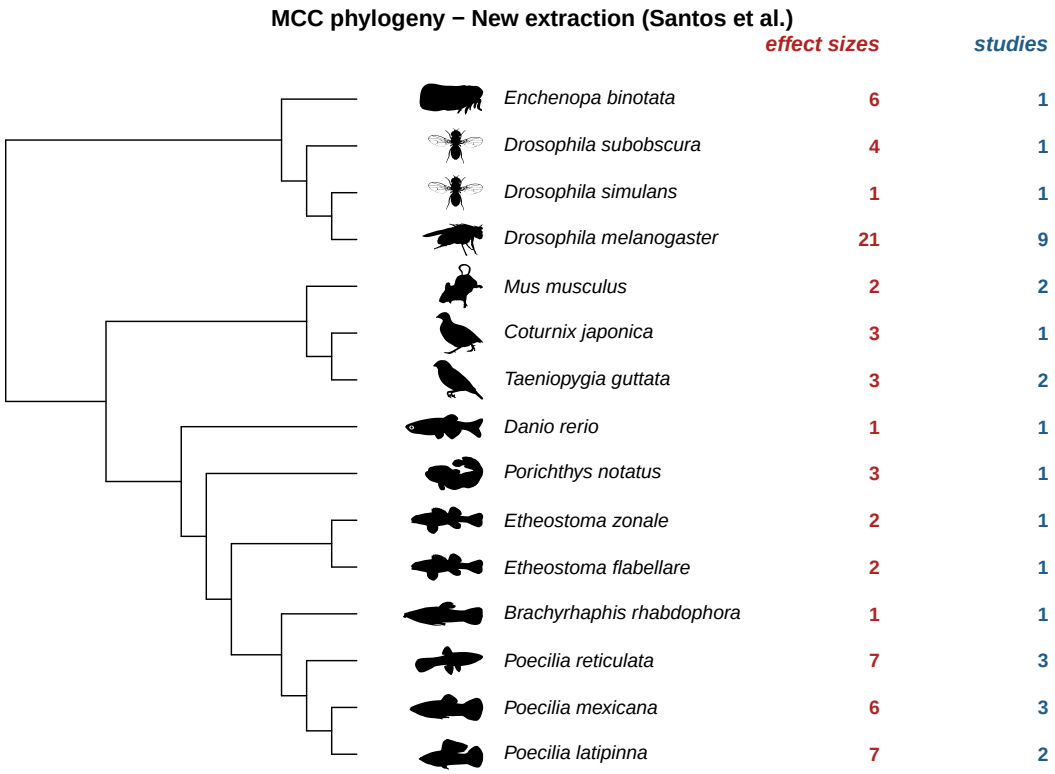

Figure 2: **Phylogenetic tree.** The phylogenetic tree of the species included in the meta-analysis, along with the number of effect sizes and the number of studies in which they are reported. Moran et al. [58] used two *Etheostoma* species, thus the number of studies column total adds to 30, when in fact the actual number of studies included is 29.

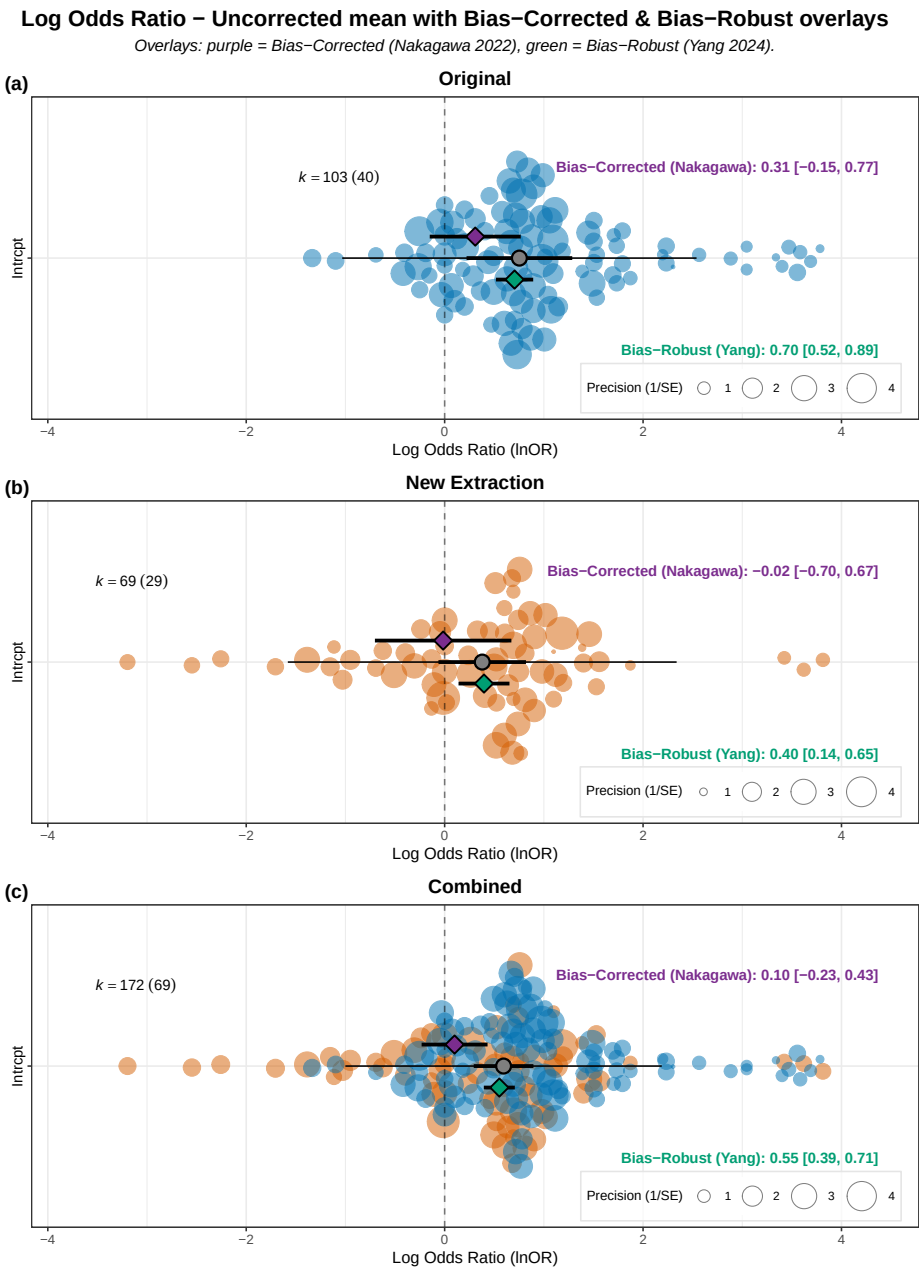

Figure 3: **Mean overall log odds ratio effect sizes.** All panels present three meta-analytic estimates: bias-corrected (Nakagawa; purple diamond), uncorrected (grey circle), and bias-robust (Yang; green diamond). Panels present the original (**a**), new (**b**, this study), and combined (**c**; original + new data) overall effect sizes. Thick horizontal lines represent 95% confidence intervals, and thin horizontal lines represent 95% prediction intervals. The points in the centre of each thick line indicate the average meta-analytic effect size. *k* is the number of effect sizes used to estimate the statistics, followed by the number of studies in the brackets. Each bubble is an individual effect size; bubble area is proportional to its precision (1/SE).

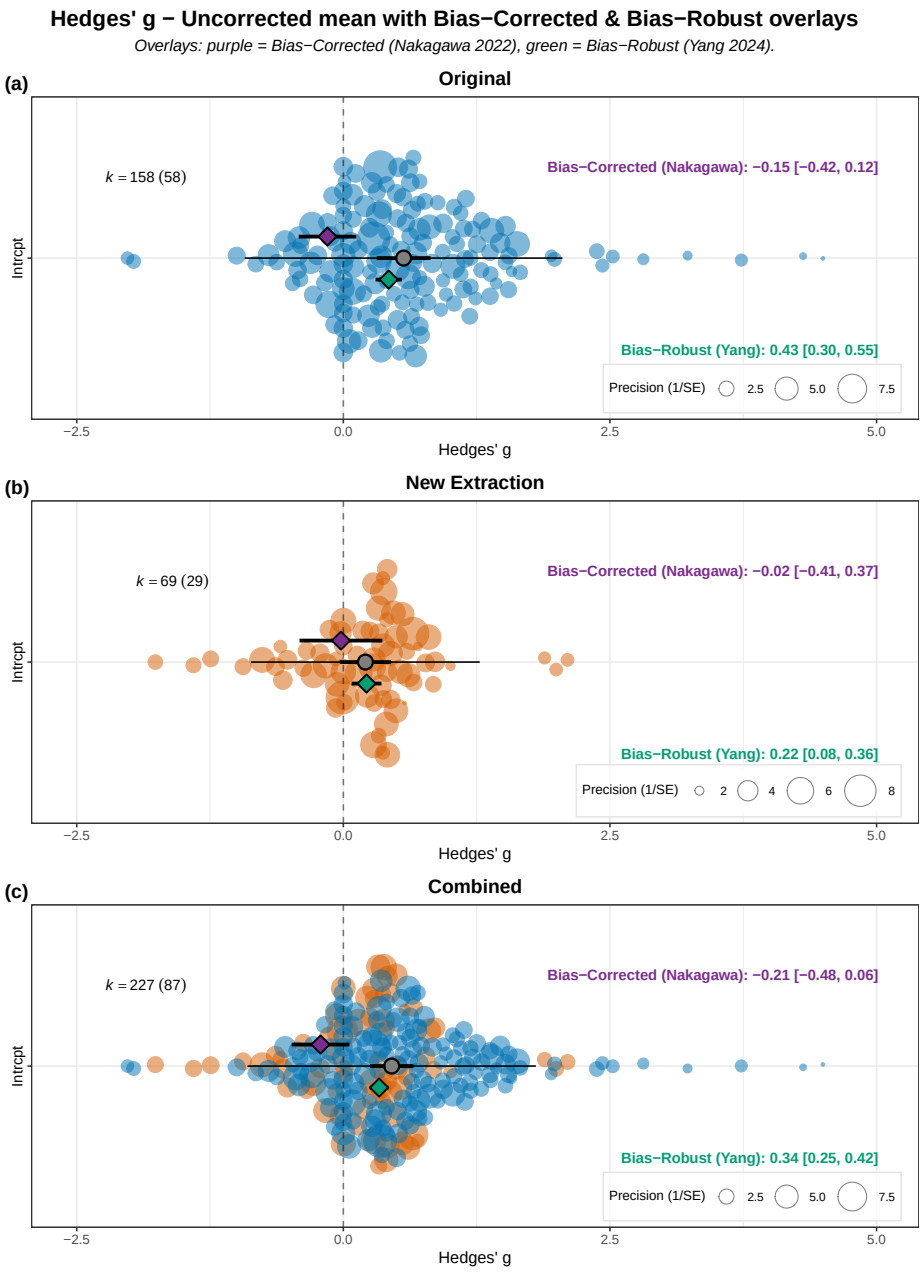

Figure 4: **Mean overall Hedges' effect sizes.** All panels present three meta-analytic estimates: bias-corrected (Nakagawa; purple diamond), uncorrected (grey circle), and bias-robust (Yang; green diamond). Panels present the original (**a**), new (**b**, this study), and combined (**c**; original + new data) overall effect sizes. Thick horizontal lines represent 95% confidence intervals, and thin horizontal lines represent 95% prediction intervals. The points in the centre of each thick line indicate the average meta-analytic effect size. *k* is the number of effect sizes used to estimate the statistics, followed by the number of studies in the brackets. Each bubble is an individual effect size; bubble area is proportional to its precision (1/SE).

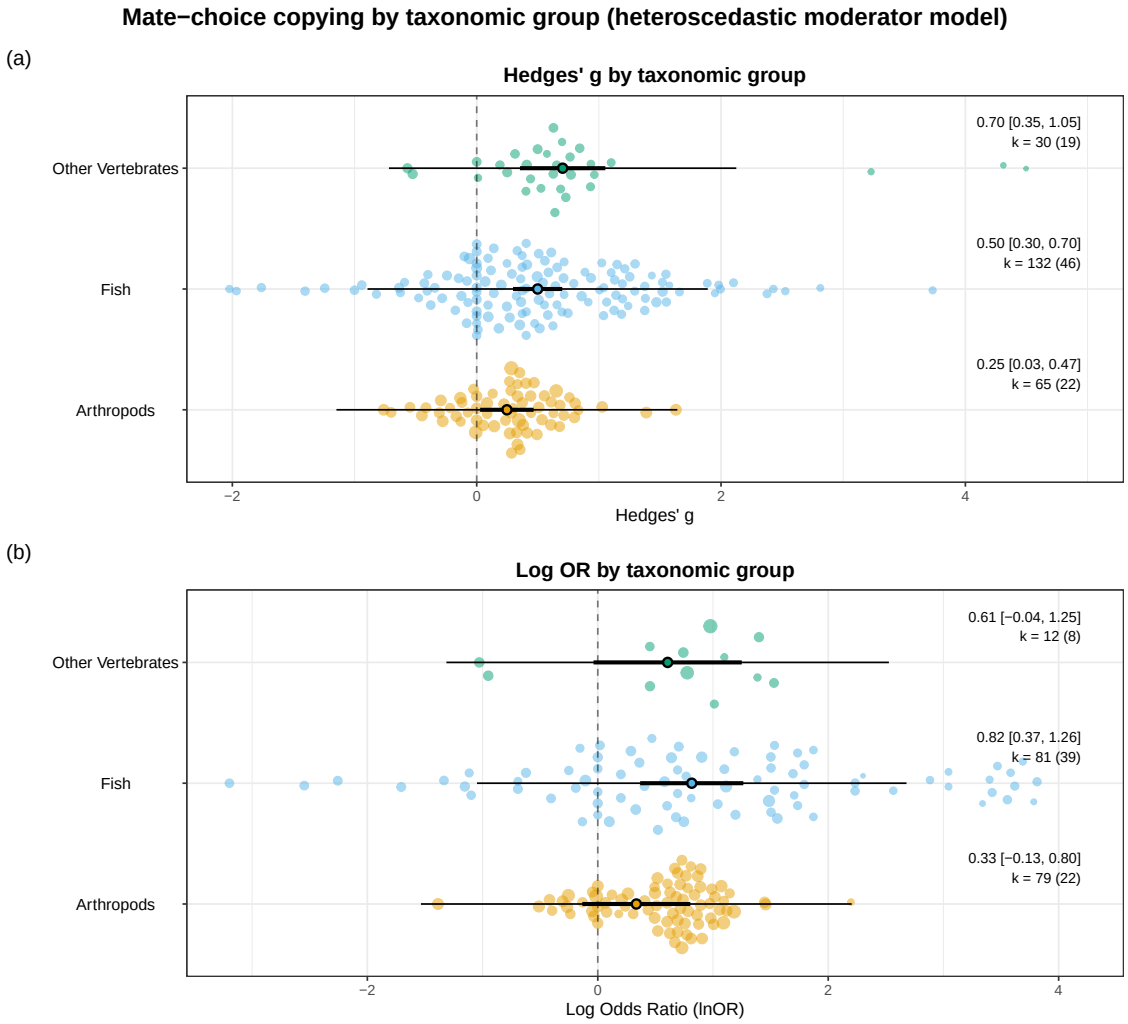

Figure 5: **Taxon moderator effect sizes.** (**a**) Hedges' *g*; and (**b**) log odds ratio. All panels present three meta-analytic estimates: other vertebrates (birds and mammals; green), fish (blue), and arthropods (yellow). Thick horizontal lines represent 95% confidence intervals, and thin horizontal lines represent 95% prediction intervals. The points in the centre of each thick line indicate the average meta-analytic effect size. *k* is the number of effect sizes used to estimate the statistics, followed by the number of studies in the brackets.

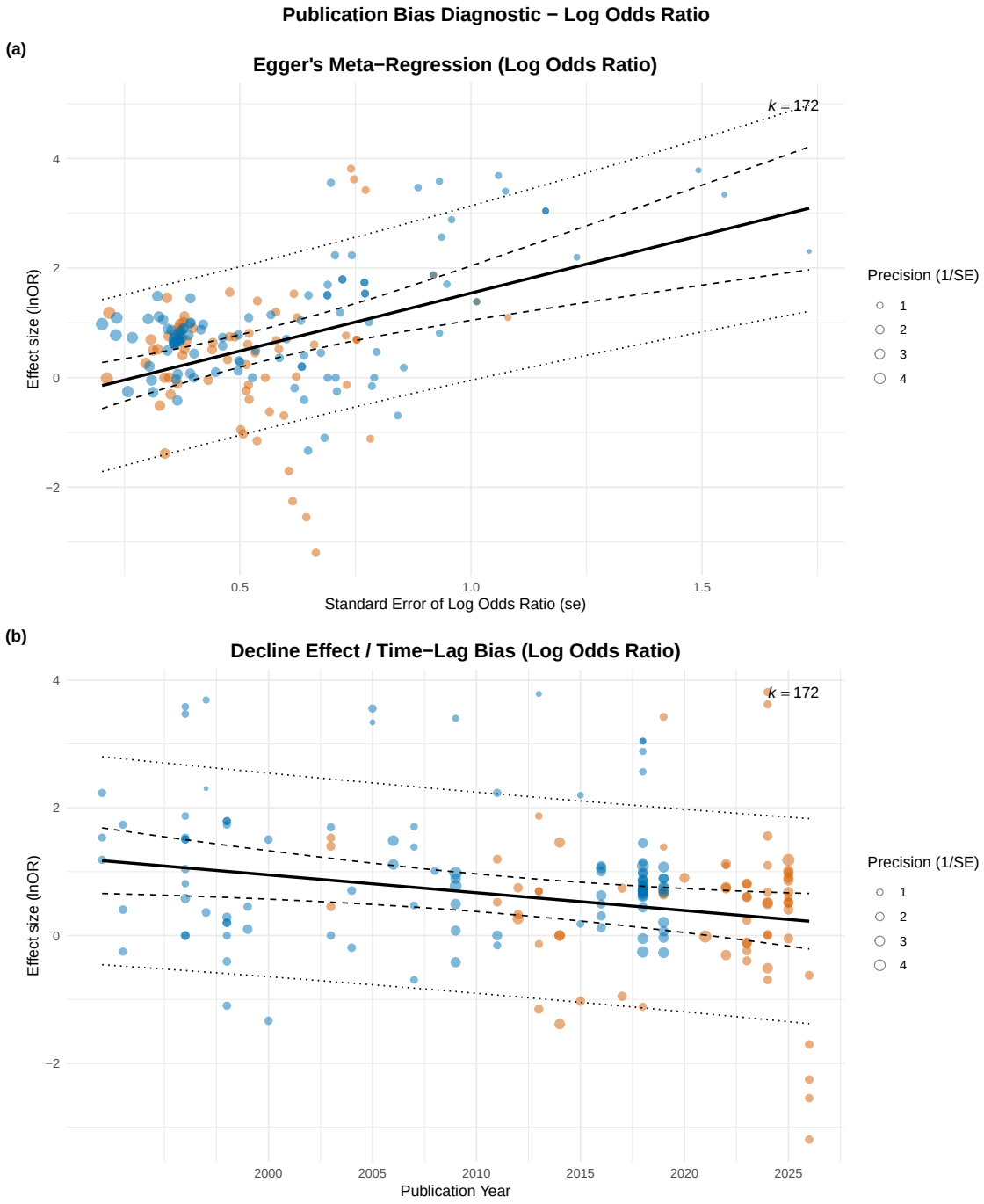

Figure 6: **Publication bias diagnostic - Odds ratio.** Associations between effect sizes and (**a**) the square root of the sampling variance and (**b**) publication year. Point size reflects precision; original data (blue) and new data (orange); solid lines show fitted meta-regressions, and dashed and dotted lines indicate 95% confidence and prediction intervals, respectively.

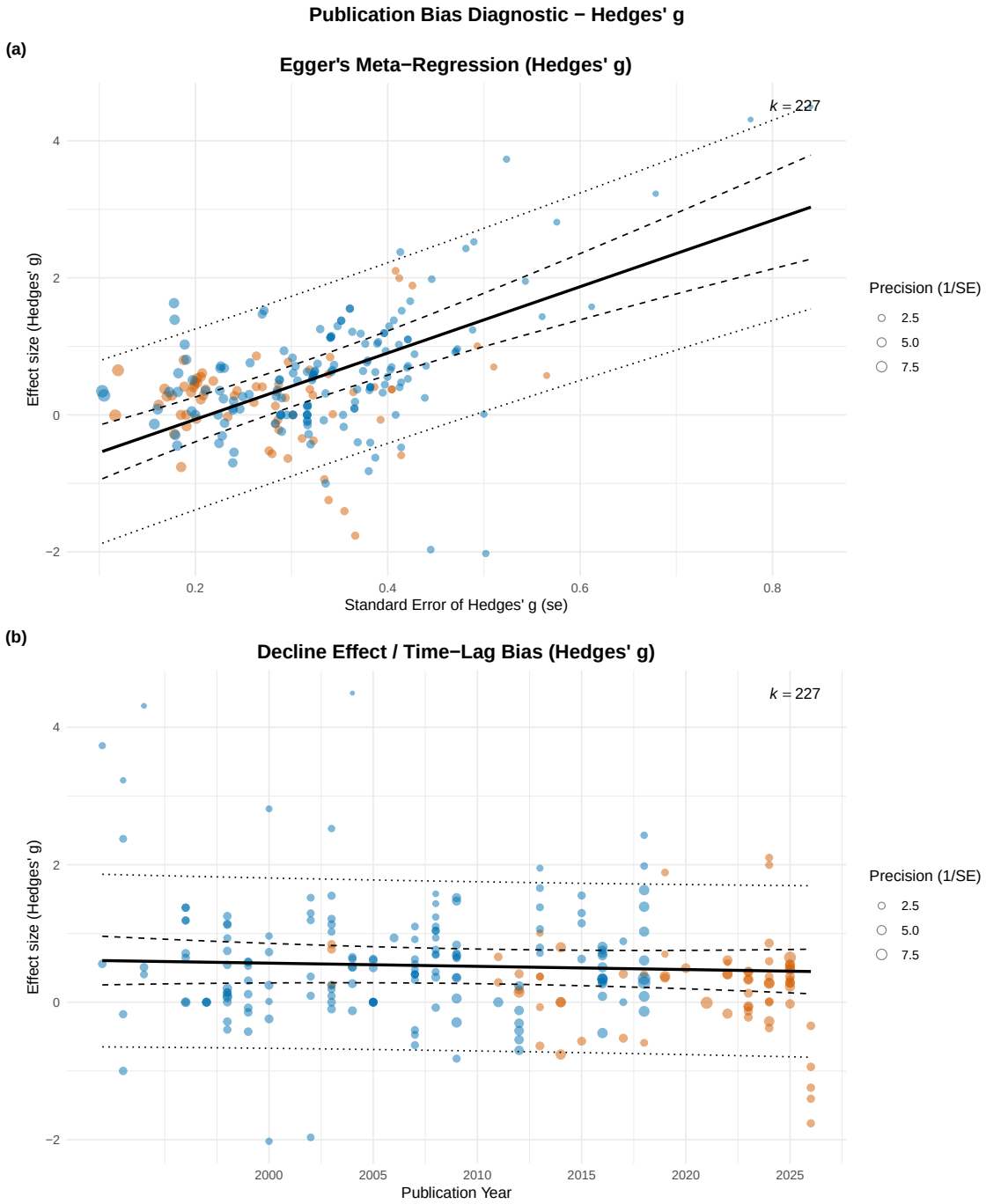

Figure 7: **Publication bias diagnostic - Hedges' g.** Associations between effect sizes and (**a**) the square root of the sampling variance and (**b**) publication year. Point size reflects precision; original data (blue) and new data (orange); solid lines show fitted meta-regressions, and dashed and dotted lines indicate 95% confidence and prediction intervals, respectively.

6 Tables

Table 1: Partial I^2 (%) for each random effect across the uncorrected multilevel models. Total is the overall proportion of variance due to heterogeneity (vs. sampling error); the remaining columns partition that heterogeneity among the study, residual effect-size, non-phylogenetic species, and phylogenetic components.

Analysis	Total	Study	Effect size (residual)	Species (non-phylo.)	Phylogeny
Hedges' g — New Only	80.8	0.0	51.3	29.4	0.0
Hedges' g — Original Only	85.8	4.1	77.1	0.0	4.6
Hedges' g — Combined	85.5	10.5	69.8	1.9	3.3
Log OR — New Only	80.8	0.0	51.3	29.5	0.0
Log OR — Original Only	75.9	0.0	23.2	45.6	7.1
Log OR — Combined	74.7	4.3	45.4	23.9	1.1

Funding

ESAS and SN are supported by the Canada Excellence Research Chairs (CERC) program (grant number CERC-2022-00074). ML is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant (RGPIN-2026-05996). CSosiak is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowship (599737-2025). AA was supported by the Polish National Agency for Academic Exchange (NAWA) under the Bekker NAWA programme (grant number BPN/BEK/2024/1/00075/U/00001).

Competing interest

No competing interests declared.

Acknowledgements

We are grateful to Davies et al. and to Jones and DuVal for making the data from their meta-analyses openly available, without which this update would not have been possible, and to the Open Tree of Life project and its contributors.

Data and resource availability

All data and code used in this study are available at https://github.com/esantos2ua/mate_choice_meta and were archived at Zenodo, DOI: 10.5281/zenodo.21829677.

Author Contributions

All authors contributed to screening records for eligibility (investigation) and to writing—review and editing. Additional contributions were as follows. **Eduardo S. A. Santos**: conceptualisation, data curation, formal analysis, methodology, project administration, software, visualisation, and writing—original draft. **Malgorzata Lagisz**: conceptualisation, methodology, and data validation. **Shinichi Nakagawa**: conceptualisation, funding acquisition, and methodology. Non-English records were screened by the authors with the relevant language expertise: AMizuno

(Japanese), ML (Polish and Russian), ESAS (Portuguese), HQ (Simplified and Traditional Chinese), and SO (Spanish).

References

1. Rosenthal, G. G. *Mate Choice: The Evolution of Sexual Decision Making from Microbes to Humans* 648 pp. (Princeton University Press, 2017).

2. Vakirtzis, A. Mate Choice Copying and Nonindependent Mate Choice: A Critical Review. *Annales Zoologici Fennici* **48,** 91–107. doi:10.5735/086.048.0202 (2011).

3. Wade, M. J. & Pruett-Jones, S. G. Female Copying Increases the Variance in Male Mating Success. *Proceedings of the National Academy of Sciences* **87,** 5749–5753. doi:10.1073/pnas.87.15.5749 (1990).

4. Valone, T. J. & Templeton, J. J. Public Information for the Assessment of Quality: A Widespread Social Phenomenon. *Philosophical Transactions of the Royal Society B: Biological Sciences* **357,** 1549–1557. doi:10.1098/rstb.2002.1064 (2002).

5. Jones, B. C. & DuVal, E. H. Mechanisms of Social Influence: A Meta-Analysis of the Effects of Social Information on Female Mate Choice Decisions. *Frontiers in Ecology and Evolution* **7,** 390. doi:10.3389/fevo.2019.00390 (2019).

6. Davies, A. D., Lewis, Z. & Dougherty, L. R. A Meta-Analysis of Factors Influencing the Strength of Mate-Choice Copying in Animals. *Behavioral Ecology* **31,** 1279–1290. doi:10.1093/beheco/araa064 (2020).

7. Verzijden, M. N. *et al.* The Impact of Learning on Sexual Selection and Speciation. *Trends in Ecology & Evolution* **27,** 511–519. doi:10.1016/j.tree.2012.05.007. PMID: 22705159 (2012).

8. Pruett-Jones, S. Independent versus Nonindependent Mate Choice: Do Females Copy Each Other? *American Naturalist* **140,** 1001–1009. doi:10.1086/285452 (1992).

9. Varela, S. A. M., Matos, M. & Schlupp, I. The Role of Mate-Choice Copying in Speciation and Hybridization. *Biological Reviews* **93,** 1304–1322. doi:10.1111/brv.12397 (2018).

10. Nakagawa, S. & Parker, T. H. Replicating Research in Ecology and Evolution: Feasibility, Incentives, and the Cost-Benefit Conundrum. *BMC Biology* **13,** 88. doi:10.1186/s12915-015-0196-3 (2015).

11. Sánchez-Tójar, A. *et al.* Meta-Analysis Challenges a Textbook Example of Status Signalling and Demonstrates Publication Bias. *eLife* **7** (eds Tautz, D. & Parker, T.) e37385. doi:10.7554/eLife.37385 (2018).

12. Yang, Y. *et al.* Robust Point and Variance Estimation for Meta-Analyses with Selective Reporting and Dependent Effect Sizes. *Methods in Ecology and Evolution* **15,** 1593–1610. doi:10.1111/2041-210x.14377 (2024).

13. Santos, E., Nakagawa, S. & Lagisz, M. Mate-Choice Copying in Non-Human Animals: Protocol for the Update of Two Meta-Analyses [Protocol]. *OSF*. doi:10.17605/osf.io/vfahk (2026).

14. O'Dea, R. E. *et al.* Preferred Reporting Items for Systematic Reviews and Meta-Analyses in Ecology and Evolutionary Biology: A PRISMA Extension. *Biological Reviews* **96,** 1695–1722. doi:10.1111/brv.12721 (2021).

15. Nakagawa, S. *et al.* Method Reporting with Initials for Transparency (MeRIT) Promotes More Granularity and Accountability for Author Contributions. *Nature Communications* **14,** 1788. doi:10.1038/s41467-023-37039-1 (2023).

16. Richardson, W. S., Wilson, M. C., Nishikawa, J. & Hayward, R. S. The Well-Built Clinical Question: A Key to Evidence-Based Decisions. *ACP Journal Club* **123,** A12. doi:10.7326/ACPJC-1995-123-3-A12 (1995).

17. Foo, Y. Z., O'Dea, R. E., Koricheva, J., Nakagawa, S. & Lagisz, M. 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. doi:10.1111/2041-210X.13654 (2021).

18. Lagisz, M., Yang, Y., Young, S. & Nakagawa, S. A Practical Guide to Evaluating Sensitivity of Literature Search Strings for Systematic Reviews Using Relative Recall. *Research Synthesis Methods* **16,** 1–14. doi:10.1017/rsm.2024.6 (2025).

19. Harzing, A. W. *Publish or Perish* https://harzing.com/resources/publish-or-perish.

20. Haddaway, N. R., Page, M. J., Pritchard, C. C. & McGuinness, L. A. PRISMA2020: An R Package and Shiny App for Producing PRISMA 2020-Compliant Flow Diagrams, with Interactivity for Optimised Digital Transparency and Open Synthesis. *Campbell Systematic Reviews* **18,** cl2.1230. doi:10.1002/cl2.1230 (2022).

21. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. *Introduction to Meta-Analysis* 351 pp. Google Books: JQg9jdrq26wC (John Wiley & Sons, 2011).

22. Chinn, S. A Simple Method for Converting an Odds Ratio to Effect Size for Use in Meta-Analysis. *Statistics in Medicine* **19,** 3127–3131. doi:10.1002/1097-0258(20001130)19:22<3127::aid-sim784>3.0.co;2-m. PMID: 11113947 (2000).

23. *Handbook of Meta-Analysis in Ecology and Evolution* (eds Koricheva, J., Gurevitch, J. & Mengersen, K.) doi:10.1515/9781400846184 (Princeton University Press, 2013).

24. Hinchliff, C. E. *et al.* Synthesis of Phylogeny and Taxonomy into a Comprehensive Tree of Life. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 12764–12769. doi:10.1073/pnas.1423041112. PMID: 26385966 (2015).

25. Michonneau, F., Brown, J. W. & Winter, D. J. Rotl: An R Package to Interact with the Open Tree of Life Data. *Methods in Ecology and Evolution* **7,** 1476–1481. doi:10.1111/2041-210X.12593 (2016).

26. Nakagawa, S. *et al.* prepR4pcm: An R Package for Preparing Data and Trees for Phylo-Genetic Comparative Methods. *EcoEvoRxiv*. doi:10.32942/X2468Z (2026).

27. Cinar, O., Nakagawa, S. & Viechtbauer, W. Phylogenetic Multilevel Meta-Analysis: A Simulation Study on the Importance of Modelling the Phylogeny. *Methods in Ecology and Evolution* **13,** 383–395. doi:10.1111/2041-210X.13760 (2022).

28. Grafen, A. The Phylogenetic Regression. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **326,** 119–157. doi:10.1098/rstb.1989.0106 (1989).

29. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and Evolution in R Language. *Bioinformatics* **20,** 289–290. doi:10.1093/bioinformatics/btg412. PMID: 14734327 (2004).

30. Viechtbauer, W. Conducting Meta-Analyses in R with the Metafor Package. *Journal of Statistical Software* **36,** 1–48. doi:10.18637/jss.v036.i03 (2010).

31. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2025).

32. Nakagawa, S. & Santos, E. S. A. Methodological Issues and Advances in Biological Meta-Analysis. *Evolutionary Ecology* **26,** 1253–1274. doi:10.1007/s10682-012-9555-5 (2012).

33. Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring Inconsistency in Meta-Analyses. *BMJ* **327,** 557–560. doi:10.1136/bmj.327.7414.557. PMID: 12958120 (2003).

34. Nakagawa, S. *et al.* orchaRd 2.0: An R Package for Visualising Meta-Analyses with Orchard Plots. *Methods in Ecology and Evolution* **14,** 2003–2010. doi:10.1111/2041-210x.14152 (2023).

35. Borenstein, M., Higgins, J. P. T., Hedges, L. V. & Rothstein, H. R. Basics of Meta-Analysis: I^2 Is Not an Absolute Measure of Heterogeneity. *Research Synthesis Methods* **8,** 5–18. doi:10.1002/jrsm.1230 (2017).

36. IntHout, J., Ioannidis, J. P. A., Rovers, M. M. & Goeman, J. J. Plea for Routinely Presenting Prediction Intervals in Meta-Analysis. *BMJ OPEN* **6,** e010247. doi:10.1136/bmjopen-2015-010247 (2016).

37. Nakagawa, S. *et al.* Methods for Testing Publication Bias in Ecological and Evolutionary Meta-Analyses. *Methods in Ecology and Evolution* **13,** 4–21. doi:10.1111/2041-210X.13724 (2022).

38. Stanley, T. D. & Doucouliagos, H. *Meta-Regression Analysis in Economics and Business* 176 pp. doi:10.4324/9780203111710 (Routledge, London, 2012).

39. Stanley, T. D. & Doucouliagos, H. Meta-Regression Approximations to Reduce Publication Selection Bias. *Research Synthesis Methods* **5,** 61–78. doi:10.1002/jrsm.1095 (2014).

40. Pick, J. L., Nakagawa, S. & Noble, D. W. A. Reproducible, Flexible and High-Throughput Data Extraction from Primary Literature: The metaDigitise r Package. *Methods in Ecology and Evolution* **10,** 426–431. doi:10.1111/2041-210X.13118 (2019).

41. Santiago Araújo, R., Nöbel, S., Antunes, D. F., Danchin, E. & Isabel, G. A Social Learning Primacy Trend in Mate-Copying: An Experiment in *Drosophila Melanogaster*. *Royal Society Open Science* **11,** 240408. doi:10.1098/rsos.240408 (June 2024).

42. Belkina, E. G., Shiglik, A., Sopilko, N. G., Lysenkov, S. N. & Markov, A. V. Mate Choice Copying in *Drosophila* Is Probably Less Robust than Previously Suggested. *Animal Behaviour* **176,** 175–183. doi:10.1016/j.anbehav.2021.04.007 (2021).

43. Bierbach, D., Kronmarck, C., Hennige-Schulz, C., Stadler, S. & Plath, M. Sperm Competition Risk Affects Male Mate Choice Copying. *Behavioral Ecology and Sociobiology* **65,** 1699–1707. doi:10.1007/s00265-011-1177-3 (2011).

44. Cirino, L. A., Gallagher, I. D., Desjonquères, C. & Rodríguez, R. L. S. Robust Mate Preferences despite Means and Opportunity for Mate Choice Copying in an Insect. *Animal Behaviour* **200,** 137–146. doi:10.1016/j.anbehav.2023.03.018 (2023).

45. Clemente, S. *Cultural Transmission of Sexual Preferences in* Drosophila melanogaster*: Testing for the Durability of Socially Learnt Preferences* M.Sc. Thesis (Universidade de Lisboa, 2012).

46. Duffy, A. G., Pew, M. G. & Johnson, J. B. Mate Choice Copying Behaviour in the Livebearing Fish Brachyrhaphis Rhabdophora. *Animal Behaviour* **222,** 123148. doi:10.1016/j.anbehav.2025.123148 (Apr. 2025).

47. Harrison-Weiss, A. *Female Choice and Social Decision-Making in a Marine Toadfish (*Porichthys notatus*)* M.Sc. Thesis (McMaster University, 2024). HDL: 11375/30296.

48. Kavaliers, M. *et al.* Conspecific Infection Threat Rapidly Biases the Social Responses of Female Mice: Involvement of Oxytocin. *Hormones and Behavior* **113,** 67–75. doi:10.1016/j.yhbeh.2019.04.016 (2019).

49. Kavaliers, M., Bishnoi, I. R., Ossenkopp, K. P. & Choleris, E. Odor-Based Mate Choice Copying in Deer Mice Is Not Affected by Familiarity or Kinship. *Animal Cognition* **25,** 241–248. doi:10.1007/s10071-021-01550-z (2022).

50. Kniel, N., Schmitz, J. & Witte, K. Quality of Public Information Matters in Mate-Choice Copying in Female Zebra Finches. *Frontiers in Zoology* **12,** 26–34. doi:10.1186/s12983-015-0119-8 (2015).

51. Kniel, N., Müller, K. & Witte, K. The Role of the Model in Mate-Choice Copying in Female Zebra Finches. *Ethology* **123,** 412–418. doi:10.1111/eth.12611 (2017).

52. Kobayashi, C. & Karino, K. Number Matters: Effect of the Number of Models on Mate-Choice Copying in Female Guppies. *Journal of Ethology* **44,** 105–112. doi:10.1007/s10164-025-00876-2 (2026).

53. MacLaren, R. D. Aesthetic Preference for Artificially Selected Color Variant Affects Mate Choice Copying Behavior in Female *Poecilia latipinna*. *PLOS One* **19,** NA. doi:10.1371/journal.pone.0298171 (2024).

54. MacLaren, R. D. Evidence of an Emerging Female Preference for an Artificial Male Trait and the Potential for Spread via Mate Choice Copying in *Poecilia latipinna*. *Ethology* **125,** 575–586. doi:10.1111/eth.12885 (2019).

55. Marcos, J. P. *Regulação Do Mate-Choice Copying Por Neurónios Apetitivos Ou Neurónios Aversivos de Dopamina Em* Drosophila melanogaster M.Sc. Thesis (ISPA, 2019).

56. Miller, A. M. *Testing for a Link between Personality and Mate-Choice Copying Tendency in the Trinidadian Guppy (*Poecilia reticulata*)* M.Sc. Thesis (Carleton Univesrity, 2013).

57. Mitchell, B. *et al.* Chemical Mate Choice Copying in *Drosophila melanogaster*. *American Naturalist* **206,** 171–178. doi:10.1086/736329 (2025).

58. Moran, R. L., Von Ende, C. N. & King, B. H. Mate Choice Copying in Two Species of Darters (Percidae: *Etheostoma*). *Behaviour* **150,** 1255–1274. doi:10.1163/1568539X-00003092 (2013).

59. Nöbel, S., Monier, M., Villa, D., Danchin, E. G. & Isabel, G. 2-D Sex Images Elicit Mate Copying in Fruit Flies. *Scientific Reports* **12,** 22127. doi:10.1038/s41598-022-26252-5 (2022).

60. Nöbel, S. *et al.* Female Fruit Flies Copy the Acceptance, but Not the Rejection, of a Mate. *Behavioral Ecology* **33,** 1018–1024. doi:10.1093/beheco/arac071 (2022).

61. Nöbel, S., Danchin, E. G. & Isabel, G. Long-Term Social Memory of Mate Copying in *Drosophila melanogaster* Is Localized in Mushroom Bodies. *Scientific Reports* **15,** 5262. doi:10.1038/s41598-025-88535-x (2025).

62. Nöbel, S. & Kaufmann, T. E. Mate Copying in *Drosophila simulans*. *Biology Letters* **21,** 20250070. doi:10.1098/rsbl.2025.0070 (2025).

63. Nöbel, S., Danchin, E. & Isabel, G. Mate Copying Requires the Coincidence Detector Rutabaga in the Mushroom Bodies of *Drosophila melanogaster*. *iScience* **26,** 107682. doi:10.1016/j.isci.2023.107682 (2023).

64. Nöbel, S. *et al.* No Evidence for Mate Copying in *Danio rerio*. *Behavioural Processes* **206,** 104837. doi:10.1016/j.beproc.2023.104837 (2023).

65. Ophir, A. G. *Mate Assessment and Non-Independent Mate Choice by Female Japanese Quail* PhD thesis (McMaster University, 2009). HDL: 11375/13081.

66. Pusiak, R. J., Auld, H. L. & Godin, J. G. J. Relative Sexual Attractiveness Does Not Influence Mate-Choice Copying in Male Trinidadian Guppies, *Poecilia reticulata*. *Animal Behaviour* **162,** 123–133. doi:10.1016/j.anbehav.2020.02.008 (2020).

67. Silva, G. F. *Can Social Learning Promote Hybridisation?: Mate-Choice Copying in* Drosophila subobscura *Populations* M.Sc. Thesis (Universidade de Lisboa, 2014). HDL: 10451/15315.

68. Witte, K., Baumgärtner, K., Röhrig, C. & Nöbel, S. Test of the Deception Hypothesis in Atlantic Mollies *Poecilia mexicana*—Does the Audience Copy a Pretended Mate Choice of Others? *Biology* **7,** 40. doi:10.3390/biology7030040 (2018).

69. Ziege, M. *et al.* A Comparison of Two Methods to Assess Audience-Induced Changes in Male Mate Choice. *Current Zoology* **58,** 84–94. doi:10.1093/czoolo/58.1.84 (2012).

70. Nakagawa, S. & Cuthill, I. Effect Size, Confidence Interval and Statistical Significance: A Practical Guide for Biologists. *Biological Reviews* **82,** 591–605. doi:10.1111/j.1469-185X.2007.00027.x (2007).

71. Danchin, E. *et al.* Cultural Flies: Conformist Social Learning in Fruitflies Predicts Long-Lasting Mate-Choice Traditions. *Science* **362,** 1025–1030. doi:10.1126/science.aat1590 (2018).

72. Galef, B. G. & Laland, K. N. Social Learning in Animals: Empirical Studies and Theoretical Models. *Bioscience* **55,** 489–499. doi:10.1641/0006-3568(2005)055[0489:SLIAES]2.0.CO;2 (2005).

73. Uehara, T., Yokomizo, H. & Iwasa, Y. Mate-Choice Copying as Bayesian Decision Making. *American Naturalist* **165,** 403–410. doi:10.1086/428301. PMID: 15729669 (2005).

74. Laland, K. N. Social Learning Strategies. *Learning and Behavior* **32,** 4–14. doi:10.3758/BF03196002 (2004).

75. Loyau, A., Blanchet, S., Van Laere, P., Clobert, J. & Danchin, E. When Not to Copy: Female Fruit Flies Use Sophisticated Public Information to Avoid Mated Males. *Scientific Reports* **2,** 768. doi:10.1038/srep00768 (2012).

76. Dugatkin, L. A., Druen, M. W. & Godin, J.-G. J. The Disruption Hypothesis Does Not Explain Mate-Choice Copying in the Guppy (*Poecilia reticulata*). *Ethology* **109,** 67–76. doi:10.1046/j.1439-0310.2003.00851.x (2003).

77. Nöbel, S. *et al.* The Importance of Population Heterogeneities in Detecting Social Learning as the Foundation of Animal Cultural Transmission. *Proceedings of the Royal Society B: Biological Sciences* **289,** 20220431. doi:10.1098/rspb.2022.0431 (2022).

78. Bridges, A. D. & Chittka, L. Animal Behaviour: Conformity and the Beginnings of Culture in an Insect. *Current Biology* **29,** R167–R169. doi:10.1016/j.cub.2019.01.023 (2019).

79. Jennions, M. & Møller, A. Relationships Fade with Time: A Meta-Analysis of Temporal Trends in Publication in Ecology and Evolution. *Proceedings of the Royal Society B: Biological Sciences* **269,** 43–48. doi:10.1098/rspb.2001.1832 (2002).

80. Koricheva, J. & Kulinskaya, E. Temporal Instability of Evidence Base: A Threat to Policy Making? *Trends in Ecology & Evolution* **34,** 895–902. doi:10.1016/j.tree.2019.05.006 (2019).

Supplementary Information

Mate choice copying in non-human animals: an update of two meta-analyses

Eduardo S. A. Santos*, on behalf of all authors[1]

[1]Collaboration for Open Science and Synthesis in Ecology and Evolution (COSSEE), Department of Biological Sciences, Faculty of Science, University of Alberta, Edmonton, Canada

This document contains the supplementary search strings, results, figures and tables referred to in the main article. Figures and tables are numbered separately from the main article (Fig. S1, Table S1, and so on) and each is cited at least once in the main text.

Contents

*Corresponding author: esantos2@ualberta.ca

Search strings

All searches were conducted across the full range of available publication dates in all databases and search platforms (1900 to 2026 inclusive). Search strings were adapted to the structure of each database or search platform.

Web of Science (combined and optimised search terms):

- Access date: 04 February 2026;
- Query: TS = ((”mate choice” AND ”copy*”) OR (”mate-choice” AND ”copy*”) OR ”mate-copy*” OR ”mate copy*” OR ”female* copy*”); and
- Total records retrieved: 891.

Scopus (Davies et al. search terms):

- Access date: 23 February 2026;
- Query: TITLE-ABS-KEY(”mate choice” AND (copying OR nonindependent OR learning)); and
- Total records retrieved: 718.

Scopus (Jones and DuVal search terms):

- Access date: 04 February 2026;
- Query: TITLE-ABS-KEY ((”mate choice” AND ”copy*”) OR (”mate-choice” AND ”copy*”) OR ”mate-copy*” OR ”mate copy*” OR ”female* copy*”); and
- Total records retrieved: 474.

BASE:

- Access date: 05 February 2026;
- Query: (”mate choice” AND ”copy*”) OR (”mate-choice” AND ”copy*”) OR ”mate-copy*” OR ”mate copy*” OR ”female* copy*”; and
- Total records retrieved: 212.

SciELO:

- Access date: 06 February 2026;

- Query: (title:(”mate choice” OR ”mate copying” OR ”escolha de parceiro” OR ”cópia de escolha”) OR abstract:(”mate choice” OR ”mate copying” OR ”escolha de parceiro” OR ”cópia de escolha”) OR subject:(”mate choice” OR ”mate copying” OR ”escolha de parceiro” OR ”cópia de escolha”)); and
- Total records retrieved: 11.

Google Scholar (Japanese):

- Access date: 26 February 2026;
- Query: ” 配偶者選択” ” コピー” OR ” 交尾選択” ” コピー” OR ” 選択のコピー” OR ” 配偶者のコピー”; and
- Total records retrieved: 2.

Google Scholar (Polish):

- Access date: 26 February 2026;
- Query: ”wybór partnera” ”kopiowanie” OR ”wybór do rozrodu” ”kopiowanie” OR ”kopiowanie wyboru” OR ”kopiowanie partnera”; and
- Total records retrieved: 77.

Google Scholar (Portuguese):

- Access date: 26 February 2026;
- Query: ”escolha de parceiro” ”cópia” OR ”escolha de acasalamento” ”cópia” OR ”cópia de escolha” OR ”cópia de parceiro”; and
- Total records retrieved: 89.

Google Scholar (Russian):

- Access date: 26 February 2026;
- Query: "выбор партнера" "копирование" OR "выбор для спаривания" "копирование" OR "копирование выбора" OR "копирование партнера"; and
- Total records retrieved: 70.

Google Scholar (Simplified Chinese):

- Access date: 26 February 2026;
- Query: ”配偶选择” ”复制” OR ”交配选择” ”复制” OR ”选择复制” OR ”配偶复制”; and
- Total records retrieved: 93.

Google Scholar (Spanish):

- Access date: 26 February 2026;
- Query: ”elección de pareja” ”copia” OR ”elección de apareamiento” ”copia” OR ”copia de elección” OR ”copia de pareja”; and
- Total records retrieved: 218.

Supplementary Methods

These are the detailed specifications of the small-study-effect, time-lag, and sensitivity analyses summarised in the Methods of the main article.

Multilevel Egger's regression and the conservative bias-corrected mean

We tested whether study precision moderated the effect size using a two-step procedure [1–3]: (1) we fitted a multilevel analogue of Egger's regression using the standard error as the precision proxy, where v_i is the sampling variance of the ith effect size and $SE_i = \sqrt{v_i}$; and (2) where the slope of that model was statistically significant—indicating a biased estimate—we re-fitted the model with the sampling variance (v_i) as the moderator, whose intercept provides the bias-adjusted overall meta-analytic mean. We used the standard error rather than the effective sample size as the precision moderator. For standardised mean differences and log odds ratios the sampling variance is partly a function of the effect-size estimate itself, so a regression on $\sqrt{v_i}$ can register asymmetry that is in part an artefact of that dependence [1]. Because this artefact inflates rather than deflates apparent small-study effects, we read the resulting adjusted mean as a conservative bound on the true average effect rather than as a point estimate—the more pessimistic of our two correction frameworks.

Bias-robust mean and its assumptions

The second correction framework follows Yang et al. [4]: a bias-robust weighting scheme combined with cluster-robust (CR2) variance estimation. The weighting scheme is a multivariate fixed-effect generalised-least-squares estimator in which the weights derive from the sampling variance–covariance matrix (assuming $\rho = 0.5$, as elsewhere); inverse-covariance weighting downweights imprecise studies reporting large effects and prevents a study contributing many correlated effect sizes from accruing weight in proportion to their number, making the resulting mean less susceptible to selective reporting than a random-effects mean. Cluster-robust variance estimation, implemented with `clubSandwich` and clustered by study, then restores valid inference under the non-independence of effect sizes within studies. Because this first-stage model contains no random effects, the CR2 small-sample correction is applicable, which it would not be for a multilevel model with crossed random terms [5]. It also follows, however, that dependence among effect sizes from the same species—crossed with, rather than nested within, study—is represented by neither the weighting nor the clustering. We therefore read the bias-robust mean as a point estimate that is robust to selective reporting but whose interval may be optimistic where species contribute effect sizes across several studies; our multilevel models, which represent species and phylogeny explicitly, supply the complementary uncorrected and bias-corrected estimates. Both corrected means are overlaid on the orchard plots in the main article using `orchaRd::pub_bias_plot()`.

Leave-one-out analyses

We assessed the influence of individual studies and species by re-fitting the combined multilevel model repeatedly, each time removing all effect sizes from a single study (leave-one-study-out) or from a single species (leave-one-species-out), and comparing each re-estimated pooled mean with the full-data estimate. Every refit used the same model structure as the main analysis: the sampling variance–covariance matrix was recomputed for the reduced dataset (assuming a within-study correlation of $\rho = 0.5$), and the phylogenetic random effect was retained, with the species correlation matrix subset to the species that remained. We ran both analyses for the continuous (Hedges' *g*) and binary (log odds ratio) syntheses using the `leave_one_out` and `orchard_leave1out` functions in the `orchaRd` package [6]. Results are shown in Supplementary Figures S4–S7.

Cross-dataset validation and the contentious effect sizes

For every study represented in both source datasets we compared the Hedges' *g* reported by Davies et al. [7] with the *g* implied by converting the odds ratio reported by Jones and DuVal [8] for the same experiment and outcome (Eqn 18 of the main article). This flagged seven studies whose two values disagreed by more than 0.5 Hedges' *g* units. We refitted each combined model under three versions of the data, all using the same multilevel structure as the main analysis: the original data as analysed (*Main*), a conservative version that discarded all effect sizes from the seven flagged studies (*Leave-out*), and a best-estimate version that retained every flagged study but substituted outcomes re-extracted from the original sources (*Corrected*). The effect-size count changes slightly relative to *Main*, because the correction was not a one-to-one value swap: three over-counted binary (odds ratio) rows were removed, and on the continuous side, one effect size that had been pooled was kept as the original two outcomes reported, while another's non-target effects were not reinstated. The results of these sensitivity analyses are reported under Supplementary Results and shown in Supplementary Figures S8 and S9.

Supplementary Results

Small-study effects, time-lag patterns, and sensitivity analyses

We found strong evidence of small-study effects, while mixed evidence of time-lag bias using the combined dataset (original + new primary studies). Funnel plots based on model residuals showed clear asymmetry for both types of effect size (Supplementary Figure S2 and Supplementary Figure S3). The Egger's-type multilevel regression slopes of sampling variance were positive and statistically significant for both datasets (Odds ratio: $\beta_{\sqrt{v_i}}$ = 2.11, 95% CI: 1.23 to 2.99, Fig. 6 of the main article; Hedges' *g*: $\beta_{\sqrt{v_i}}$ = 4.84, 95% CI: 3.59 to 6.09, Fig. 7 of the main article). For the odds ratio dataset, we found evidence that publication year was negatively associated with effect size magnitude (β_{year} = -0.028, 95% CI: -0.048 to -0.007; Fig. 6 of the main article). However, for the Hedges' *g* dataset, we found little evidence that publication year was negatively associated with effect size magnitude (β_{year} = -0.005, 95% CI: -0.018 to 0.008; Fig. 7 of the main article).

The two correction frameworks we used gave divergent adjusted means and provided smaller estimates than the uncorrected overall estimates. On the Hedges' *g* scale, the bias-corrected

intercept [1] fell to -0.21 (95% CI: -0.48 to 0.06), a mean indistinguishable from zero, whereas the bias-robust estimator [4] retained a positive but reduced mean of 0.34 (95% CI: 0.25 to 0.42). The odds ratio scale showed the same pattern: the bias-corrected mean approached the null at a log odds ratio of 0.10 (95% CI: -0.23 to 0.43; *i.e.*, OR = 1.11), while the bias-robust mean remained clearly positive at a log odds ratio of 0.55 (95% CI: 0.39 to 0.71; OR = 1.73). Both corrected means are shown alongside the uncorrected estimate in Fig. 3 of the main article and Fig. 4 of the main article.

Leave-one-out analyses confirmed the robustness of both the overall odds ratio and Hedges' *g* on the combined datasets. Sequential exclusion of individual studies yielded consistent pooled estimates, with every confidence interval overlapping the full-sample estimate and no single study exerting disproportionate influence. The largest shift in the pooled mean was only 0.02 Hedges' *g* units for the continuous outcomes and 0.04 log-odds units for the binary outcomes (Supplementary Figure S4 and Supplementary Figure S5). The results are also robust to the more rigorous leave-one-species-out analyses, which remove entire species, despite the uneven phylogenetic sampling, with all overall effect sizes remaining significantly positive (Supplementary Figure S6 and Supplementary Figure S7). Here, the largest shift was 0.04 Hedges' *g* units (when omitting *Poecilia latipinna*) and 0.07 log-odds units. Removing *Poecilia reticulata*—the most-studied and strongest-copying species, which alone contributes nearly a quarter of the continuous effect sizes—nudged the Hedges' *g* mean toward zero, as expected, but did not change the sign or significance of either synthesis.

The combined estimates showed little sensitivity to the contentious effect sizes on both scales (Hedges' *g* or odds ratio). The pooled Hedges' *g* synthesis was essentially unchanged, the original 0.45 (95% CI: 0.25 to 0.66, $k = 227$) became 0.41 (95% CI: 0.22 to 0.60, $k = 213$) when the flagged studies were dropped, and 0.42 (95% CI: 0.23 to 0.60, $k = 225$) when they were corrected, remaining significant throughout (Supplementary Figure S8). The odds ratio synthesis was also stable: the pooled odds ratio of 1.81 (95% CI: 1.34 to 2.45, $k = 172$) became 1.77 (95% CI: 1.32 to 2.38, $k = 156$) under leave-out and 1.79 (95% CI: 1.36 to 2.34, $k = 169$) under correction (Supplementary Figure S9).

Supplementary Figures

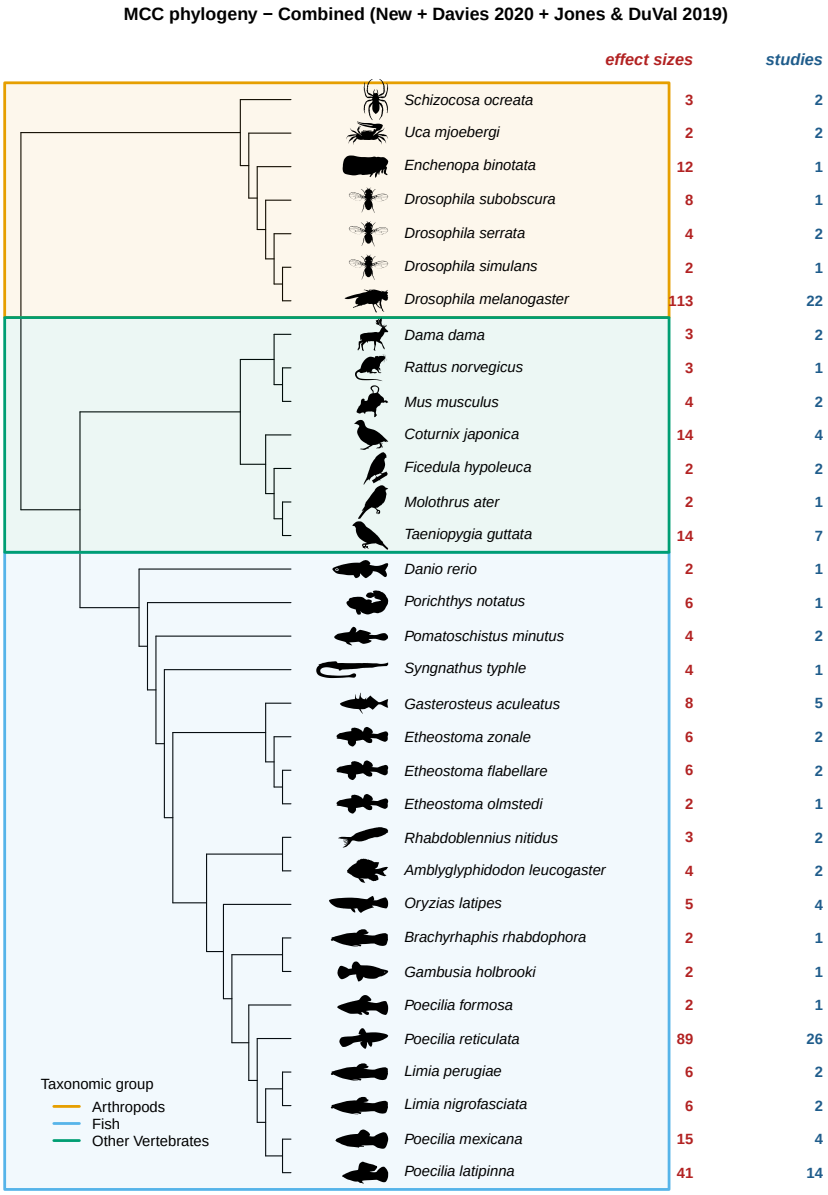

Figure S1: **Phylogenetic tree.** The phylogenetic tree of the species included in the meta-analysis using the combined dataset of original and new primary studies, along with the number of effect sizes (*k*) and the number of studies in which they are reported.

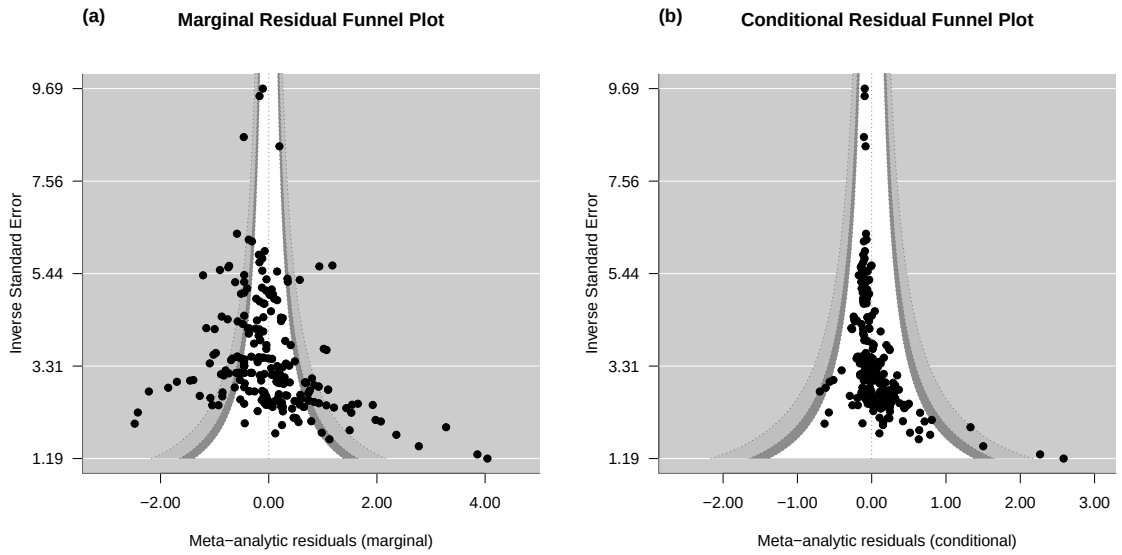

Figure S2: **Funnel plots of Hedges' g data.** Both panels plot each effect size against its precision ($1/SE$), but differ in what is removed before plotting: (**a**) Marginal residuals are the deviations of each observed effect from the fixed-effect (intercept, the overall mean) prediction, so they still carry all of the random-effects heterogeneity (between-study, between-effect-size, and between-species–*i.e.* the hierarchical non-independence of effect sizes). Asymmetry here (defined in Methods) is the classic visual signature of small-study and publication bias. (**b**) Conditional residuals additionally subtract the estimated random intercepts (study, effect-size, and species BLUPs), so the heterogeneity captured by the multilevel structure is partialled out. A symmetric, evenly-filled funnel here means that once the multilevel structure is accounted for, there is little precision-related asymmetry left, *i.e.*, the asymmetry seen in (**a**) is largely explained by the modelled random effects rather than by selective reporting. Any asymmetry that persists in (**b**) is the more credible visual indication of true small-study effects.

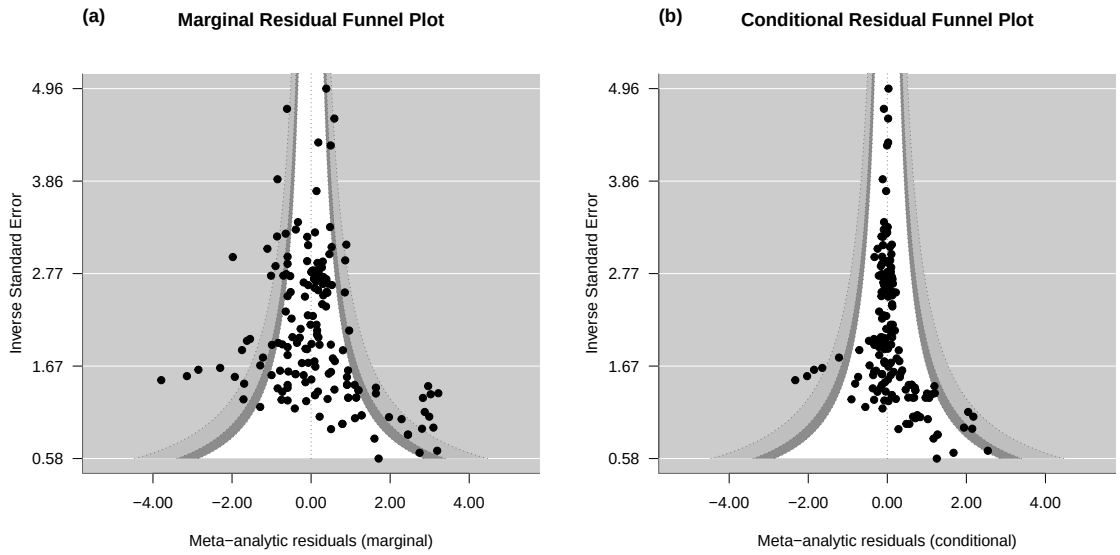

Figure S3: **Funnel plots of odds ratio data.** Both panels plot each effect size against its precision ($1/SE$), but differ in what is removed before plotting: (**a**) Marginal residuals are the deviations of each observed effect from the fixed-effect (intercept, the overall mean) prediction, so they still carry all of the random-effects heterogeneity (between-study, between-effect-size, and between-species–*i.e.* the hierarchical non-independence of effect sizes). Asymmetry here (defined in Methods) is the classic visual signature of small-study and publication bias. (**b**) Conditional residuals additionally subtract the estimated random intercepts (study, effect-size, and species BLUPs), so the heterogeneity captured by the multilevel structure is partialled out. A symmetric, evenly-filled funnel here means that once the multilevel structure is accounted for, there is little precision-related asymmetry left, *i.e.*, the asymmetry seen in (**a**) is largely explained by the modelled random effects rather than by selective reporting. Any asymmetry that persists in (**b**) is the more credible visual indication of true small-study effects.

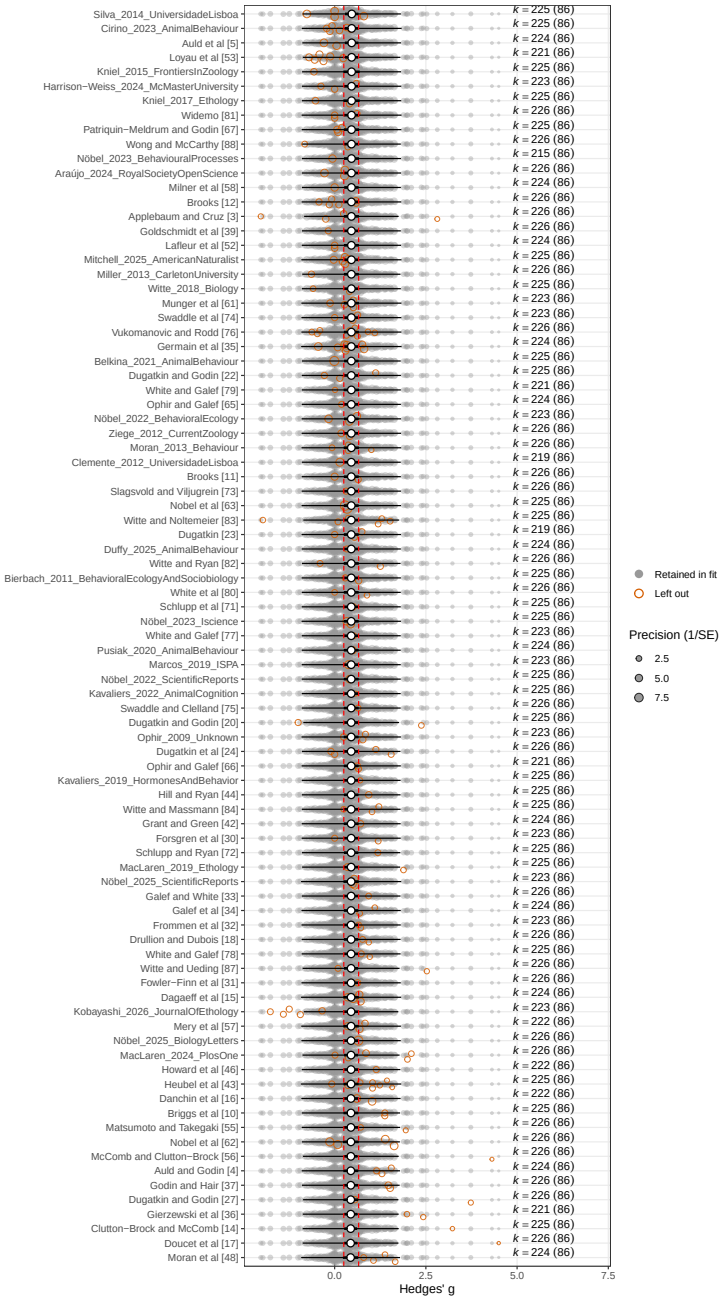

Figure S4: **Leave-one-study-out analysis of the combined Hedges' *g*.** Each row re-estimates the multilevel pooled mean after removing one study from the continuous (Hedges' *g*) dataset, with rows ordered by the leave-one-out estimate so the panel reads as an influence ranking. Grey-filled points are the effect sizes retained in that fit, and open orange rings are the effect sizes that were removed, each scaled by its precision $(1/SE)$. The white marker is the re-estimated pooled mean, the thick black bar its 95% confidence interval and the thin black whisker its 95% prediction interval. The two vertical dashed red lines mark the full-sample 95% confidence interval, so a mean falling between them indicates that omitting that study leaves the pooled estimate essentially unchanged. Labels on the right give the number of effect sizes retained (k), with the number of studies remaining in parentheses.

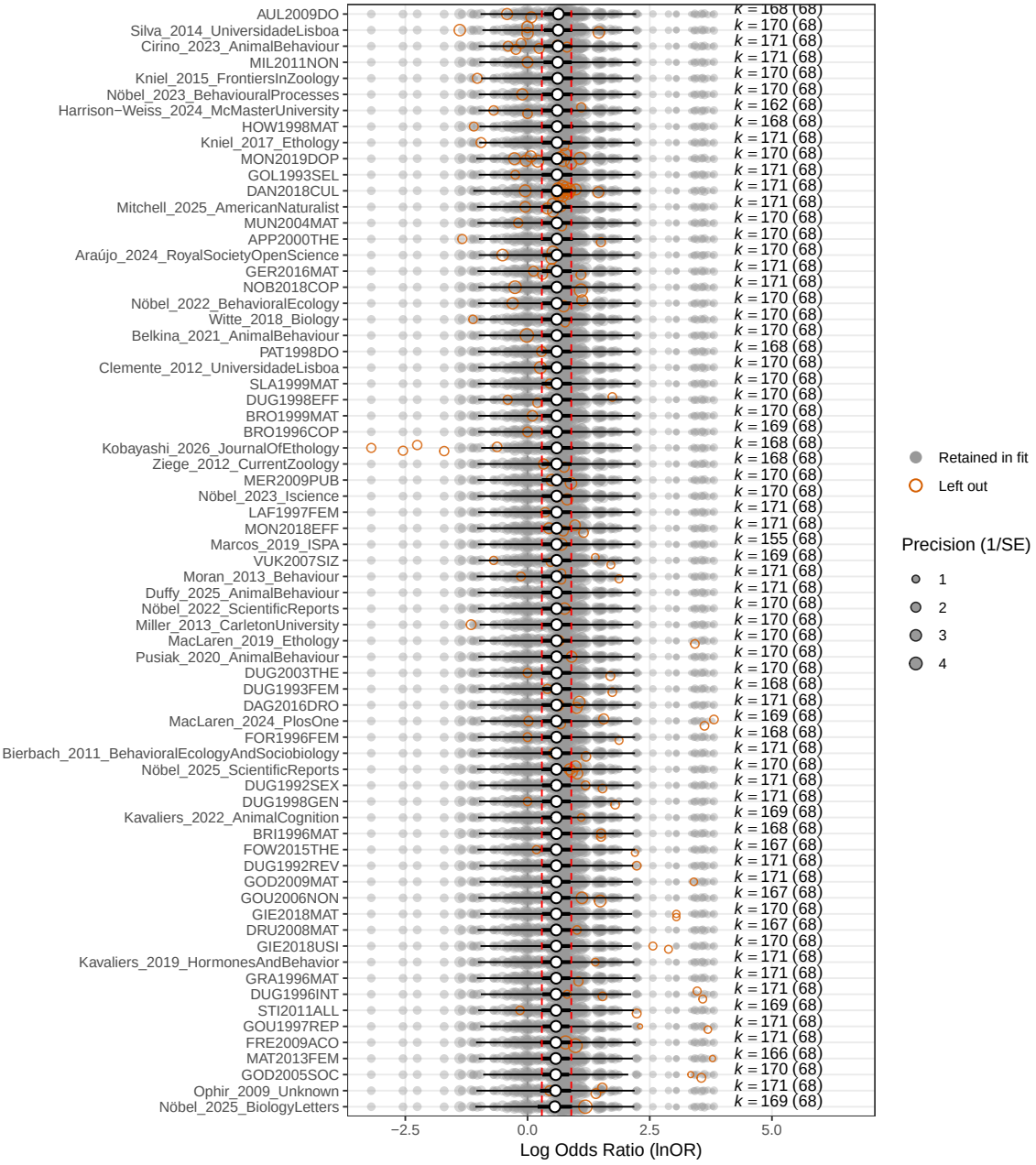

Figure S5: **Leave-one-study-out analysis of the combined odds ratio.** Each row re-estimates the multilevel pooled mean (on the log-odds-ratio scale) after removing one study from the binary dataset, with rows ordered by the leave-one-out estimate. Grey-filled points are the effect sizes retained in that fit, and open orange rings are the effect sizes that were removed, each scaled by its precision ($1/SE$). The white marker is the re-estimated pooled mean, the thick black bar its 95% confidence interval, and the thin black whisker its 95% prediction interval. The two vertical dashed red lines mark the full-sample 95% confidence interval, so a mean falling between them indicates that omitting that study leaves the pooled estimate essentially unchanged. Labels on the right give the number of effect sizes retained (k), with the number of studies remaining in parentheses.

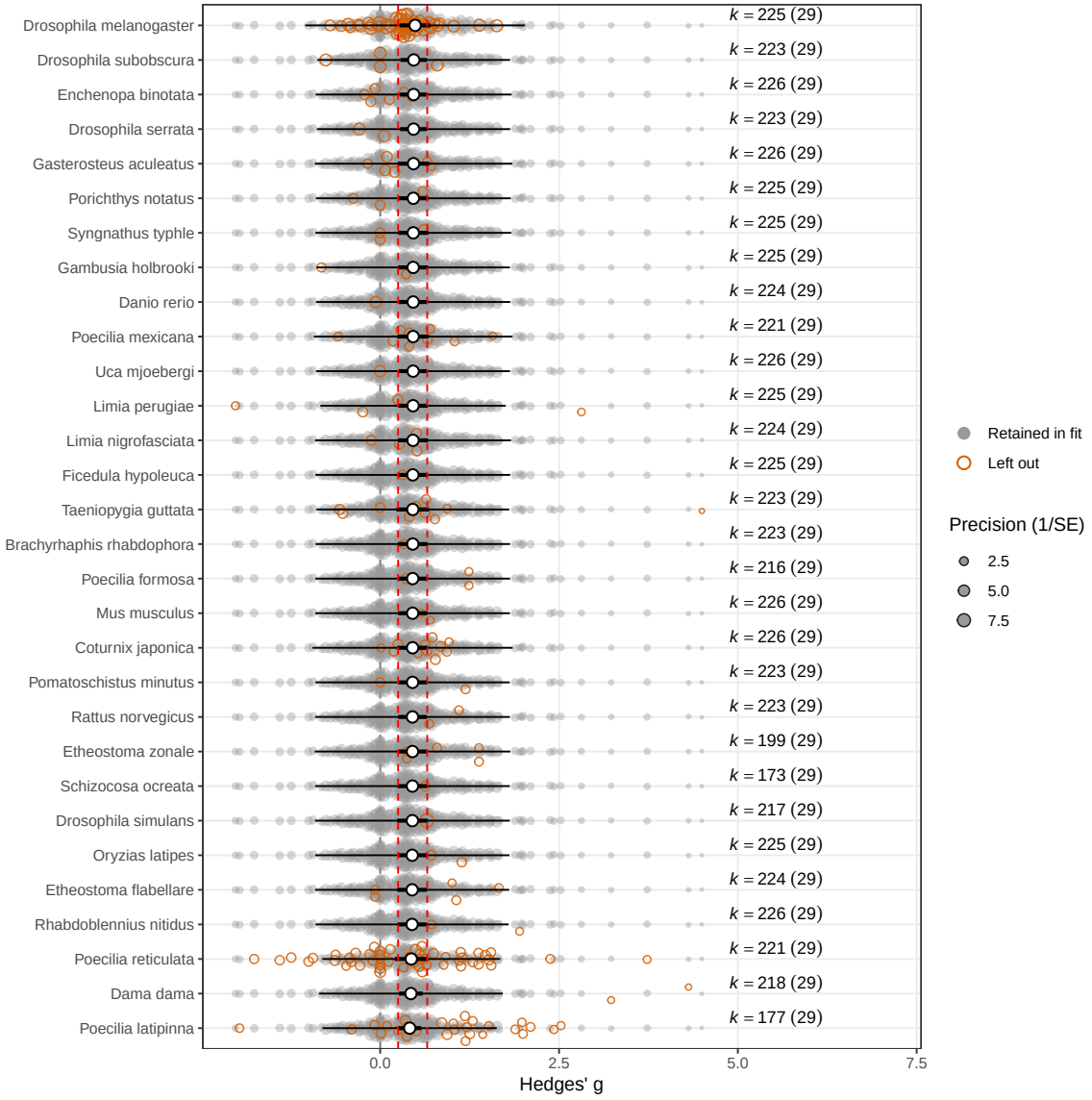

Figure S6: **Leave-one-species-out analysis of the combined Hedges' *g*.** Each row re-estimates the multilevel pooled mean after removing all effect sizes from one species in the continuous (Hedges' *g*) dataset—a stricter test than leaving out single studies, because the phylogenetic sampling is uneven. Rows are ordered by the leave-one-out estimate. Grey-filled points are the effect sizes retained in that fit, and open orange rings are the effect sizes contributed by the omitted species, each scaled by its precision ($1/SE$). The white marker is the re-estimated pooled mean, the thick black bar its 95% confidence interval, and the thin black whisker its 95% prediction interval. The two vertical dashed red lines mark the full-sample 95% confidence interval. Labels on the right give the number of effect sizes retained (k), with the number of species remaining in parentheses.

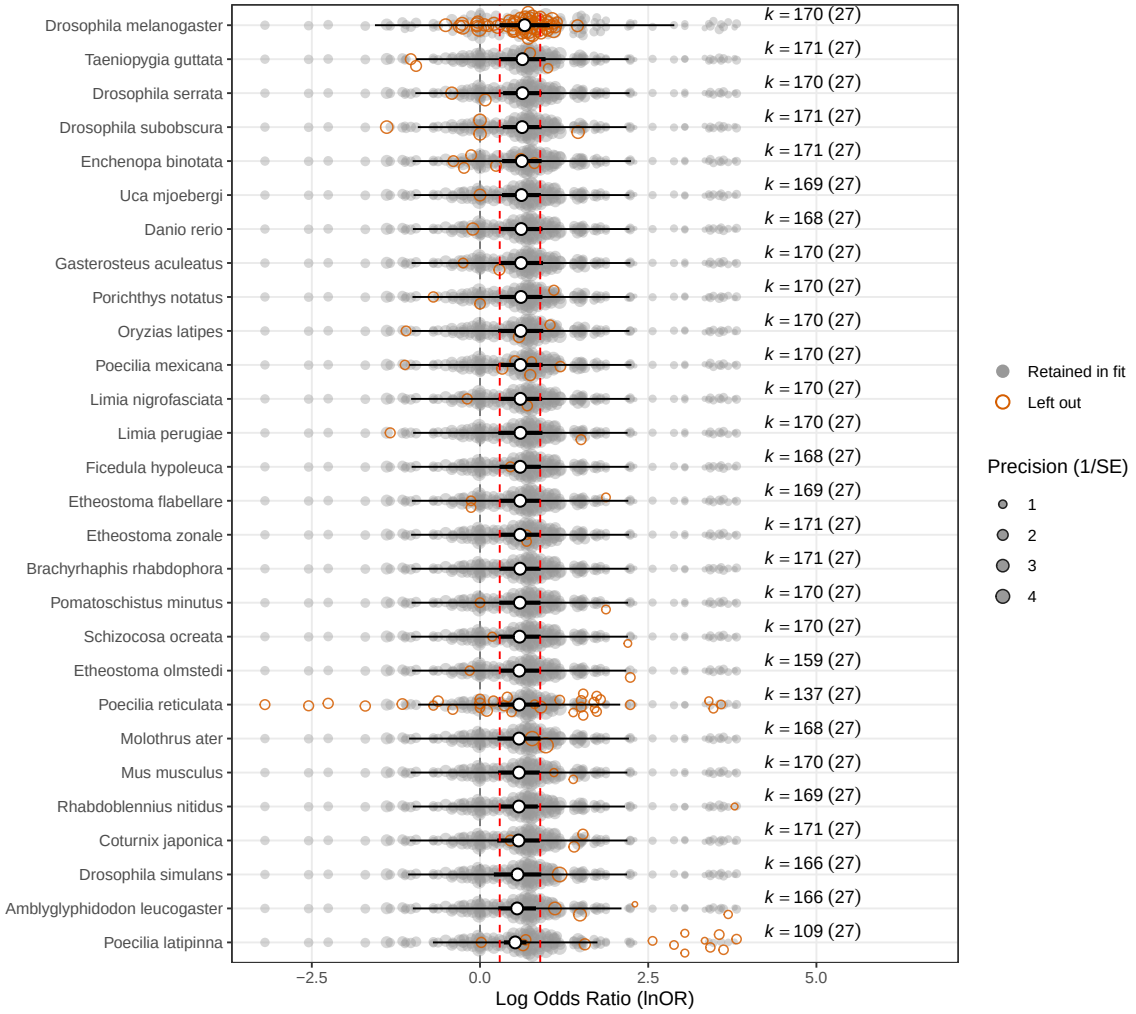

Figure S7: **Leave-one-species-out analysis of the combined odds ratio.** Each row re-estimates the multilevel pooled mean (on the log-odds-ratio scale) after removing all effect sizes from one species in the binary dataset—a stricter test than leaving out single studies, because the phylogenetic sampling is uneven. Rows are ordered by the leave-one-out estimate. Grey-filled points are the effect sizes retained in that fit, and open orange rings are the effect sizes contributed by the omitted species, each scaled by its precision ($1/SE$). The white marker is the re-estimated pooled mean, the thick black bar its 95% confidence interval, and the thin black whisker its 95% prediction interval. The two vertical dashed red lines mark the full-sample 95% confidence interval. Labels on the right give the number of effect sizes retained (k), with the number of species remaining in parentheses.

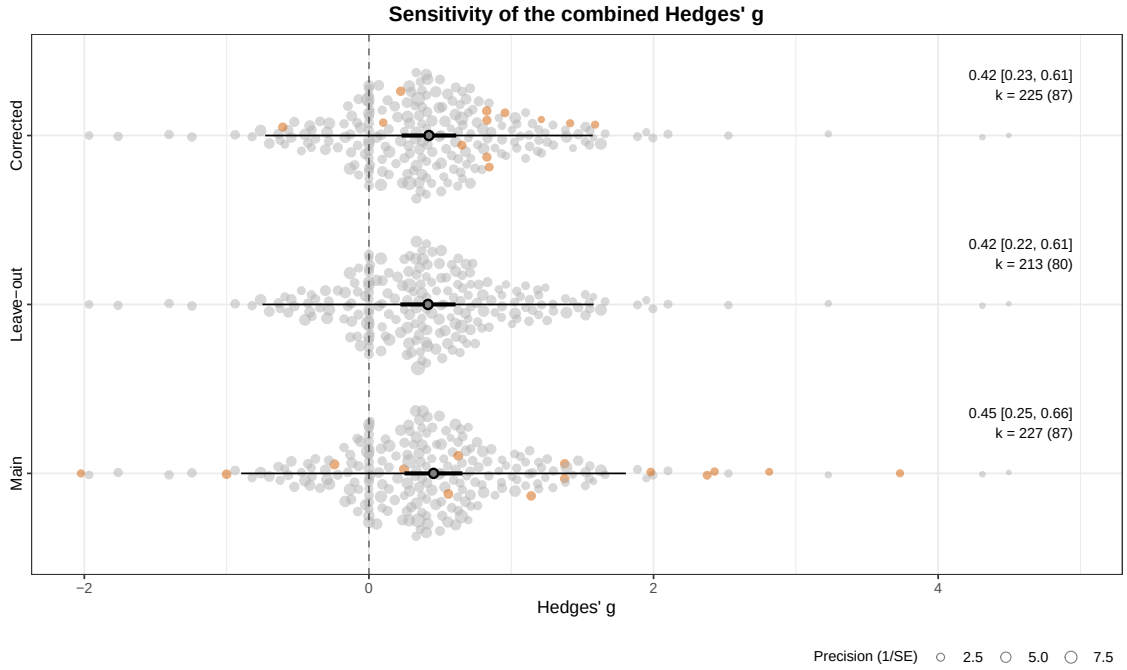

Figure S8: **Sensitivity of the combined Hedges' *g* to the contentious effect sizes.** Each row shows the combined random-effects estimate for the continuous (Hedges' *g*) dataset under one version of the data: the original analysis (*Main*), the conservative version with all seven flagged studies removed (*Leave-out*), and the best-estimate version with their effect sizes replaced by re-extractions (*Corrected*). Points are individual effect sizes scaled by their precision ($1/SE$), with the effect sizes from the seven contentious studies highlighted in orange and all others in grey; the central marker is the pooled mean, the thick bar its 95% confidence interval, and the thin whisker its 95% prediction interval. The vertical dashed line marks the no-effect line ($g = 0$). Annotations give the pooled mean [95% CI] and the number of effect sizes (k), with the number of studies in parentheses.

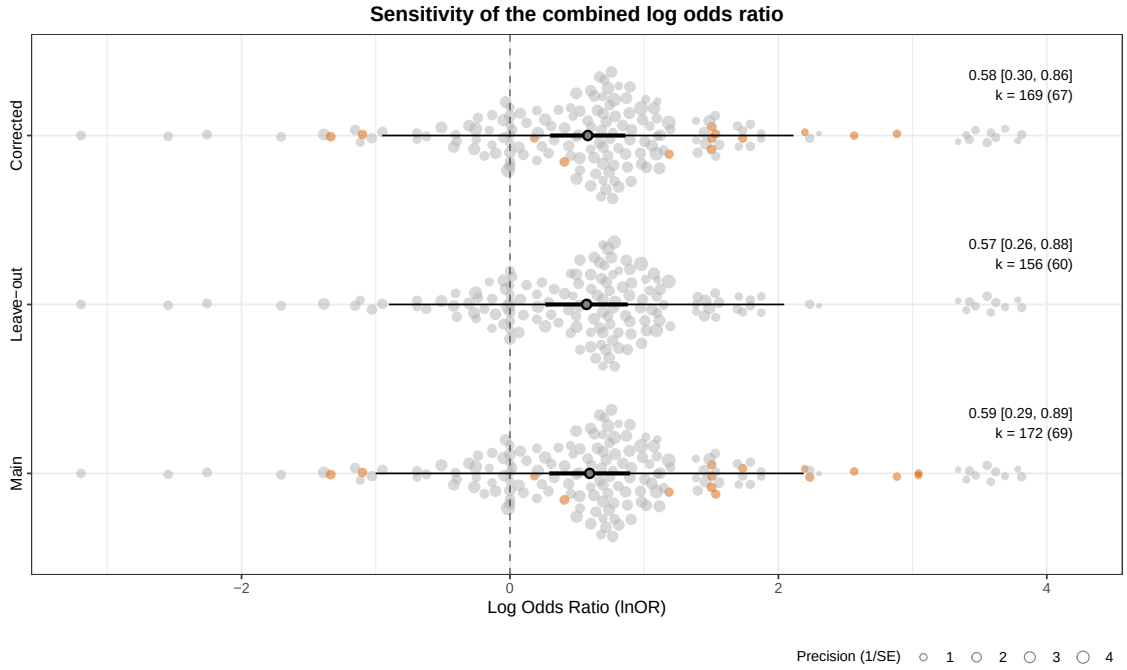

Figure S9: **Sensitivity of the combined odds ratio to the contentious effect sizes.** Each row shows the combined random-effects estimate for the binary dataset (on the log-odds-ratio scale) under one version of the data: the original analysis (*Main*), the conservative version with all seven flagged studies removed (*Leave-out*), and the best-estimate version with their effect sizes replaced by re-extractions (*Corrected*). Points are individual log-odds-ratio effect sizes scaled by their precision ($1/SE$), with the effect sizes from the contentious studies highlighted in orange and all others in grey; the central marker is the pooled mean, the thick bar its 95% confidence interval, and the thin whisker its 95% prediction interval. The vertical dashed line marks no effect (lnOR $= 0$, i.e. OR $= 1$). Annotations give the pooled mean [95% CI] and the number of effect sizes (k), with the number of studies in parentheses.

Supplementary Tables

Table S1: PRISMA-EcoEvo [9] reporting checklist for the present systematic review and meta-analysis. "Reported" status: Yes = fully reported; Partly = reported but incomplete; No = not reported; NA = not applicable. Page numbers refer to the main manuscript; "SI" denotes the supplementary information.

Item	Sub-item	Reported?	Location / notes
Title and abstract			
1.1	Identify the review as a systematic review, meta-analysis, or both	Yes	Abstract identifies it as a meta-analysis update (p. 2)
1.2	Summarise the aims and scope of the review	Yes	Abstract (p. 2)
1.3	Describe the data set	Yes	Abstract: 69 new effect sizes from 29 studies combined with original datasets (p. 2)
1.4	State the results of the primary outcome	Yes	Abstract: positive, significant effect on both scales (p. 2)
1.5	State conclusions	Yes	Abstract: effect is modest and context-dependent (p. 2)
1.6	State limitations	Yes	Abstract notes high heterogeneity and small-study effects (p. 2)
Aims and questions			
2.1	Provide a rationale for the review	Yes	Introduction (p. 3–4)
2.2	Reference any previous reviews or meta-analyses on the topic	Yes	Update of Jones & DuVal and Davies et al. (p. 3–4)
2.3	State the aims and scope of the review (including its generality)	Yes	Introduction / objectives (p. 3–4)
2.4	State the primary questions the review addresses (e.g. which moderators were tested)	Yes	Two research questions (p. 4); moderators specified (p. 13)
2.5	Describe whether effect sizes were derived from experimental and/or observational comparisons	Yes	PECOS study design: experimental audience-effect designs (p. 5)
Review registration			

continued on next page

Table S1 continued from previous page

Item	Sub-item	Reported?	Location / notes
3.1	Register aims, hypotheses, and methods in a time-stamped, publicly accessible archive and link it in the methods section	Yes	Pre-registered on OSF (p. 4, 6); protocol cited in reference list and URL given in the methods text
3.2	Describe deviations from the registered aims and methods	Yes	"Additions and deviations" section (p. 4)
3.3	Justify deviations from the registered aims and methods	Yes	Rationales given for each deviation (p. 4)
Eligibility criteria			
4.1	Report the specific criteria used for including/excluding studies at title/abstract and full-text screening	Yes	Numbered inclusion criteria + PECOS framework (p. 5–6)
4.2	Justify criteria, if necessary	Yes	Criteria follow the two source meta-analyses and study scope (p. 5–6)
Finding studies			
5.1	Define the type of search (e.g. comprehensive search, representative sample)	Yes	Comprehensive, multilingual, grey-literature–inclusive search (p. 5)
5.2	State what sources of information were sought	Yes	Scopus, ISI Web of Science, Google Scholar, SciELO, BASE; grey literature (p. 5)
5.3	Include, for each database, the exact search strings used, with keyword combinations and Boolean operators	Partly	Full search strings provided in the SI; databases named in main text (p. 5, SI)
5.4	Provide enough information to repeat the search, including the timespan covered (start and end dates)	Partly	Publication timespan 1900–2026 stated (p. 5, SI); exact date each source was last searched not reported
Study selection			
6.1	Describe how studies were selected at each stage of screening (e.g. decision trees, software)	Yes	Two-stage screening described (p. 6)
6.2	Report the number of people involved and how they contributed (e.g. independent parallel screening)	Partly	Disagreements resolved by consensus (p. 6); number of screeners and independence not explicitly stated

continued on next page

Table S1 continued from previous page

Item	Sub-item	Reported?	Location / notes
Data collection process			
7.1	Describe where in the reports data were collected from (e.g. text or figures)	Yes	Text and figures (p. 7)
7.2	Describe how data were collected (e.g. software used to digitise figures)	Yes	Data extracted from figures (Google Gemini), validated with metaDigitise (p. 7)
7.3	Describe moderator variables constructed from collected data	Yes	Copying mechanism, cue modality, observer mating status, relative demonstrator age (p. 7, 13)
7.4	Report how missing or ambiguous information was dealt with	Yes	Studies with missing/ambiguous data excluded; effect sizes derived from test statistics (p. 7)
7.5	Report who collected data	Yes	Individuals who performed extraction identified (p. 7)
7.6	State the number of extractions checked for accuracy by co-authors	Yes	Percentage of records checked for accuracy reported (p. 7)
Data items			
8.1	Describe the key data sought from each study	Yes	Citation, experiment details, effect-size data (p. 7)
8.2	Describe items that do not appear in main results or could not be extracted	Partly	Non-convertible statistics (F, χ^2, GLMM outputs) excluded (p. 9)
8.3	Describe main assumptions or simplifications made	Yes	Equal-probability assumption ($C = D = n/2$) when no control; outcomes grouped as "mate-choice" (p. 10)
8.4	Describe the type of replication unit	Yes	Observation-level effect sizes nested within studies/species (p. 12)
Assessment of individual study quality			
9.1	Describe whether the quality of included studies was assessed	No	No formal study-level risk-of-bias/quality tool used (publication bias assessed instead)
9.2	Describe how information about study quality was incorporated into analyses	No	Not applicable; study-level quality not assessed

continued on next page

Table S1 continued from previous page

Item	Sub-item	Reported?	Location / notes
Effect size measures			
10.1	Describe effect size(s) used	Yes	Hedges' g and odds ratio (p. 7, 9)
10.2	Provide a reference to the equation of each effect size and its sampling variance	Yes	Equations with citations (Borenstein et al.); delta-method variance (p. 8–10)
10.3	If no reference exists, derive the equations and state the assumed sampling distribution(s)	Yes	Conversions derived, logistic-distribution assumption stated (p. 10)
Missing data			
11.1	Describe steps taken to deal with missing data during analysis	Partly	Studies with missing data excluded; converted values flagged for sensitivity analysis (p. 7, 10)
11.2	Justify the decisions made to deal with missing data	Partly	Conversions noted as approximate and checked in sensitivity analyses (p. 10)
Meta-analytic model description			
12.1	Describe the models used for synthesis of effect sizes	Yes	Multilevel mixed-effects models via `rma.mv` (p. 12)
12.2	If not a random-effects model, justify the choice	Partly	Multilevel random-effects model used; rationale conveyed implicitly rather than as an explicit justification (p. 12)
Software			
13.1	Describe the statistical platform used for inference	Yes	R (p. 12)
13.2	Describe the packages used to run models	Yes	`metafor`, `orchaRd` (p. 12)
13.3	Describe the functions used to run models	Yes	`rma.mv` (p. 12)
13.4	Describe any arguments that differed from default settings	No	Non-default arguments not described in the main text
13.5	Describe the version numbers of all software used	Yes	R v4.6.0, `metafor` v5.0-1 (p. 12)
Non-independence			

continued on next page

Table S1 continued from previous page

Item	Sub-item	Reported?	Location / notes
14.1	Describe the types of non-independence encountered	Yes	Multiple effect sizes per study, species, and phylogenetic non-independence (p. 12)
14.2	Describe how non-independence has been handled	Yes	Random effects for study/species/observation ID and a phylogenetic correlation matrix (p. 12)
14.3	Justify decisions made	Yes	Model structure described as accounting for each dependency source (p. 12)
Meta-regression and model selection			
15.1	Provide a rationale for the inclusion of moderators evaluated in meta-regressions	Yes	Taxonomic group plus four candidate biological moderators (p. 12–13)
15.2	Justify the number of parameters estimated relative to effect sizes and studies	Partly	Uni-moderator analyses used; explicit parameter-count justification not detailed (p. 13)
15.3	Describe any process of model selection	Partly	Series of uni-moderator models; no formal model-selection procedure (p. 13)
Publication bias and sensitivity analyses			
16.1	Describe assessments of risk of bias due to missing results	Yes	Funnel-plot asymmetry, multilevel Egger's regression, time-lag test (p. 13)
16.2	Describe steps taken to investigate the effects of such biases	Yes	Bias-corrected intercept and bias-robust estimator (p. 14–15, 19)
16.3	Describe any other analyses of robustness of the results	Yes	Leave-one-out (study and species), contentious effect-size handling (p. 15, 19, 21)
Clarification of post hoc analyses			
17.1	Acknowledge hypotheses formulated after data analysis	Yes	Four moderators explicitly labelled exploratory/candidate (p. 13)
Metadata, data, and code			

continued on next page

Table S1 continued from previous page

Item	Sub-item	Reported?	Location / notes
18.1	Share metadata (data descriptions)	Yes	GitHub repository and supplementary website (p. 16, 33)
18.2	Share data required to reproduce the results presented	Yes	Data available on GitHub (p. 16, 33)
18.3	Share additional data not presented in the manuscript	Yes	Supplementary materials / website (p. 16, 33)
18.4	Share analysis scripts	Yes	Annotated R code on GitHub / website (p. 16, 33)
Results of study selection process			
19.1	Report the number of studies screened	Yes	PRISMA flowchart, Figure 1 (p. 25)
19.2	Report the number of studies excluded at each stage of screening	Yes	PRISMA flowchart, Figure 1 (p. 25)
19.3	Report brief reasons for exclusion from the full-text stage	Yes	Reasons listed (n = 19 excluded) (p. 25)
19.4	Present a PRISMA-like flowchart	Yes	Figure 1 (p. 25)
Sample sizes and study characteristics			
20.1	Report the number of studies and effect sizes for data included in meta-analyses	Yes	69 effect sizes from 29 new studies, combined with source datasets (p. 17)
20.2	Report the number of studies and effect sizes for subsets included in meta-regressions	Partly	Overall counts reported; per-subset counts for each moderator not fully reported (p. 17–19)
20.3	Provide a summary of key characteristics for reported outcomes	Yes	Dataset characteristics and Table S2 (p. 17, SI)
20.4	Provide a summary of limitations of included moderators (e.g. collinearity/overlap)	No	Moderator collinearity/overlap not summarised
20.5	Provide a summary of characteristics related to individual study quality (risk of bias)	No	Not applicable; study-level quality not assessed
Meta-analysis			
21.1	Provide a quantitative synthesis with mean effect size and confidence/credible intervals	Yes	Overall OR = 1.81 (95% CI 1.34–2.45) and Hedges' g (p. 2, 17)
Heterogeneity			

continued on next page

Table S1 continued from previous page

Item	Sub-item	Reported?	Location / notes
22.1	Report indicators of heterogeneity (e.g. I^2, τ^2, variance components)	Yes	Total $I^2 > 74\%$ and partial I^2 by random factor (p. 18, 32)
Meta-regression			
23.1	Provide estimates of meta-regression slopes and confidence/credible intervals	Yes	Uni-moderator estimates reported (p. 18–19)
23.2	Include estimates and intervals for all moderator variables assessed (complete reporting)	Yes	Taxonomic and exploratory moderators reported (p. 18–19)
23.3	Report interactions, if included	NA	No interaction terms included (uni-moderator analyses only)
23.4	Describe outcomes from model selection, if done	NA	No formal model selection performed
Outcomes of publication bias and sensitivity analyses			
24.1	Provide results for the assessments of the risks of bias	Yes	Funnel plots and Egger's-type regression results (p. 19, 30–31)
24.2	Provide results for the robustness of the review's results	Yes	Leave-one-out and contentious-effect analyses confirm robustness (p. 19)
Discussion			
25.1	Summarise the main findings in terms of the magnitude of effect	Yes	Discussion (p. 19–20)
25.2	Summarise the main findings in terms of the precision of effects	Yes	Discussion of intervals/significance (p. 19–20)
25.3	Summarise the main findings in terms of their heterogeneity	Yes	Substantial, largely unexplained heterogeneity (p. 20)
25.4	Summarise the main findings in terms of their biological/practical relevance	Yes	Implications for sexual selection discussed (p. 20–22)
25.5	Compare results with previous reviews on the topic	Yes	Direct comparison with the two source meta-analyses (p. 19–20)
25.6	Consider limitations and their influence on the generality of conclusions	Partly	Evidence limitations discussed (p. 22–24); limitations of the review process itself not explicitly discussed
Contributions and funding			

continued on next page

Table S1 continued from previous page

Item	Sub-item	Reported?	Location / notes
26.1	Provide names, affiliations, and funding sources of all co-authors	Yes	CERC and NSERC grants (RGPIN-2026-05996) (p. 33)
26.2	List the contributions of each co-author	Yes	CRediT-style author contributions statement (p. 33–34)
26.3	Provide contact details for the corresponding author	Yes	Corresponding-author contact located in the title page (p. 1)
26.4	Disclose any conflicts of interest	Yes	"No competing interests declared" (p. 33)
References			
27.1	Provide a reference list of all studies included in the review	Yes	Included studies cited [42–70] (p. 17, references)
27.2	List included studies as referenced sources	Yes	Included studies appear in the reference list rather than only in a table (p. 17, 38–41)

Table S2: Studies included in the updated meta-analysis.

Author	Year	Title	Journal / Institution	DOI / Link
Araújo et al.	2024	A Social Learning Primacy Trend in Mate-Copying: An Experiment in *Drosophila melanogaster*.	Royal Society Open Science	10.1098/rsos.240408
Belkina et al.	2021	Mate Choice Copying in *Drosophila* Is Probably Less Robust than Previously Suggested.	Animal Behaviour	10.1016/j.anbehav.2021.04.007
Bierbach et al.	2011	Sperm Competition Risk Affects Male Mate Choice Copying.	Behavioral Ecology and Sociobiology	10.1007/s00265-011-1177-3
Cirino et al.	2023	Robust Mate Preferences despite Means and Opportunity for Mate Choice Copying in an Insect.	Animal Behaviour	10.1016/j.anbehav.2023.03.018
Clemente	2012	Cultural Transmission of Sexual Preferences in *Drosophila melanogaster*: Testing for the Durability of Socially Learnt Preferences.	Universidade de Lisboa	ProQuest
Duffy et al.	2025	Mate Choice Copying Behaviour in the Livebearing Fish *Brachyrhaphis*.	Animal Behaviour	10.1016/j.anbehav.2025.123148
Harrison-Weiss	2024	Female Choice and Social Decision-Making in a Marine Toadfish (*Porichthys notatus*).	Thesis	hdl
Kavaliers et al.	2022	Odor-Based Mate Choice Copying in Deer Mice Is Not Affected by Familiarity or Kinship.	Animal Cognition	10.1007/s10071-021-01550-z
Kavaliers et al.	2019	Conspecific Infection Threat Rapidly Biases the Social Responses of Female Mice: Involvement of Oxytocin.	Hormones and Behavior	10.1016/j.yhbeh.2019.04.016
Kniel et al.	2017	The Role of the Model in Mate-Choice Copying in Female Zebra Finches.	Ethology	10.1111/eth.12611
Kniel et al.	2015	Quality of Public Information Matters in Mate-Choice Copying in Female Zebra Finches.	Frontiers in Zoology	10.1186/s12983-015-0119-8
Kobayashi and Karino	2026	Number Matters: Effect of the Number of Models on Mate-Choice Choice Copying in Female Guppies.	Journal of Ethology	10.1007/s10164-025-00876-2

Continued on next page

Table S2 – continued from previous page

Author	Year	Title	Journal / Institution	DOI / Link
MacLaren	2024	Aesthetic Preference for Artificially Selected Color Variant Affects Mate Choice Copying Behavior in Female *Poecilia latipinna*.	PLOS One	10.1371/journal.pone.0298171
MacLaren	2019	Evidence of an Emerging Female Preference for an Artificial Male Trait and the Potential for Spread via Mate Choice Copying in *Poecilia latipinna*.	Ethology	10.1111/eth.12885
Marcos	2019	Regulação Do Mate-Choice Copying Por Neurónios Apetitivos Ou Neurónios Aversivos de Dopamina Em *Drosophila melanogaster*.	ISPA	ProQuest
Miller	2013	Testing for a Link between Personality and Mate-Choice Copying Tendency in the Trinidadian Guppy (*Poecilia reticulata*).	Carleton Univesrity	Carleton
Mitchell et al.	2025	Chemical Mate Choice Copying in *Drosophila melanogaster*.	American Naturalist	10.1086/736329
Moran et al.	2013	Mate Choice Copying in Two Species of Darters (Percidae: *Etheostoma*).	Behaviour	10.1163/1568539X-00003092
Nöbel et al.	2025	Long-Term Social Memory of Mate Copying in *Drosophila melanogaster* Is Localized in Mushroom Bodies.	Scientific Reports	10.1038/s41598-025-88535-x
Nöbel et al.	2023	Mate Copying Requires the Coincidence Detector Rutabaga in the Mushroom Bodies of *Drosophila melanogaster*.	iScience	10.1016/j.isci.2023.107682
Nöbel and Kaufmann	2025	Mate Copying in *Drosophila simulans*.	Biology Letters	10.1098/rsbl.2025.0070
Nöbel et al.	2022	Female Fruit Flies Copy the Acceptance, but Not the Rejection, of a Mate.	Behavioral Ecology	10.1093/beheco/arac071
Nöbel et al.	2022	2-D Sex Images Elicit Mate Copying in Fruit Flies.	Scientific Reports	10.1038/s41598-022-26252-5
Nöbel et al.	2023	No Evidence for Mate Copying in *Danio rerio*.	Behavioural Processes	10.1016/j.beproc.2023.104837
Ophir	2009	Mate Assessment and Non-Independent Mate Choice by Female Japanese Quail.	McMaster University	hdl

Continued on next page

Table S2 – continued from previous page

Author	Year	Title	Journal / Institution	DOI / Link
Pusiak et al.	2020	Relative Sexual Attractiveness Does Not Influence Mate-Choice Copying in Male Trinidadian Guppies, *Poecilia reticulata*.	Animal Behaviour	10.1016/j.anbehav.2020.02.008
Silva	2014	Can Social Learning Promote Hybridisation? Mate-Choice Copying in *Drosophila subobscura* Populations.	Universidade de Lisboa	hdl
Witte et al.	2018	Test of the Deception Hypothesis in Atlantic Mollies *Poecilia mexicana*—Does the Audience Copy a Pretended Mate Choice of Others?	Biology	10.3390/biology7030040
Ziege et al.	2012	A Comparison of Two Methods to Assess Audience-Induced Changes in Male Mate Choice.	Current Zoology	10.1093/czoolo/58.1.84

Table S3: Studies excluded from the meta-analysis after full-text assessment, with reasons for exclusion.

Author	Year	Title	Journal	DOI / Source	Reason for exclusion
Bridges et al.	2019	Animal Behaviour: Conformity and the Beginnings of Culture in an Insect.	Current Biology	10.1016/j.cub.2019.01.023	Wrong study type
Chen et al.	2021	Love the Love of Other Men: Male's Mate-Choice Copying under the Activation of Mating Motives; 君子不夺人所好？择偶动机启动下的男性择偶复制	Journal of Psychological Science	10.16719/j.cnki.1671-6981.20210424	Study on humans
Cruz	2014	Communication Networks in the Zebrafish, *Danio rerio*.	Dissertation Universidade de Lisboa	https://search.proquest.com/openview/aaeafb1c91f2d0dfbd0dbd892c8ce4fe/1?pq-origsite=gscholar&cbl=2026366&diss=y	Wrong study design
Durey et al.	2008	Do Female Siamese Fighting Fish Copy the Mate Choice of Others?	European Conference on Behavioural Biology	https://researchprofiles.ku.dk/da/publications/b2bc1be0-ddbe-11dd-b5fc-000ea68e967b	No full-text available
Gilman et al.	2020	Demonstrating Mate Choice Copying in Spiders Requires Further Research	Current Zoology	10.1093/cz/zoz033	Wrong study type
Harrison-Weiss et al.	2025	Assessing the Cues Required for Mate Choice Copying in the Plainfin Midshipman Fish.	Animal Behaviour	10.1016/j.anbehav.2025.123292	Same data as in Harrison-Weiss 2024 Dissertation
Huckvale	2016	An Investigation into the Influence of Hunger-State on the Mate-Choice Copying Phenomenon in Trinidadian Guppies (*Poecilia reticulata*).	Thesis	figshare	Does not contain data to calculate effect sizes

Continued on next page

Table S3 – continued from previous page

Author	Year	Title	Journal	DOI / Source	Reason for exclusion
Imperio et al.	2020	Female Mating Tactics in Lekking Fallow Deer	Scientific Reports	10.1038/s41598-020-58681-5	Wrong study design
李卫卫	2017	择偶复制的影响因素及心理机制	心理技术与应用	10.16842/j.cnki.issn2095-5588.2017.03.008	Study on humans
Mendelson et al.	2023	Could Sexual Selection Be Driven by the Mistaken Inferences of Young Females?	PLOS Biology	10.1371/journal.pbio.3002321	Wrong study type
Nöbel et al.	2022	The Importance of Population Heterogeneities in Detecting Social Learning as the Foundation of Animal Cultural Transmission.	Proceedings of the Royal Society B: Biological Sciences	10.1098/rspb.2022.0431	Does not contain data to calculate effect sizes
Plath et al.	2008	Male Fish Deceive Competitors about Mating Preferences.	Current Biology	10.1016/j.cub.2008.06.067	Wrong study design
Plath et al.	2009	Audience Effect Alters Male but Not Female Mating Preferences.	Behavioral Ecology and Sociobiology	10.1007/s00265-008-0672-7	Wrong study design
Roberts and Gosling	2004	Manipulation of Olfactory Signaling and Mate Choice for Conservation Breeding: A Case Study of Harvest Mice	Conservation Biology	10.1111/j.1523-1739.2004.00514.x	Wrong study design
Sapage and Varela	2020	Two Research Avenues for Future Mate-Choice Copying Studies: A Comment on Davies et al.	Behavioral Ecology	10.1093/beheco/araa075	Wrong study design
Schütz and Taborsky	2000	Giant Males or Dwarf Females: What Determines the Extreme Sexual Size Dimorphism in *Lamprologus callipterus*?	Journal of Fish Biology	10.1111/j.1095-8649.2000.tb00485.x	Wrong study design
Vega-Trejo et al.	2024	Fear of Mating out	Scientific Reports	10.1038/s41598-024-83465-6	Wrong study design

Continued on next page

Table S3 – continued from previous page

Author	Year	Title	Journal	DOI / Source	Reason for exclusion
Yang et al.	2025	Evolutionary Divergence of Developmental Plasticity and Learning of Mating Tactics in Trinidadian Guppies.	Journal of Animal Ecology	10.1111/1365-2656.14043	Wrong study design
Ziege et al.	2009	Audience Effects in the Atlantic Molly (*Poecilia mexicana*)–Prudent Male Mate Choice in Response to Perceived Sperm Competition Risk?	Frontiers in Zoology	10.1186/1742-9994-6-17	Wrong study design

Supplementary references

1. Nakagawa, S. *et al.* Methods for Testing Publication Bias in Ecological and Evolutionary Meta-Analyses. *Methods in Ecology and Evolution* **13,** 4–21. doi:10.1111/2041-210X.13724 (2022).

2. Stanley, T. D. & Doucouliagos, H. *Meta-Regression Analysis in Economics and Business* 176 pp. doi:10.4324/9780203111710 (Routledge, London, 2012).

3. Stanley, T. D. & Doucouliagos, H. Meta-Regression Approximations to Reduce Publication Selection Bias. *Research Synthesis Methods* **5,** 61–78. doi:10.1002/jrsm.1095 (2014).

4. Yang, Y. *et al.* Robust Point and Variance Estimation for Meta-Analyses with Selective Reporting and Dependent Effect Sizes. *Methods in Ecology and Evolution* **15,** 1593–1610. doi:10.1111/2041-210x.14377 (2024).

5. Williams, C., Yang, Y., Warton, D. I. & Nakagawa, S. Modelling Approaches for Meta-Analyses with Dependent Effect Sizes in Ecology and Evolution: A Simulation Study. *Methods in Ecology and Evolution* **16,** 2362–2379. doi:10.1111/2041-210x.70156 (2025).

6. Nakagawa, S. *et al.* orchaRd 2.0: An R Package for Visualising Meta-Analyses with Orchard Plots. *Methods in Ecology and Evolution* **14,** 2003–2010. doi:10.1111/2041-210x.14152 (2023).

7. Davies, A. D., Lewis, Z. & Dougherty, L. R. A Meta-Analysis of Factors Influencing the Strength of Mate-Choice Copying in Animals. *Behavioral Ecology* **31,** 1279–1290. doi:10.1093/beheco/araa064 (2020).

8. Jones, B. C. & DuVal, E. H. Mechanisms of Social Influence: A Meta-Analysis of the Effects of Social Information on Female Mate Choice Decisions. *Frontiers in Ecology and Evolution* **7,** 390. doi:10.3389/fevo.2019.00390 (2019).

9. O'Dea, R. E. *et al.* Preferred Reporting Items for Systematic Reviews and Meta-Analyses in Ecology and Evolutionary Biology: A PRISMA Extension. *Biological Reviews* **96,** 1695–1722. doi:10.1111/brv.12721 (2021).