# Do egg hormones have fitness consequences in wild birds? A systematic review and meta-analysis

Lucia Mentesana[1,2,3*], Michaela Hau[1,4], Pietro B. D'Amelio[5], Nicolas M. Adreani[2,6], Alfredo Sánchez-Tójar[3]

[1] Max Planck Institute for Biological Intelligence, Germany.
[2] Biology Department, Facultad de Ciencias, Universidad de la República, Uruguay.
[3] Department of Evolutionary Biology, Faculty of Biology, Bielefeld University, Germany.
[4] Department of Biology, University of Konstanz, Germany.
[5] Department of Biology, Reed College, 3203 SE Woodstock Blvd, Portland, OR, 97202, USA.
[6] Department of Collective Behavior, Max Planck Institute of Animal Behaviour, Konstanz, Germany.

Dr. Lucia Mentesana: luciamentesana@gmail.com (https://orcid.org/0000-0003-4090-2570)
Dr. Michaela Hau: michaela.hau@bi.mpg.de (https://orcid.org/0000-0002-3836-4083)
Dr. Pietro B. D'Amelio: pie.damelio@googlemail.com (https://orcid.org/0000-0002-4095-6088)
Dr. Nicolas M. Adreani: mn.adreani@gmail.com (https://orcid.org/0000-0002-5043-0389)
Dr. Alfredo Sánchez-Tójar: alfredo.tojar@gmail.com (https://orcid.org/0000-0002-2886-0649)

* Author for correspondence: Dr. Lucia Mentesana
Address: Igua 4225, Montevideo (11400), Uruguay. Telephone: (+598) 2525-8617.

## Abstract

Egg-laying species are key models for understanding the adaptive significance of maternal effects, with egg hormones proposed as an important underlying mechanism. However, even thirty years after their discovery, the evolutionary consequences of hormone-mediated maternal effects remain unclear. Using evidence synthesis, we tested the extent to which increased prenatal maternal hormone deposition in eggs relates to fitness in wild birds (19 species, 438 effect sizes, 57 studies). Egg androgens, glucocorticoids, and thyroid hormones showed an overall near-zero mean effect for both maternal and offspring fitness proxies. However, heterogeneity was high, suggesting that egg hormone effects on fitness are context-dependent. Hormone type and age did not explain much of the observed variance, nor did methodological factors such as type of study or experimental design. Heterogeneity decomposition showed that differences in effect sizes were mostly driven by within-study variability and phylogenetic relationships. Our study provides the most comprehensive investigation to date of the relationship between egg hormones and fitness in vertebrates. By synthesizing current knowledge, we aim to overcome theoretical shortcomings in the field of maternal effects via egg hormone deposition and inspire new research into its many intriguing aspects.

## 1. Introduction

Maternal (genetic and nongenetic) effects are widespread across nature, occurring in plants, invertebrates, and vertebrates. They occur when mothers shape their offspring's phenotype by influencing the environment they provide (Mousseau & Fox 1998; Reinhold 2002; Uller 2008, 2012; Wolf & Wade 2016). The general contribution of maternal effects to offspring phenotypic variation within populations was found to be half that (~ 11%) of additive genetic effects (~ 21%), influencing offspring morphology, physiology, and phenology, among other traits (meta-analysis by Moore *et al.* 2019; but see also Sánchez-Tójar *et al.* 2020; Yin *et al.* 2019). Mothers can exert these effects both before and after birth, but prenatal maternal effects are generally expected to have larger and more irreversible impacts on offspring than postnatal effects, with consequences for offspring survival and reproductive success (Groothuis *et al.* 2005a; Marshall & Uller 2007; Mousseau & Fox 1998; Uller 2008; Wolf *et al.* 2011).

Egg-laying species are key models for understanding the adaptive significance of maternal effects and their underlying mechanisms. Unlike in most mammals, embryos develop outside the mother's body, limiting the time for direct maternal influence on the offspring. In egg-laying species, mothers can influence the offspring's prenatal environment by depositing compounds into the egg (e.g., Lovern & Wade 2001; McCormick 1999; Mentesana *et al.* 2019; Valcu *et al.* 2019; Weiss *et al.* 2011). In 1993, Schwabl first reported maternal hormones in bird eggs. These findings then were extended to other egg-laying vertebrates like fishes, reptiles, and amphibians (e.g., Groothuis *et al.* 2005a). Since then, many studies have suggested that hormones, especially steroid hormones present in the yolk, shape offspring growth, behavior (e.g., begging, aggression, boldness, activity, food competitiveness), and physiological systems (e.g., immune system, hypothalamic-pituitary-adrenal axis, oxidative condition; von Engelhardt & Groothuis 2011; Gil 2008; Groothuis *et al.* 2005a; Lovern & Wade 2001; Mentesana *et al.* 2021a). Yet, whether such phenotypic changes translate into fitness benefits for the mother, the father, or the offspring remains under debate (reviewed by Bebbington & Groothuis 2021), especially for wild populations (e.g., Tschirren *et al.* 2014).

Egg hormone-mediated maternal effects are often considered an adaptive mechanism, where mothers are assumed to adjust hormone concentrations in response to environmental cues to shape offspring phenotypes and optimize maternal and/or offspring fitness in variable but predictable environments (reviewed by Groothuis *et al.* 2005a). However, maternal egg hormone concentrations may also lack adaptive value and instead reflect unavoidable physiological processes, such as changes in maternal hormone concentrations that passively diffuse into the egg during formation, leading to varying hormone allocation (reviewed by Gil 2008; Groothuis &

Schwabl 2008). Thirty years after the discovery of maternal hormones in eggs, the evolutionary consequences of hormone-mediated maternal effects remain unclear.

There are several reasons why the evolutionary consequences of maternal egg hormones remain unresolved. First, the influence of maternal hormones on fitness is expected to vary greatly, depending not only on the group of hormones studied but also on the fitness proxy used, which can result in contrasting fitness effects. For example, egg androgens (i.e., androstenedione, testosterone, and 5-α dihydrotestosterone) and thyroid hormones (i.e., TH3 and TH4) can lead to increased behavioral competitiveness, boost early growth, and ultimately survival (von Engelhardt & Groothuis 2011; Groothuis *et al.* 2005a; Hsu *et al.* 2017; Ruuskanen *et al.* 2018). However, high concentrations of egg androgens can also impair the offspring's immune system (reviewed by Gil 2008) and oxidative condition (e.g., Treidel *et al.* 2013), potentially decreasing survival. Regarding egg glucocorticoids, they have been reported to alter behaviors such as antipredator response, impair physiological systems beneficial for offspring (e.g., immune system), and decrease offspring body condition, growth rate, and ultimately survival (Haussmann *et al.* 2012; Polich *et al.* 2018; Stier *et al.* 2009). Second, egg hormones might have sex-specific effects on offspring fitness. A meta-analysis of bird studies investigating how hormonal manipulations of eggs affect offspring traits showed that experimentally increasing androgen and corticosterone concentrations via egg injection tended to lead to larger effect sizes in male compared to female offspring (Podmokła *et al.* 2018). However, the generality of this finding is debated, as the study could only examine the relationship between experimentally injected eggs and offspring fitness proxies in seven populations of wild birds - six of which provided data only on androgens, and one on corticosterone. Third, the influence of maternal hormones on fitness might be dependent on ontogenetic state. Maternal effects are expected and have been observed to weaken throughout offspring ontogeny; however, a meta-analysis examining the contribution of additive genetic phenotypic variation determined by maternal effects, found only partial support for this pattern (Moore *et al.* 2019). Similarly, another meta-analysis that considered parental effects more broadly reported inconclusive evidence for such a weakening effect (Sánchez-Tójar *et al.* 2020). Fourth, theory predicts that the evolution of maternal effects generally, and of hormone allocation specifically, should be shaped by conflicting fitness optima between offspring, mothers, and fathers (Bebbington & Groothuis 2021; Marshall & Uller 2007). For example, high concentrations of egg androgens might increase maternal reproductive success and offspring survival by boosting offspring growth. But at the same time, high concentrations of egg androgens might decrease maternal life expectancy by increasing her workload or even reduce reproductive success if the increased workload cannot be fulfilled through parental care. Thus, if conflicting

fitness optima between offspring and mothers occur and are not considered, we expect considerable variation in the net fitness consequences of egg hormone concentrations. Finally, besides biological reasons, methodological aspects might also add uncertainty to our understanding of the evolutionary importance of maternal egg hormones (reviewed by Groothuis *et al.* 2020; Groothuis & Von Engelhardt 2005). A meta-analysis examining the effects of different experimental manipulations of egg androgens and glucocorticoids in wild and captive bird populations provided evidence that factors such as study type (correlational *vs* experimental), manipulation methods (e.g., maternal *vs* egg manipulation), and hormone doses (e.g., within natural range *vs* supra-physiological concentrations) can influence the relationship between egg hormones and fitness proxies (Podmokła *et al.* 2018).

Here, we used an evidence synthesis approach to test the extent to which an increase in prenatal maternal egg hormone concentrations relates to fitness in wild populations of egg-laying vertebrates and to understand the patterns of the effects (i.e., the sources of heterogeneity). We tested six pre-registered predictions, which we divided into biological and methodological categories (pre-registration: Mentesana *et al.* 2021b). By pre-registering our study, we enhanced the distinction between hypothesis testing and hypothesis generation, thereby improving the credibility of our research findings (Ihle *et al.* 2017). Note that we mostly focused on the relationship between egg hormones and offspring fitness rather than parental fitness because the former has received the most empirical attention. Indeed, due to a lack of studies, we could not investigate the relationship between egg hormones and paternal fitness and could only conduct an exploratory analysis for egg hormones and maternal fitness.

1. Biological predictions:

BP.1. Before testing our pre-registered predictions for each group of hormones, we first calculated the overall relationship between egg hormones and offspring fitness proxies, regardless of hormone group, and estimated heterogeneity across effect sizes. Next, for each hormone group, we based our predictions on the prevailing hypothesis in the literature. We predicted a positive association between egg androgens (BP.1.1) and thyroid hormones (BP.1.3) with offspring fitness proxies, and a negative one with egg glucocorticoids (BP.1.2). We did not have directional predictions for progesterone and estradiol (BP.1.4 & BP.1.5) due to insufficient data (reviewed by Gil 2008; Groothuis *et al.* 2019). For the remaining predictions, we assumed the same hormone-specific directional predictions as detailed in BP.1.

BP.2. Given some early evidence that androgens and glucocorticoids may have a stronger impact on males than females (Podmokła *et al.* 2018), we predicted larger effect

sizes in males. For the remaining hormones (thyroids, progesterone, and estradiol), we did not have directional sex-specific predictions. Unfortunately, the number of studies that have addressed whether egg hormones have sex-specific fitness consequences for the offspring in wild populations was too limited to provide reliable conclusions (Supplementary Information S1).

BP.3. Since there is partial evidence that maternal effects weaken throughout ontogeny (Moore *et al.* 2019), we predicted the relationship between egg hormones and offspring fitness proxies to decline from early to late life stages in both altricial and precocial species, weakening from hatching through the nesting period (excluding hatching day) and into adulthood (i.e., before and after offspring independence; BP.3.1). Also, for altricial species, we predicted that the association would become weaker throughout the nesting phase (BP.3.2).

2. Methodological predictions:

MP.1. We predicted stronger associations in experimental studies than in correlational ones because manipulations should reduce/eliminate confounding factors and/or because physiological constraints or trade-offs might prevent females from depositing optimal hormone concentrations into the egg. An alternative (non-preregistered) prediction would be that manipulating hormone concentrations might incur fitness costs by shifting concentrations away from adaptive values (e.g., von Engelhardt & Groothuis 2011).

MP.2. Since females can potentially modify the transfer of circulating hormones to their eggs (reviewed by Groothuis & Schwabl 2008), we predicted that the association between egg hormones and fitness proxies would be stronger for direct egg manipulations than for female manipulations.

MP.3. The dose of hormones used may differentially affect offspring fitness. While high doses of hormones within the natural range are expected to result in stronger fitness effects, doses beyond naturally-occurring concentrations may lead to weaker or even opposite effects on fitness (e.g., due to toxicity; Norton & Wira 1977). Following Podmokła *et al.* (2018), we predicted that manipulations within the natural range would lead to stronger fitness effects than supra-physiological concentrations for androgens, but not glucocorticoids, whereas we did not have directional predictions for the remaining hormones (thyroids, progesterone, and estradiol).

## 2. Materials and Methods

## 2.1. Protocol

The study protocol was pre-registered on the Open Science Framework before data extraction began (Mentesana *et al.* 2021b; https://doi.org/10.17605/OSF.IO/KU47W). Our pre-registration specified the purpose of the study, our a priori hypotheses and predictions, the data collection procedures, and the entire analytical plan. Unless stated otherwise, we adhered to these pre-registered plans. Throughout, we follow the Preferred Reporting Items for Systematic reviews and Meta-Analyses in ecology and evolutionary biology (PRISMA-EcoEvo Checklist: Supplementary Figure 1 & Supplementary Information S2; O'Dea *et al.* 2021).

## 2.2. Information sources and search

We conducted a systematic search of the literature from 1993 (i.e., the year when the first paper on maternal hormones was published) until the 22nd of August 2022 in both Web of Science Core Collection and Scopus. The search query was designed to find studies that: (1) tested inter- and/or transgenerational (plasticity) effects; (2) measured or experimentally manipulated egg hormone concentrations; and (3) measured maternal and/or offspring fitness proxies. The full keyword search string with further details is provided in Supplementary Information S3.

## 2.3. Study selection and eligibility criteria

Our searches yielded 2404 records which we deduplicated using the R package revtool v.0.3.0 (Westgate 2018). We screened the titles and abstracts of all records using Rayyan (Ouzzani *et al.* 2016). In total, 178 articles passed the title-and-abstract screening and were subjected to full-text screening (for the decision tree used, see Supplementary Figure 2). We included studies conducted in free-living populations of vertebrates that either experimentally manipulated maternal hormone concentrations during egg laying, affecting hormone deposition into eggs, or measured/manipulated egg hormone concentrations during egg laying, provided they also reported fitness proxies. For both mother and offspring, we considered all fitness components related to their reproduction and/or survival (full list of fitness proxies used in Supplementary Information 4). For the offspring, we also considered morphological information because they are commonly accepted as predictors of post-fledging survival (e.g., meta-analysis by Ronget *et al.* 2018).

After full-text screening, we identified 79 articles meeting our criteria (see PRISMA diagram in Supplementary Figure 3). Three observers (LM, NMA and PBD) evenly performed all the screening and double-screened ~20% of the records to confirm the reproducibility of the

procedure (12% for title-and-abstract screening and 30% for full-text screening). The three observers collectively discussed and resolved conflicting decisions (~7% conflicted papers).

Although our pre-registered plan aimed to test for the relationship between egg hormones and fitness proxies in all egg-laying vertebrate groups, only 2 out of the 79 potentially suitable studies found were conducted on egg-laying vertebrates other than birds; thus, we decided to focus our study solely on birds.

## 2.4. Data collection and extraction

The data extraction was evenly performed by three observers (LM, NMA and PBD). Two observers (LM and AST) double-checked 36% of the extracted data to confirm the reproducibility of this process, LM double-checked the remaining data in search of potential errors, and both observers discussed and fixed any typos found.

We extracted data from text, tables or figures. For the latter, we used the R package metaDigitise v.1.0.1 (Pick *et al.* 2019). In addition, when necessary, we also obtained raw data directly from supplementary materials and/or from published datasets. Complete data extraction was possible for 22 articles (29% of the dataset). To obtain the missing data from the remaining articles, we contacted 31 authors (54 articles) using a standardized email template (Supplementary Information S5), which allowed us to obtain complete data from an additional 20 articles (37% of the dataset). Eight authors communicated that the data were lost, one author expressed that they were working outside science and did not have time to search for the data, another could not provide the data due to being on sabbatical leave, and five authors did not reply. In total, we obtained complete data from 47 articles and partial data for 10 articles (N = 57); that is, we had to exclude 21 in-principle suitable articles (27%) due to preventable reporting issues (for more discussion, see Hennessy *et al.* 2022). During author correspondence and the review process, one author and one reviewer informed us of two studies that although originally fell outside our systematic search (Duckworth *et al.* 2015: k = 1) or did not fulfill our title-and-abstract screening inclusion/exclusion criteria (Remeš 2011: k = 2) contained data for our meta-analysis and were, therefore, subsequently added to our dataset. For a complete list of studies and extracted variables, see Table 1 and the provided data (https://github.com/ASanchez-Tojar/meta-analysis_egg_hormones_and_fitnes).

Table 1: List of studies (N = 57) from which data was extracted and included in our meta-analysis on egg hormone concentrations and maternal and offspring fitness in wild birds.

| Study | Year | Journal | Number of effect sizes | Sample size[a] | Species | References |
|---|---|---|---|---|---|---|
| 1 | 2000 | Proceedings of the Royal Society B: Biological Sciences | 21 | 16 (11,19) | *Falco sparverius* | Sockman & Schwabl 2000 |
| 2 | 2003 | Animal Behaviour | 1 | 534 | *Chroicocephalus ridibundus* | Eising & Groothuis 2003 |
| 3 | 2003 | Proceedings of the Royal Society B: Biological Sciences | 3 | 58 | *Larus fuscus* | Verboven et al. 2003 |
| 4 | 2004 | Hormones and Behavior | 1 | 50 | *Junco hyemalis* | Clotfelter *et al.* 2004 |
| 5 | 2004 | Hormones and Behavior | 5 | 49 (28,60) | *Sturnus vulgaris* | Pilz et al. 2004 |
| 6 | 2004 | Behavioral Ecology and Sociobiology | 2 | 20.5 (17,24) | *Passer domesticus* | Strasser & Schwabl 2004 |
| 7 | 2005 | American Naturalist | 2 | 73 | *Sturnus vulgaris* | Love *et al.* 2005 |
| 8 | 2005 | Proceedings of the Royal Society B: Biological Sciences | 1 | 210 | *Chroicocephalus ridibundus* | Müller *et al.* 2005 |
| 9 | 2005 | Physiological and Biochemical Zoology | 4 | 12.5 (10,18) | *Sialia sialis* | Navara *et al.* 2005 |
| 10 | 2005 | Hormones and Behavior | 17 | 202 (40,387) | *Larus michahellis* | Rubolini et al. 2005 |
| 11 | 2005 | Ecology Letters | 2 | 424.5 (376,473) | *Ficedula albicollis* | Tschirren *et al.* 2005 |
| 12 | 2005 | Biology Letters | 1 | 16 | *Chroicocephalus ridibundus* | Groothuis *et al.* 2005a |
| 13 | 2006 | Journal of Evolutionary Biology | 21 | 81.5 (28,163) | *Hirundo rustica* | Gil *et al.* 2006 |
| 14 | 2006 | Physiological and Biochemical Zoology | 2 | 29 (28,30) | *Haemorhous mexicanus* | Navara *et al.* 2006 |
| 15 | 2006 | Behavioral Ecology and Sociobiology | 5 | 123 (79,161) | *Larus michahellis* | Rubolini et al. 2006 |
| 16 | 2007 | Behavioral Ecology | 24 | 43 (43,49) | *Sturnus unicolor* | Müller et al. 2007 |
| 17 | 2007 | American Naturalist | 3 | 500 (488,623) | *Parus major* | Tschirren *et al.* 2007 |
| 18 | 2008 | Behavioral Ecology and Sociobiology | 9 | 9 (6,9) | *Falco sparverius* | Sockman *et al.* 2008 |
| 19 | 2009 | Behavioural Processes | 3 | 20 | *Sturnus unicolor* | López-Rull & Gil 2009 |
| 20 | 2009 | Journal of Animal Ecology | 1 | 49 | *Chroicocephalus ridibundus* | Müller *et al.* 2009 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 21 | 2009 | Journal of Avian Biology | 1 | 603 | *Ficedula albicollis* | Pitala *et al.* 2009 |
| 22 | 2009 | Hormones and Behavior | 1 | 66 | *Ficedula albicollis* | Ruuskanen *et al.* 2009 |
| 23 | 2010 | Journal of Experimental Biology | 2 | 30.5 (19,42) | *Parus major* | Galván & Alonso-Alvarez 2010 |
| 24 | 2010 | Hormones and Behavior | 2 | 57.5 (57,58) | *Ficedula albicollis* | Ruuskanen & Laaksonen 2010 |
| 25 | 2011 | Journal of Experimental Biology | 2 | 46.5 (46,47) | *Troglodytes aedon* | Clairardin et al. 2011 |
| 26 | 2011 | Behavioral Ecology | 16 | 94 (23,141) | *Ficedula albicollis* | Hegyi *et al.* 2011 |
| 27 | 2011 | Behavioral Ecology and Sociobiology | 2 | 128 | *Parus major* | Remeš 2011 |
| 28 | 2012 | PLOS One | 6 | 119 (110,400) | *Eudyptes chrysocome* | Poisbleau *et al.* 2012 |
| 29 | 2012 | Behavioral Ecology and Sociobiology | 4 | 124.5 (27,617) | *Ficedula albicollis* | Ruuskanen *et al.* 2012a |
| 30 | 2012 | Functional Ecology | 11 | 18 (10,606) | *Ficedula albicollis* | Ruuskanen *et al.* 2012b |
| 31 | 2013 | Hormones and Behaviour | 3 | 20 (8,20) | *Tachycineta bicolor* | Bentz *et al.* 2013 |
| 32 | 2013 | General and Comparative Endocrinology | 9 | 63.3 (50,185) | *Sturnus unicolor* | Muriel *et al.* 2013 |
| 33 | 2013 | Hormones and Behaviour | 1 | 91 | *Larus michahellis* | Noguera *et al.* 2013 |
| 34 | 2013 | PLOS One | 10 | 33 (26,33) | *Parus major* | Podlas *et al.* 2013 |
| 35 | 2013 | Hormones and Behavior | 4 | 54 (38,79) | *Ficedula albicollis* | Rice et al. 2013 |
| 36 | 2013 | Journal of Avian Biology | 2 | 75 (57,93) | *Ficedula hypoleuca* | Ruuskanen & Laaksonen 2013 |
| 37 | 2014 | Ecology Letters | 10 | 256 | *Ficedula albicollis* | Tschirren *et al.* 2014 |
| 38 | 2015 | American Naturalist | 2 | 133 | *Troglodytes aedon* | Bowers et al. 2015 |
| 39 | 2015 | Science | 1 | 18 | *Sialia mexicana* | Duckworth *et al.* 2015 |
| 40 | 2015 | Journal of Experimental Biology | 53 | 51.7 (14,131) | *Sturnus unicolor* | Muriel et al. 2015a |

| | | | | | | |
|---|---|---|---|---|---|---|
| 41 | 2015 | Journal of Evolutionary Biology | 16 | 175 (44,545) | *Sturnus unicolor* | Muriel *et al.* 2015b |
| 42 | 2015 | PLOS One | 5 | 49 (31,184) | *Parus major* | Tschirren 2015 |
| 43 | 2016 | Ecology and Evolution | 4 | 18.5 (17,19) | *Sialia sialis* | Bentz *et al.* 2016 |
| 44 | 2016 | Hormones and Behavior | 2 | 23 (22,24) | *Troglodytes aedon* | Bowers et al. 2016 |
| 45 | 2016 | Science of Nature | 4 | 88 (39,174) | *Ficedula albicollis* | Nagy *et al.* 2016 |
| 46 | 2016 | Hormones and Behavior | 5 | 41 (32,317) | *Larus michahellis* | Possenti *et al.* 2016 |
| 47 | 2016 | Hormones and Behavior | 18 | 66 (62,212) | *Parus major* | Ruuskanen et al. 2016 |
| 48 | 2016 | Journal of Experimental Zoology Part A: Ecological Genetics and Physiology | 10 | 28 (18,148) | *Troglodytes aedon* | Strange *et al.* 2016 |
| 49 | 2016 | Hormones and Behavior | 4 | 97 (40,187) | *Parus major* | Tilgar *et al.* 2016 |
| 50 | 2018 | PLOS One | 1 | 201 | *Larus michahellis* | Parolini *et al.* 2018 |
| 51 | 2018 | Functional Ecology | 15 | 233 (37,234) | *Troglodytes aedon* | Weber *et al.* 2018 |
| 52 | 2019 | Journal of Avian Biology | 16 | 126.5 (119,705) | *Ficedula albicollis* | Hsu *et al.* 2019 |
| 53 | 2020 | Scientific Reports | 5 | 113 (60,249) | *Ficedula hypoleuca* | Hsu *et al.* 2020 |
| 54 | 2020 | Molecular Ecology | 1 | 46 | *Larus michahellis* | Noguera et al. 2020 |
| 55 | 2020 | Ecology and Evolution | 12 | 13 (11,14) | *Turdus merula* | Partecke *et al.* 2020 |
| 56 | 2020 | Journal of Experimental Biology | 2 | 37 | *Passer domesticus* | Schwabl & Partecke 2020 |
| 57 | 2021 | Frontiers in Zoology | 48 | 315.5 (91,597) | *Parus major* | Mentesana *et al.* 2021a |

[a] For studies where we extracted more than one effect size and sample sizes (i.e., offspring or nest number) varied, we report the median sample size, with the minimum and maximum values in brackets.

## 2.5. Extracted variables

From each article we extracted information regarding the study (e.g., title, publication year, author information, journal), the study subjects (e.g., species, developmental mode, mean clutch size), the study area (e.g., location, latitude, longitude and elevation of study area), the methodology

(e.g., egg collection, sampling method, subject of the experiment, method and dose use), the fitness proxy studied (e.g., maternal/offspring, trait measured, offspring life history stage), the results (e.g., statistical test used, statistical value, mean values for control and experimental groups), and information on how data were provided (e.g., data location, data source). For the complete list of continuous and categorical moderators see Supplementary Information S6.

Note that for some fitness proxies, it is difficult to fully disentangle maternal and offspring fitness (e.g., hatching or fledging success; discussed by Wolf & Wade 2001). Following the recommendations of Wolf & Wade (2001), we took a conservative approach: fitness proxies measured in offspring that are largely or entirely controlled by the mother (e.g., hatching success, which is strongly influenced by maternal factors such as egg size; meta-analysis by Krist 2011) were assigned to the mother.

## 2.6. Effect size calculation

To quantify the relationship between egg hormones and fitness, we combined both Pearson's correlation coefficients *r* and biserial correlations $r_{bis}$ – throughout we refer to simply *r* for simplicity. First, for correlation coefficients other than Pearson's *r* (e.g., Spearman's rho correlation coefficient), percentages, absolute values or inferential statistics (e.g., F-value, t-value, Chi-square value), we calculated Pearson's *r* using the equations from Nakagawa & Cuthill (2007) and Lajeunesse (2013; Supplementary Information S7). For tests that lacked information on directionality, we determined it from the text, figures, or by directly asking the authors via email. We then calculated the variance in Pearson's *r* (*Vr*) as $Vr = (1 - r^2)^2 \div (N - 1)$ (Borenstein *et al.* 2009). Second, for experimental studies in which two groups were compared (control *vs* treatment), we transformed the means, SDs, and sample sizes of the fitness proxies of interest of each group into biserial correlations coefficients and calculated their sampling variance following Jacobs & Viechtbauer (2017) with the function 'escalc' from the R package "metafor" v.2.4-0 (Viechtbauer 2010). Biserial correlations are conceptually equivalent and directly comparable to Pearson's *r* (Jacobs & Viechtbauer 2017). Meta-analyses involving both Pearson's and biserial correlation coefficients need to be based on the raw coefficients; this is why we did not use Fisher's *r*-to-*z* transformation (Jacobs & Viechtbauer 2017). When more than two groups were compared (e.g., control *vs* treatment 1 *vs* treatment 2), we calculated effect sizes between the control and each experimental group and divided the sample size of the control group by the number of times the control group was used (N/2 in the previous example) to account for shared-control non-independence. Whenever more than one effect size could be extracted or calculated for the same data (e.g., when the same result was provided using more than one statistical

model), we chose one source using the following order of preference: (a) Pearson's *r*; (b) other correlation coefficients, (c) mean/percentages/absolute differences between control and experimental groups, and (d) inferential statistics. The sample size assigned to each effect size (i.e., number of clutches or number of offspring) reflected the replication level for the specific trait and experimental design reported by the authors in the papers. If the sample size was not reported at that level or the reported sample size did not correspond to the correct replication level, we used the sample size for the other replication level (if available) or estimated it from the reported degrees of freedom (df + 1) whenever possible (18 effect sizes).

In all cases, effect sizes were coded so that fitness proxies were positively correlated with overall fitness. Specifically, for 'egg mortality,' 'hatching failure,' and 'offspring mortality,' the sign of the effect sizes (k = 5 effect sizes, N = 3 studies) was inverted to ensure comparability across all fitness proxies.

## 2.7. Phylogenies and variance-covariance matrices

We built phylogenetic trees by searching for species in the Open Tree Taxonomy (Rees & Cranston 2017) and extracting their phylogenetic information from the Open Tree of Life (Hinchliff *et al.* 2015) using the R package "rotl" v.3.0.5 (Michonneau *et al.* 2016). We calculated tree branch length following Grafen (1989) with the R package "ape" v.5.4-1 (Paradis & Schliep 2019) and constructed a phylogenetic correlation matrix that was included in all the phylogenetic multilevel models (Figure 1). In addition, we specified sampling variances as a variance-covariance matrix rather than a vector of sampling variances. This variance-covariance matrix had the sampling variance for each effect size on the diagonal and the covariance between these measures, which was calculated assuming a 0.5 correlation between effect size sampling variances from the same study (Noble *et al.* 2017), as off-diagonal elements (see Supplementary Information S8 for a sensitivity analysis using a vector of sampling variances rather than the variance-covariance matrix).

## 2.8. General information on analyses

All phylogenetic multilevel meta-analyses and meta-regressions were run using the R package "metafor" v.2.4-0 (Viechtbauer 2010) and included the following random effects: (a) study ID, which encompasses effect sizes extracted from the same study, (b) laboratory ID, which contains effect sizes obtained from the same laboratory, (c) population ID, which considers effect sizes derived from the same population, (d) species name, which encompasses effect sizes derived from the same species, (e) phylogeny, which consists of the phylogenetic correlation matrix and

informs about the relatedness between species, and (f) effect ID, which is a unit-level effect and represents residual/within-study variance.

We present our results by first describing the direction and magnitude of the effect, followed by the statistical significance, and finally by reporting the heterogeneity around the mean estimates. This approach places less focus on dichotomizing the results (discussed by e.g., Amrhein *et al.* 2019) and provides crucial insights into the variance across effect sizes, helping assess the generalizability of our findings (discussed by e.g., Gurevitch *et al.* 2018). We provide information on the effect size, both 95% confidence intervals (CI) and 95% prediction intervals (PI), and p-values. Confidence intervals show the range in which the overall average effect is likely to be found, whereas prediction intervals (which incorporate heterogeneity) estimate the likely range in which 95% of effects are expected to occur in studies (past or future) similar to those included in the current study (IntHout *et al.* 2016). For the intercept-only models, we first performed the Cochrane's *Q* test and then calculated total heterogeneity ($\sigma^2$), its source (or relative heterogeneity; $I^2$) and two different metrics to understand its magnitude – a mean-standardized metric (*CVH2*) and a variance-mean-standardized metric (*M2*; more in Yang *et al.* 2024). For meta-regression models, we estimated the percentage of heterogeneity explained by the moderators as $R^2_{marginal}$ (Nakagawa & Schielzeth 2013). We graphically presented the results as orchard plots using the R package "orchaRd" v.2.0 (Nakagawa *et al.* 2023), which visualizes the overall effect size (i.e., meta-analytic mean), along with its 95% confidence and prediction intervals, and individual effect sizes scaled by their precision (i.e., 1/SE).

## 2.9. Main effect model

To estimate the overall effect size for the association between egg hormones and fitness across studies (BH.1), we fitted a phylogenetic multilevel intercept-only meta-analytic model for our entire dataset (k = 438 effect sizes, N = 57 studies). Note that this model was not pre-registered and includes data for all hormones. We used this approach to understand the overall heterogeneity in our data before we proceeded to explain it with our meta-regressions (see below).

To assess the robustness of our results to the choice of effect size, we additionally ran four sensitivity models that, although not stated in our pre-registration, we considered important to confirm that our results were robust to the choice of effect size. For that, we ran four intercept-only models using three subsets of the data: i) Pearson's *r* and ii) Fisher's *Zr* (k = 164, N = 39), iii) biserial correlations (k = 272, N = 38), and iv) log-response ratio coefficients (lnRR, k = 269, N = 37; calculated using the function 'escalc' from the R package "metafor" v.2.4-0, Viechtbauer

2010). These four sensitivity models reached similar conclusions as the main effect model (Supplementary Information S9).

2.10. Biological predictions

To investigate the association between maternal egg hormones and offspring fitness proxies (BP.1), we use data only including effect sizes on offspring fitness (k = 352, N = 42), and then split it into three subsets, one per hormone type: androgens (i.e., androstenedione, 5α-dihydrotestosterone and testosterone; BP.1.1), glucocorticoids (BP.1.2) and thyroids (i.e., TH3 and TH4; BH.1.3). Next, we fitted three multilevel intercept-only meta-analytic models. Note that in our pre-registration we also proposed to study egg progesterone and estradiol (BP.1.4 & BP.1.5); however, we could not because no study fulfilled our eligibility criteria. To further assess the robustness of our results to our choice of analytical strategy for these three hypotheses (BP.1.1 - BP.1.3), we ran a phylogenetic multilevel meta-regression with 'hormone type' (levels: androgens, glucocorticoids and thyroids) as a moderator. Although the mean estimate for glucocorticoids (k = 56, N = 8) and thyroids (k = 36, N = 3) was slightly different between the two analytical approaches, likely due to the much smaller sample sizes for those hormones compared to androgens (k = 260, N = 33), the estimates remained statistically non-significant (Supplementary Information S8 & S10).

To test if maternal egg hormones have age-specific fitness consequences for the offspring (BP.3) we fitted two meta-regression models. In the first model (BP.3.1), we originally planned to test if the hormone-fitness relationship differed depending on 'hormone type' (levels: androgens, glucocorticoids, and thyroids), 'offspring life history stage' (levels: hatching, before independence, and after independence), and their interaction. However, our final model contained only androgen data and the moderator 'offspring life history stage' (k = 260, N = 33) because we did not have enough data points for the levels glucocorticoids, thyroids, and after independence (i.e., we had less than 5 data points per level; more in our pre-registration Mentesana *et al.* 2021b). In the second model (BP.3.2), we included 'hormone type' (levels: androgens, glucocorticoids and thyroids), 'offspring relative age' as a continuous moderator, and their interaction (k = 218, N = 19).

2.11. Methodological predictions

We performed three phylogenetic multilevel meta-regressions to investigate the impact of methodological differences on the relationship between maternal egg hormones and both maternal and offspring fitness proxies. In the first model (MP.1), we included 'hormone type'

(levels: androgens, glucocorticoids and thyroids), ‘study type’ (levels: correlational and experimental), and their interaction as moderators (k = 397, N = 53). In the second model (MH.2), we fitted ‘hormone type’, ‘experimental study object’ (levels: egg and mother), and their interaction as moderators (k = 247, N = 35). Note that due to a lack of correlational studies testing for the effects of thyroids on fitness and of experimental studies manipulating maternal thyroid hormones, we could only test the first two models (MP.1 & MP.2) for androgens and glucocorticoids. In the third model (MP.3), we originally planned to include ‘hormone type’, ‘experimental dose’ (levels: within natural range and supraphysiological), and their interaction as moderators; however, we could only do so for androgens (k = 89, N = 19) due to a lack of data for the other two hormones, and thus, our final model only included ‘experimental dose’ as the only moderator.

## 2.12. Exploratory hypotheses

We performed three additional pre-registered biological and methodological exploratory meta-regressions for which we did not have directional predictions (Mentesana *et al.* 2021b). To test the effect of maternal egg hormones on maternal fitness (BEH.1) we only included effect sizes on maternal fitness proxies (k = 86, N = 46). We originally planned to run a phylogenetic multilevel meta-regression with ‘hormone type’ (levels: androgens, glucocorticoids and thyroids) as a moderator. However, we could only test this for androgens and glucocorticoids (k = 83, N = 44). To test if the effect of maternal egg hormones on offspring fitness differs depending on the matrix in which the hormone was measured/manipulated (BEH.2), we planned to run a phylogenetic multilevel meta-regression using ‘hormone type’, ‘egg site measured’ (levels: albumen and yolk), and their interaction as moderators. Unfortunately, the data set covered information only about glucocorticoids and for a limited number of studies (k = 38, N = 4), preventing us from drawing reliable conclusions (Supplementary Information S11). Finally, to test the effect of maternal egg hormones on fitness depending on the sampling technique used, we ran a phylogenetic multilevel meta-regression using ‘egg sampling method’ (levels: biopsy and entire egg removed) as moderator (MEH.1), and we were able to test this only for androgens (k = 91, N = 9).

## 2.13. Publication bias tests

We tested for evidence of publication bias (following Nakagawa *et al.* 2022; Sánchez-Tójar *et al.* 2018). We did not find any compelling evidence of publication bias either when testing for small study, decline or reporting effects (see Supplementary Information S12 for further details on our approach and the results).

## 3. Results

Overall, we obtained 438 effect sizes from 57 studies across 19 bird species, 4 bird orders, and 34 different populations (Figure 1). All studies but one were conducted in the Northern Hemisphere. Three species accounted for 61% of the effect sizes (i.e., *Sturnus unicolor*, k = 105 effect sizes; *Parus major*, k = 94; *Ficedula albicollis*, k = 69). Androgens were the most studied hormones (k = 333), followed by glucocorticoids (k = 66) and thyroid hormones (k = 39), and most effect sizes (76%) were obtained from experiments (k = 335). The fitness proxies studied corresponded mainly to offspring (k = 352) rather than mothers (k = 86), and those proxies were primarily obtained before offspring independence (k = 223) and by pooling offspring sexes (k = 309).

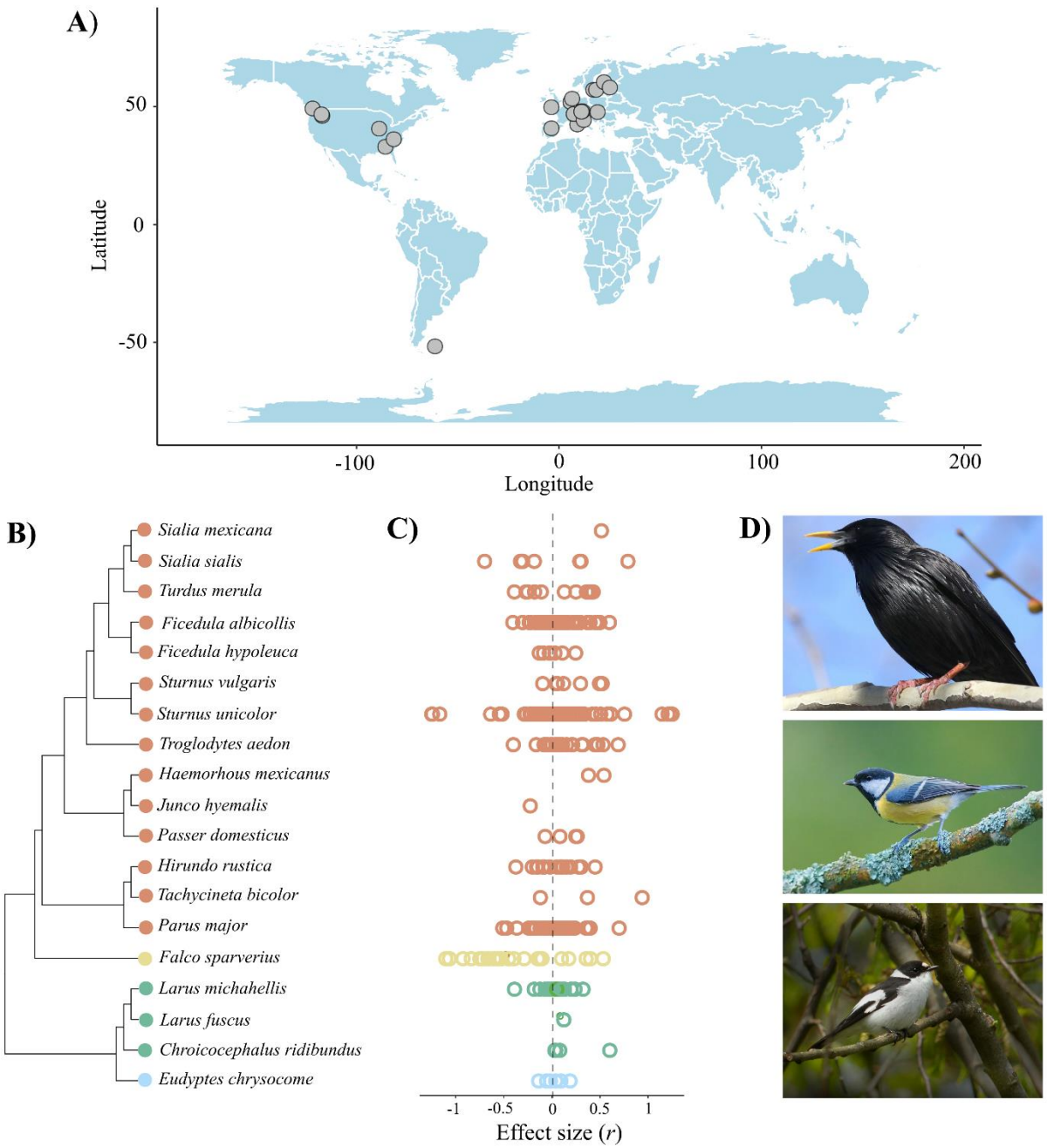



Figure 1: Overview of the meta-analytic data analyzed to test the fitness consequences of maternal egg hormones. A) Global map showing the location of studies included in the meta-analysis that provided longitude and latitude data. Each point represents one study area for which one or more effect sizes were obtained. (B) Phylogenetic tree of the 19 bird species included in the meta-analysis, along with (C) their respective effect sizes. Each color represents a different bird order (red: Passeriformes, yellow: Falconiformes, green: Charadriiformes, and blue: Sphenisciformes). D) Images of the bird species that contributed to most of the data (top to

bottom: *Sturnus unicolor*, *Parus major*, *Ficedula albicollis*). All images are copyright-free (CC BY-SA 4.0 DEED. Authors: Peterwchen, Luc Viatour, Frank Vassen, respectively) and were extracted from www.commons.wikipedia.org

3.1. Main effect model

The intercept-only model revealed a negative, near-zero, and statistically nonsignificant association between egg hormones and fitness - both offspring and maternal - across wild birds (Figure 2; $r$ = -0.072, 95% CI: [-0.331, 0.186], 95% PI: [-0.851, 0.707], p-value = 0.583, k = 438, n = 57). Total absolute heterogeneity ($\sigma^2$ = 0.14; $Q$ = 48122, p-value < 0.001), relative heterogeneity ($I^2_{total}$ = 95.3%), and the two magnitude metrics ($CVH2_{total}$ = 26.7, $M2_{total}$ = 0.96) showed high values of heterogeneity that were beyond the 75$^{th}$ percentile of empirically derived ones in ecological and evolutionary meta-analyses (more in Yang *et al.* 2024; for details see Supplementary Figure 4 & Supplementary Information S13).

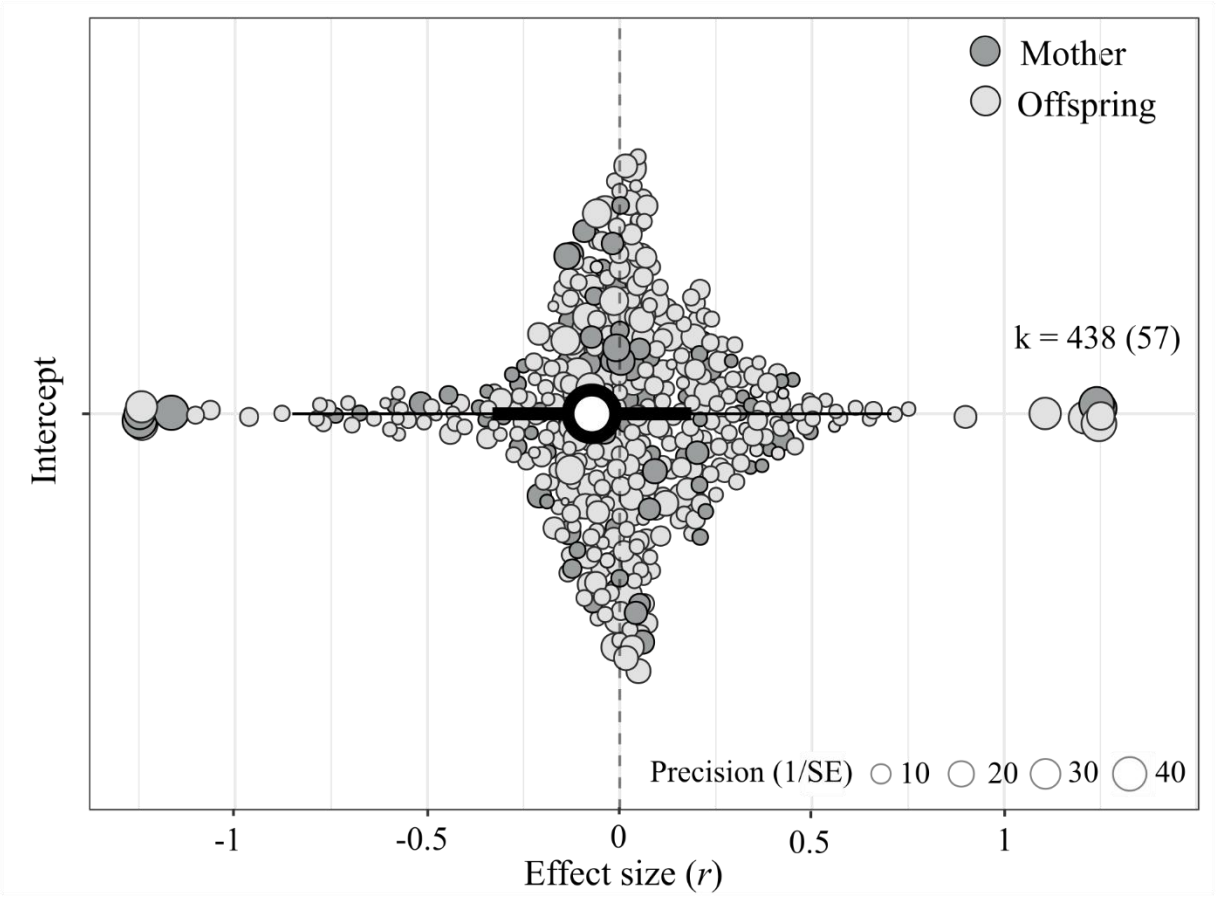



Figure 2: Overall, egg hormone concentrations (androgens, glucocorticoids, and thyroid hormones) show a near-zero association with both maternal (dark gray circles; ca. 20% of effect

sizes) and offspring (light gray circles; ca. 80% of effect sizes) fitness in wild birds. Orchard plot of the phylogenetic multilevel (intercept-only) model showing the meta-analytic mean (black circle), 95% confidence intervals (thick whisker), 95% prediction intervals (thin whisker), and individual effect sizes scaled by their precision (gray circles). k is the number of individual effect sizes, the number of studies is shown in brackets.

## 3.2. Biological predictions

### 3.2.1. Offspring fitness effect sizes (BP.1)

A non-pre-registered intercept-only model confirmed the negative, near-zero, and statistically nonsignificant association between egg hormones and offspring fitness across wild birds ($r$ = -0.063, [95% CI = -0.410, 0.283], [95% PI = -0.928, 0.801], p-value = 0.719, k = 352; N = 42) as well as its high levels of heterogeneity ($\sigma^2$ = 0.16; Q = 22407, p-value < 0.001; $I^2_{total}$ = 95.5%; $CVH2_{total}$ = 40.4, $M2_{total}$ = 0.9). Moreover, effect sizes for the association between all three types of egg hormones and offspring fitness remained small and statistically nonsignificant (Figure 3; details in Supplementary Information S14) and the moderator hormone type explained negligible heterogeneity ($R^2_{marginal}$ = 0.3%). Moreover, none of the three estimates differed statistically from each other (p-value androgens *vs* glucocorticoids = 0.400; androgens *vs* thyroids = 0.533; glucocorticoids *vs* thyroids = 0.970).

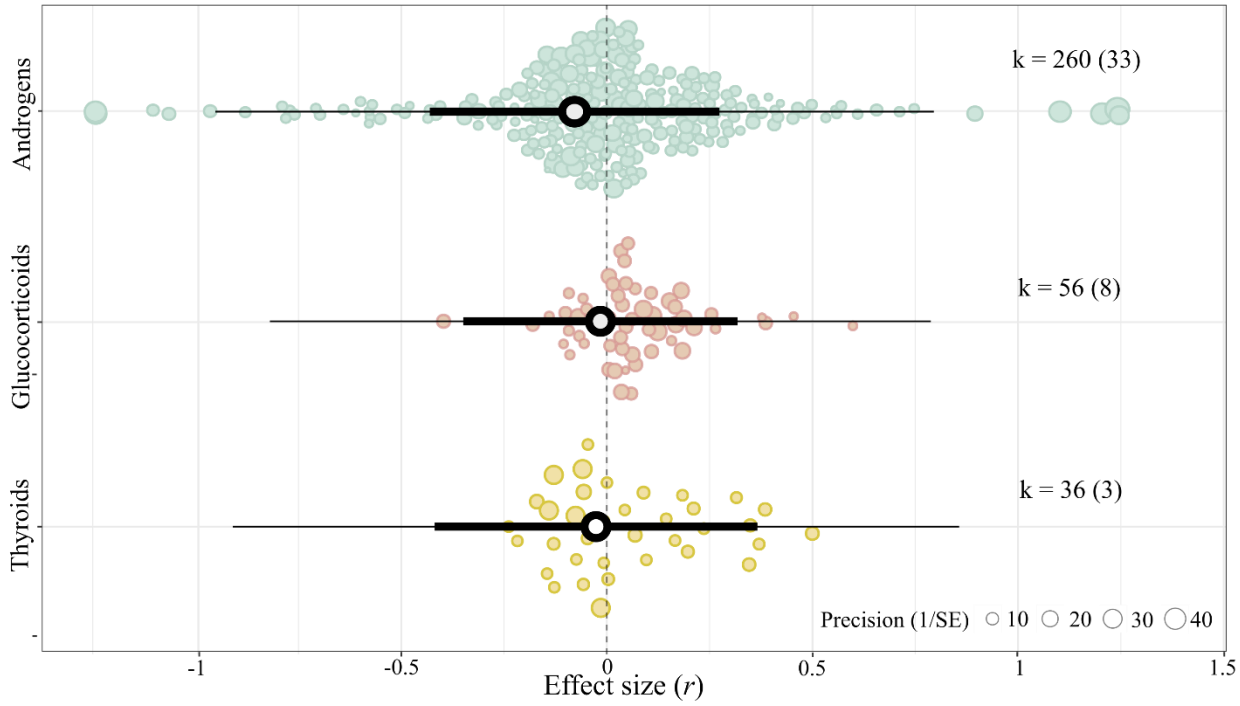



Figure 3: Regardless of hormone type, egg hormone concentrations show a near-zero association with offspring fitness in wild birds. Orchard plot of the phylogenetic multilevel meta-regression showing the mean estimates (black circles), 95% confidence intervals (thick whisker), 95% prediction intervals (thin whisker), and individual effect sizes scaled by their precision (coloured circles). k is the number of individual effect sizes, the number of studies is shown in brackets.

### 3.2.3. Age-specific offspring fitness effect sizes (BP.3)

For both altricial and precocial species, androgens (the only hormone for which we had enough data at different categorical age stages to perform this test; BP.3.1) seemed to be positively associated with offspring fitness at hatching and negatively associated before and after independence. However, effect sizes were small and statistically nonsignificant (Figure 4A; hatching: $r$ = 0.009, [95% CI = -0.389, 0.407], [95% PI = -0.937, 0.956], p-value = 0.963, k = 21, N = 7; before independence: $r$ = -0.076, [95% CI = -0.446, 0.294], [95% PI = -1.012, 0.859], p-value = 0.684 , k = 223, N = 30; after independence: $r$ = -0.174, [95% CI = -0.618, 0.271], [95% PI = -1.141, 0.793], p-value = 0.442, k = 16, N = 4), and the moderator explained negligible heterogeneity ($R^2_{marginal}$ = 0.6%). None of the three age stages differed statistically from each other (p-value hatching *vs* before independence = 0.334; hatching *vs* after independence = 0.240; before *vs* after independence = 0.448).

In altricial species, the association between egg androgen or thyroid concentrations and offspring fitness proxies appeared to weaken throughout the nesting phase, while the association

with glucocorticoids showed the opposite trend (BP.3.2); however, none of those slopes were statistically significant (Figure 4B-D; Table 2) and the moderator only explained little heterogeneity ($R^2_{marginal}$ = 1.7%).

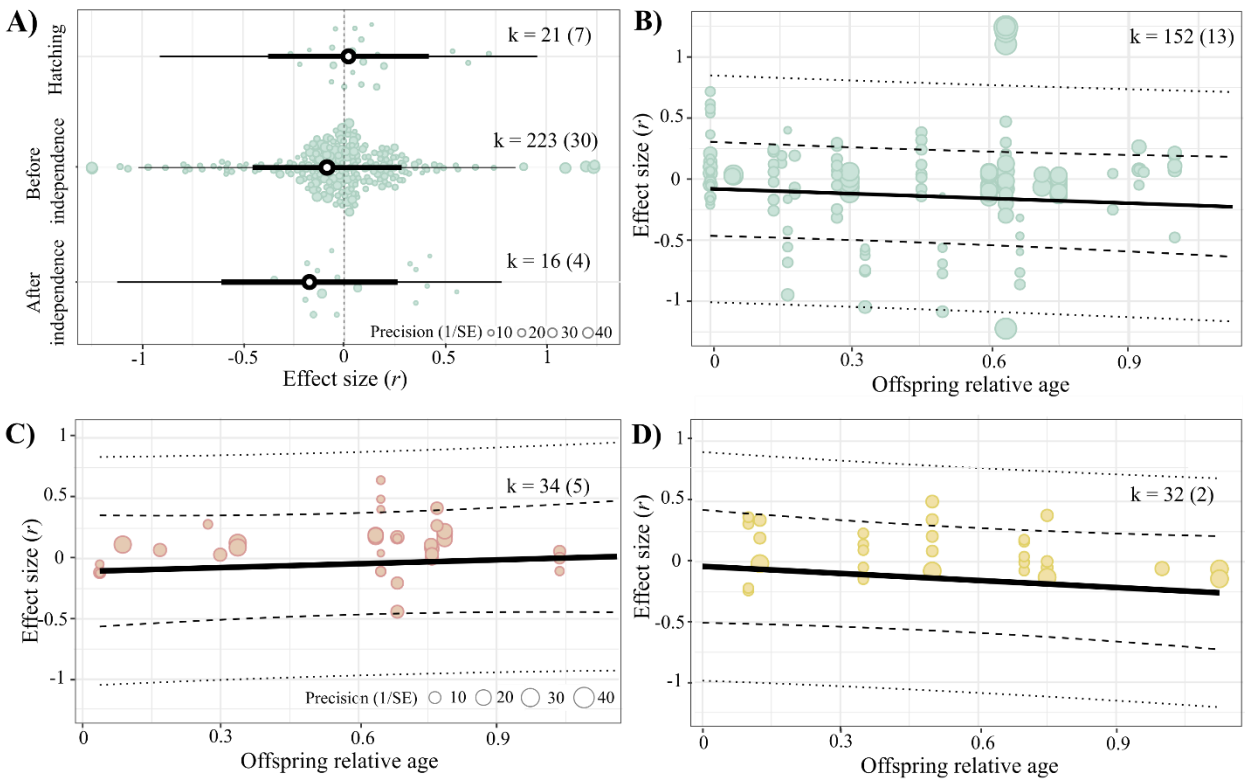



Figure 4: Regardless of offspring age and hormone type, egg hormone concentrations show a near-zero association with offspring fitness in wild birds. A) Egg androgen concentrations and offspring fitness in both altricial and precocial species. Orchard plot of the phylogenetic multilevel meta-regression displaying mean estimates (black circles), 95% confidence intervals (thick whiskers), 95% prediction intervals (thin whiskers), and individual effect sizes scaled by precision (colored circles). Panels B-D refer to B) androgen, C) glucocorticoid and D) thyroid egg hormone concentrations with offspring fitness in altricial species. We plotted a separate bubble plot for each of the hormone types analyzed in a single phylogenetic multilevel meta-regression, showing model estimates (solid lines), 95% CIs (thick dashed lines), 95% prediction intervals (thin dashed lines), and individual effect sizes by precision (colored circles: green for androgens, red for glucocorticoids, yellow for thyroids). k is the number of individual effect sizes, the number of studies is shown in brackets.

Table 2: Results of phylogenetic multilevel meta-regression models testing whether the association between egg hormone concentrations and offspring fitness vary throughout the nesting phase in altricial bird species[‡].

| | Estimate | SE | t-value | df | p-value | L 95% CI | U 95% CI |
|---|---|---|---|---|---|---|---|
| Intercept (Androgens) | -0.088 | 0.197 | -0.445 | 212 | 0.657 | -0.476 | 0.301 |
| Glucocorticoids | -0.005 | 0.152 | 0.033 | 212 | 0.974 | -0.295 | 0.305 |
| Thyroids | 0.071 | 0.153 | 0.465 | 212 | 0.643 | -0.231 | 0.374 |
| Offspring age | -0.132 | 0.099 | -1.331 | 212 | 0.185 | -0.327 | 0.063 |
| Glucocorticoids * Offspring age | 0.224 | 0.220 | 1.019 | 212 | 0.309 | -0.209 | 0.658 |
| Thyroids * Offspring age | -0.060 | 0.209 | -0.289 | 212 | 0.773 | -0.473 | 0.352 |

[‡] df: degrees of freedom; CI: confidence interval. Offspring age was standardized as the age of the offspring divided by the average nesting time and was calculated for each study based on information reported by the authors (more in Table S5.1).

## 3.3. Methodological predictions

### 3.3.1. Fitness effect sizes in correlational *vs* experimental studies (MP.1)

For the two hormone types for which sufficient data existed, androgen and glucocorticoids, the associations between concentrations and both maternal and offspring fitness seemed negative, near-zero, and statistically nonsignificant (Figure 5A; Supplementary Information S15A). The moderator explained little heterogeneity ($R^2_{marginal}$ = 1.4%), and the differences between correlational and experimental were statistically nonsignificant regardless of hormone type (p-value androgens = 0.142; glucocorticoids = 0.340).

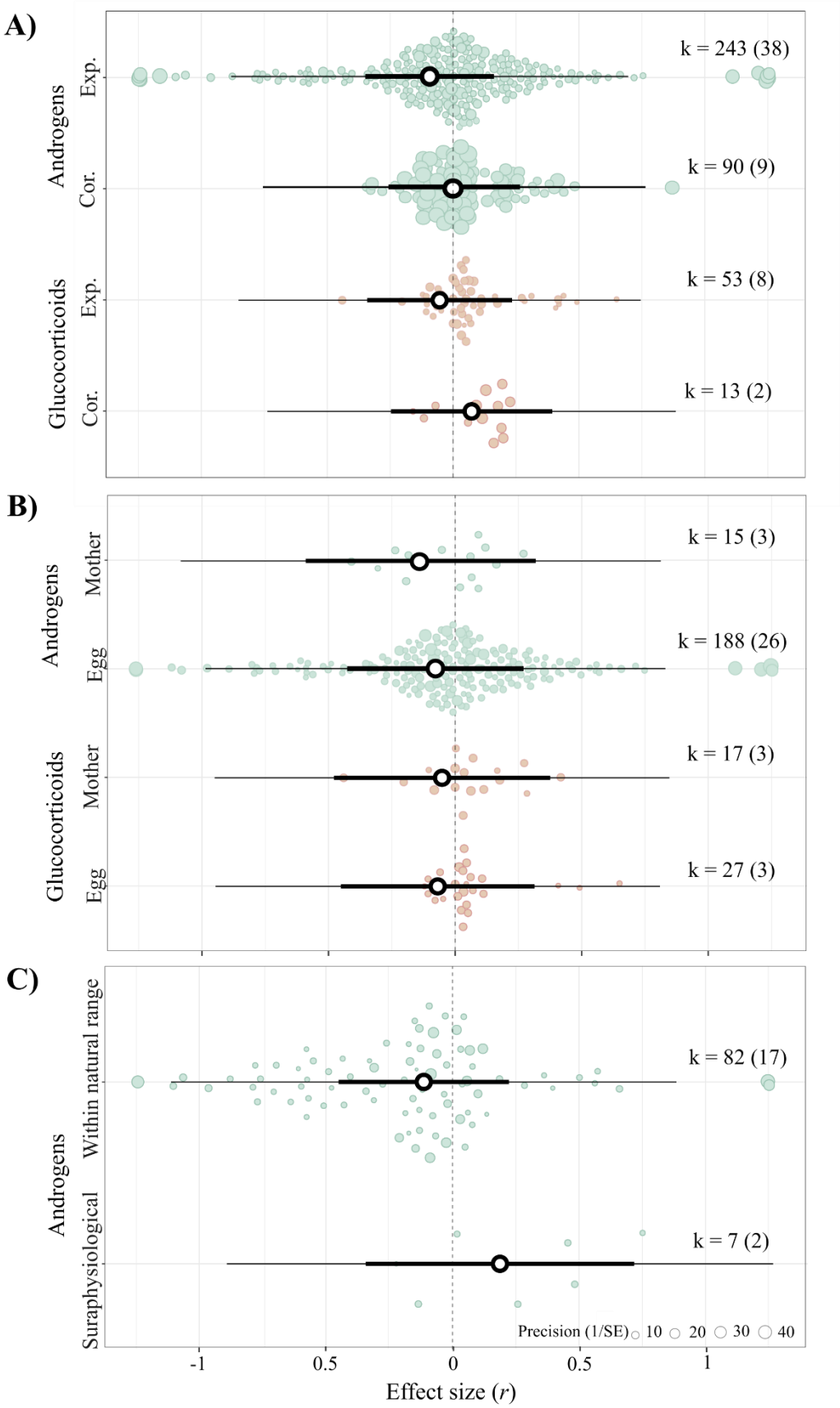
### 3.3.2. Fitness effect sizes in experimental studies performed on the mother *vs* the egg (MP.2)

The negative but statistically nonsignificant association between egg androgen or glucocorticoid concentrations and offspring fitness remained independent of whether hormonal manipulations were performed on the mother or the egg (Figure 5B; Supplementary Information S15B; p-value

mother *vs* eggs = androgens 0.654, glucocorticoids 0.985). The moderator explained negligible heterogeneity ($R^2_{marginal}$ = 0.2%).

3.3.3. Fitness effect sizes in experimental studies conducted within *vs* above the natural physiological range (MP.3)

Although the association between egg androgens and offspring fitness appeared to be positive at doses above the physiological range and seemed negative at doses within the natural range, none of the estimates were statistically significant (Figure 5C; Supplementary Information S15C) or differed from each other (p-value = 0.175). Moreover, the moderator explained little heterogeneity ($R^2_{marginal}$ = 2.9%).


A)
Androgens
Exp.
Cor.
Glucocorticoids
Exp.
Cor.
k = 243 (38)
k = 90 (9)
k = 53 (8)
k = 13 (2)
B)
Androgens
Mother
Egg
Glucocorticoids
Mother
Egg
k = 15 (3)
k = 188 (26)
k = 17 (3)
k = 27 (3)
C)
Androgens
Within natural range
Suraphysiological
k = 82 (17)
k = 7 (2)
Precision (1/SE) 10 20 30 40
-1
0.5
0
0.5
1
Effect size (r)

541

Figure 5: The relationship between androgen and glucocorticoid egg hormone concentrations and fitness is near-zero, irrespective of A) the type of study performed (i.e., correlational or experimental), B) the study object on which the experiment was performed (i.e., the mother or the egg), and, solely for androgens, C) the experimental dose used (i.e., within or above the natural range). Orchard plots of three phylogenetic multilevel meta-regression models showing mean estimates (black circles), 95% confidence intervals (thick whisker), 95% prediction intervals (thin whisker), and individual effect sizes scaled by their precision (coloured circles). k is the number of individual effect sizes, the number of studies is shown in brackets.

### 3.4. Exploratory biological and methodological hypotheses

The association between egg hormones and maternal fitness was statistically nonsignificant both for androgens and glucocorticoids (BEH.1; androgens: $r$ = -0.052, [95% CI = -0.159, 0.054], [95% PI = -0.826, 0.722], p-value = 0.332, k = 73, N = 36; glucocorticoids: $r$ = 0.082, [95% CI = -0.183, 0.348], [95% PI = -0.729, 0.893], p-value = 0.540, k = 10, N = 8). The moderator explained negligible heterogeneity ($R^2_{marginal}$ = 1.3%), and the two hormone groups did not differ statistically from each other (p-value = 0.351).

We found that the association between maternal egg androgen concentrations on fitness did not differ statistically between the two sampling techniques used (p-value = 0.196; MEH.1). When androgens were measured using an egg biopsy, they seemed to be negatively associated with fitness, whereas when the entire egg was removed, they appeared to be positively associated with fitness (egg biopsy: $r$ = -0.202, [95% CI = -0.681, 0.278], [95% PI = -1.025, 0.622], p-value = 0.406, k = 19, N = 2; entire egg removed: $r$ = 0.144, [95% CI = -0.089, 0.377], [95% PI = -0.565, 0.853], p-value = 0.223, k = 74, N = 8; $R^2_{marginal}$ = 14.7%).

Due to small sample sizes, these results should be interpreted cautiously as the models likely are heavily underpowered.

## 4. Discussion

Our pre-registered systematic review and meta-analysis found that the fitness effects of egg hormones across 19 wild bird species are neither as strong nor as clear as previously suggested. The overall mean effect of androgens, glucocorticoids, and thyroid hormones on both maternal and offspring fitness was statistically indistinguishable from zero. However, those effects were also highly heterogeneous, suggesting that the magnitude and direction of the relationship between hormone concentrations and fitness is context specific. We tested several biological and methodological moderators that we predicted could explain the differences in fitness effects for

offspring, for which most of the data exists, but they only accounted for a small percentage of the heterogeneity. Below, we discuss our findings in detail, examining our hypotheses in specific contexts, and highlighting knowledge gaps, biases in the current literature, and areas for future research.

The influence of egg hormone concentrations on fitness has historically been expected to vary by hormone group: increased egg androgens and thyroid hormones are predicted to enhance fitness, while increased glucocorticoids are expected to reduce it (e.g., von Engelhardt & Groothuis 2011; Groothuis *et al.* 2005a; Haussmann *et al.* 2012; Hsu *et al.* 2017; Polich *et al.* 2018; Ruuskanen *et al.* 2018; Stier *et al.* 2009). However, our results showed that the relationship between hormone concentration and offspring fitness proxies was similarly small and statistically indistinguishable from zero for all three hormone groups, with 'hormone type' explaining a negligible 0.4% of the heterogeneity among effect sizes. Our exploratory analysis testing the relationship with maternal fitness for the two hormones for which sufficient data were available (i.e., androgens and glucocorticoids) also showed a small and statistically nonsignificant relationship. In all, our results, including all sensitivity analyses, robustly show no support for a consistent overall relationship between egg hormones and fitness in wild bird populations.

Egg hormone concentrations are also suggested to have sex-specific effects on offspring fitness (reviewed by e.g., von Engelhardt & Groothuis, 2011). While we aimed to test this hypothesis (see BP.2 in the Introduction), data on offspring sex-specific effects were available in only 12.2% of cases (43 out of 352 offspring effect sizes, 7 studies). This limited dataset included only two hormone types (androgens and thyroid hormones) and three bird species (*Parus major*, *Ficedula albicollis*, and *Falco sparverius*). The relative availability of sex-specific data in our study was even lower than the 32% reported by Podmokła et al. (2018) for experimental studies conducted primarily in captive birds. Notably, while Podmokła et al. (2018) included data on androgens and glucocorticoids, we found information only for androgens and thyroid hormones, and overall, 56.1% of the studies in our dataset were not included in theirs. Altogether, this clearly indicates the current lack of sex-specific data for all egg hormones in both captive and wild birds, which limits our understanding of whether maternal egg hormones have sex-specific fitness consequences.

The relationship between egg hormones and offspring fitness proxies showed no overall change with age. Theory predicts that maternal effects should weaken as offspring encounter more environments throughout ontogeny (e.g., Mousseau & Dingle 1991), and a meta-analysis of maternal (genetic) effects on vertebrate species supported this prediction (Moore *et al.* 2019). The absence of an ontogenetic decline in our study might be because egg hormones already had

a weak effect at hatching, despite being expected to be strongest at this stage compared to later ones. Moreover, Moore *et al.* (2019) found that maternal (genetic) effects declined more for traits with higher maternal influence such as morphological traits compared to other traits such as behavior and physiology. Most studies providing offspring life-stage or age focused on androgens and, thus, we urge researchers to gather more data on glucocorticoids and thyroid hormones along developmental stages to better understand the relationship between hormones and fitness across life stages.

In wild bird populations, the overall null effect between hormones and fitness proxies appears consistent, even when accounting for methodological differences across studies. Specifically, we found that egg androgens and glucocorticoids had no overall effect on fitness proxies regardless of study type (correlational or experimental) or experimental focus (mother or egg). The experimental dose being within the natural range or above physiological concentrations, and the sampling technique used (egg biopsy or entire egg collection) explained little to negligible amounts of the observed heterogeneity. These results are also relevant, as they suggest that manipulating egg hormone concentrations neither leads to fitness gains, as expected when females face physiological constraints or trade-offs, nor incurs fitness costs, as predicted if maternal egg hormones are adaptive (von Engelhardt & Groothuis 2011). A methodological issue that requires further consideration is the variation in extraction methods across laboratories, particularly for glucocorticoids. In immunoassays, glucocorticoid antibodies can cross-react with progesterone and its precursors, leading to an overestimation of yolk glucocorticoid concentrations (Rettenbacher *et al.* 2009, 2013). Column chromatography is recommended to prevent this, but likely due to the complexity of this method, only 2 of 10 laboratories studying egg glucocorticoids and fitness proxies employed it (Bowers et al. 2016; Mentesana et al. 2021a). Overestimated glucocorticoid concentrations resulting from methodological issues might contribute to discrepancies across labs in the observed relationships between egg glucocorticoids and fitness proxies. Supporting this idea, we found that relative heterogeneity was highly explained by differences among laboratories ($I^2_{\text{among-laboratory}}$ = 42.17%; Supplementary Information Table S10.2).

By using a pluralistic approach to estimate total heterogeneity and stratify it across the different hierarchical levels explored, we found that the overall association between egg hormones and both maternal and offspring fitness was highly heterogeneous and stemming from within-study variation and phylogenetic relationships among species (Supplementary Information S10.2 & S13). The high levels of heterogeneity associated with within- compared to among-study variation, the latter being zero in all models (and robust to sensitivity tests), suggest that the small

and statistically nonsignificant overall effect found by our meta-analysis (i.e., our meta-analytic mean) would be generalizable among studies as long as (i) we define replication as the testing of the null hypothesis at the among-study level and, (ii) importantly, if within-study level variation - both biological and methodological, are accounted for (Yang *et al.* 2024). Our moderator analyses testing for both biological and methodological differences, however, did not explain a substantial amount of heterogeneity, limiting our potential for generalization. We hypothesized that such differences may stem from researchers reporting multiple fitness proxies within studies (mean ± SD = 2.96 ± 1.86) and conducted an additional (non-pre-registered) meta-regression using the entire dataset. None of the fitness proxy-specific mean estimates were statistically significantly different from zero (see Supplementary Figure 5; Supplementary Information S16), and the fitness proxy used accounted for a small but somewhat non-negligible percentage of the heterogeneity ($R^2_{marginal}$ = 3.5%). Among-species differences, and particularly, phylogenetic relationships among the 19 bird species were associated with high levels of heterogeneity, which indicates that the size and magnitude of the overall association between egg hormones and fitness proxies differ among taxonomic groups (Figure 1C). We explored whether species-specific characteristics that are strongly correlated with phylogeny may explain part of that heterogeneity. For this, we ran a (non-pre-registered) model to compare precocial *vs* altricial species and found that developmental mode explains little heterogeneity ($R^2_{marginal}$ = 0.3%; see Supplementary Information S17). To better understand the factors driving these differences, future research could focus on understanding if species-specific characteristics (e.g., life-history, ecology) might be driving the phylogenetic signal observed in our study as well as investigate whether egg hormones and fitness are associated in other bird species not yet studied.

Three additional factors might explain the heterogeneity that we found across effect sizes. First, one complex factor that we could not account for is the environment experienced by mothers and offspring. Theory predicts that anticipatory maternal effects evolve when mothers and offspring encounter similar environments, or when environmental fluctuations are low (Hoyle & Ezard 2012; Kuijper & Hoyle 2015). Egg hormones may significantly influence fitness proxies only when maternal and offspring environments match, and/or when mothers can predict the environment their offspring will encounter. Conversely, maternal effects may be non-anticipatory, with heterogeneity in fitness outcomes arising from the cost-benefit balance of egg hormones under varying environmental scenarios experienced solely by the offspring (reviewed by Groothuis *et al.* 2020). Second, maternal effects are multivariate. Mothers influence their offspring's prenatal environment by depositing various compounds into the egg besides hormones (e.g., immunoglobulins, antioxidants, fatty acids, proteins; Lovern & Wade 2001; McCormick

1999; Mentesana *et al.* 2019; Valcu *et al.* 2019; Weiss *et al.* 2011), which may interact (Groothuis *et al.* 2019; Mentesana *et al.* 2021a; Williams & Groothuis 2015). If these egg components do not act independently on fitness traits, and their correlations are weak or different combinations have similar outcomes, testing the effects of single components independently, as we did, is likely to result in high heterogeneity. Third, there is growing evidence that embryos are not passive responders to maternal signals (Bebbington & Groothuis 2021; Groothuis *et al.* 2019). Embryos seem adapted to receiving maternal hormone signals, as they have hormone receptors before producing their own hormones (Godsave *et al.* 2002; Kumar *et al.* 2019). However, recent evidence suggests that embryos can rapidly metabolize maternal hormones in the egg into potentially inactive steroid forms (Kumar *et al.* 2018; Vassallo *et al.* 2014). If the hormone concentrations in the egg do not reflect the amount actually incorporated by the offspring, this could explain the variation in effect sizes and the overall lack of effect that manipulating egg hormone concentrations has on offspring fitness.

Egg hormone-mediated maternal effects are considered an adaptive mechanism, where mothers adjust hormone concentrations in response to environmental cues to optimize maternal and/or offspring fitness. This assumes that mothers can modify hormone levels based on the environment, in addition to the maternal effects on fitness explored in this study. However, it is debated whether hormone deposition is truly female-controlled (von Engelhardt & Groothuis 2011; Gil 2008; Groothuis & Schwabl 2008). Groothuis & Schwabl (2008) proposed three mechanisms: (1) egg hormone concentrations mirror maternal plasma levels, (2) eggs act as a sink for excess steroids, or (3) egg and maternal hormone levels are independently regulated. The first two mechanisms suggest egg hormones reflect maternal responses to the environment or physiology, not being offspring-specific. If mothers cannot adjust hormone transfer, adaptive maternal effects may be constrained. Understanding which mechanism is operating and whether it is common among species is pivotal to the study of the correlations between egg hormones concentrations and fitness.

## 5. Conclusions

Thirty years after the discovery of egg hormones, our meta-analysis showed that an increase in their concentrations has a near zero and highly variable overall mean effect on fitness proxies in wild bird populations. We could not identify any predictor explaining much of the high levels of heterogeneity found among effect sizes, even after accounting for various biological (e.g., hormone type and age) and methodological (e.g., study type, experimental subject, dose) factors.

We hope our work helps advance the field of maternal effects via egg hormone deposition and inspires new research into its many intriguing aspects. Key areas for future research include studying phylogenetically distant species, investigating whether the near-zero mean effect of egg hormones on both maternal and offspring fitness proxies can be explained by environmental context, studying sex-specific effects of egg hormones, and clarifying the mechanisms of maternal transfer to eggs. Such research questions are essential for understanding the adaptive significance of maternal effects via egg deposition, so we urge researchers in the field to address them.

## Acknowledgements

We thank all the authors who replied to our email, addressed our questions, and kindly shared their data: Alexandra Bentz, Amber Rice, Barbara Tschirren, Renee Duckworth, Bin-Yan Hsu, E. Diego Rubolini, Keith Bowers, Gergely Hegyi, Ismael Galvan, Jaime Muriel, János Török, Jose Carlos Noguera, Keith W. Sockman, Kristen J. Navara, Marco Parolini, Scott Sakaluk, and Suvi Ruuskanen. We also want to thank Yefeng Yang for his feedback on interpreting heterogeneity results. This research was funded by the Max Planck Gesellschaft (to MH).

## Author's contributions

Conceptualization: LM, MH and AST. Data curation: LM and AST. Formal analysis: LM and AST. Funding acquisition: MH. Investigation: LM, PBD, and NMA. Methodology: LM and AST. Project administration: LM. Software: LM and AST. Supervision: LM and AST. Validation: LM, PBD, NMA, and AST. Visualization: LM and NMA. Writing – original draft: LM. Writing – review & editing: all authors.

## Data availability statement

All data and code are available on GitHub (https://github.com/ASanchez-Tojar/meta-analysis_egg_hormones_and_fitness). Upon acceptance of this work, the repository will be assigned with a DOI via Zenodo.

## References

Amrhein, V., Greenland, S. & Blake McShane. (2019). Scientists rise up against statistical significance." Nature 567.7748 (2019): 305-307. *Nature*, 567, 305–307.

Bebbington, K. & Groothuis, T.G.. (2021). Who listens to mother ? A whole-family perspective on the evolution of maternal hormone allocation. *Biol. Rev.*, 4.

Bentz, A.B., Navara, K.J. & Siefferman, L. (2013). Phenotypic plasticity in response to breeding density in tree swallows: An adaptive maternal effect? *Horm. Behav.*, 64, 729–736.

Bentz, A.B., Sirman, A.E., Wada, H., Navara, K.J. & Hood, W.R. (2016). Relationship between maternal environment and DNA methylation patterns of estrogen receptor alpha in wild Eastern Bluebird (Sialia sialis) nestlings: a pilot study. *Ecol. Ev*.

Borenstein, M., Hedges, L. V. & Higgins, J. P. T., Rothstein, H.R. (2009). *Introduction to meta-analysis.*

Bowers, E.K., Bowden, R.M., Sakaluk, S.K. & Thompson, C.F. (2015). Immune activation

generates corticosterone-mediated terminal reproductive investment in a wild bird. *Am. Nat.*, 185, 769–783.

Bowers, E.K., Bowden, R.M., Thompson, C.F. & Sakaluk, S.K. (2016). Elevated corticosterone during egg production elicits increased maternal investment and promotes nestling growth in a wild songbird. *Horm. Behav.*, 83, 6–13.

Clairardin, S.G., Barnett, C.A., Sakaluk, S.K. & Thompson, C.F. (2011). Experimentally increased in ovotestosterone leads to increased plasma bactericidal activity and decreased cutaneous immune response in nestling house wrens. *J. Exp. Biol.*, 214, 2778–2782.

Clotfelter, E.D., O'Neal, D.M., Gaudioso, J.M., Casto, J.M., Parker-Renga, I.M., Snajdr, E.A., *et al.* (2004). Consequences of elevating plasma testosterone in females of a socially monogamous songbird: Evidence of constraints on male evolution? *Horm. Behav.*, 46, 171–178.

Duckworth, R.A., Belloni, V. & Anderson, S.R. (2015). Cycles of species replacement emerge from locally induced maternal effects on offspring behavior in a passerine bird. *Science (80-. ).*, 347, 875–877.

Eising, C.M. & Groothuis, T.G.G. (2003). Yolk androgens and begging behaviour in black-headed gull chicks: an experimental field study. *Anim. Behav.*, 66, 1027–1034.

von Engelhardt, N. & Groothuis, T.G.G. (2011). *Maternal Hormones in Avian Eggs*. *Horm. Reprod. Vertebr. - Vol. 4*. Elsevier Inc.

Galván, I. & Alonso-Alvarez, C. (2010). Yolk testosterone shapes the expression of a melanin-based signal in great tits: An antioxidant-mediated mechanism? *J. Exp. Biol.*, 213, 3127–3130.

Gil, D. (2008). *Chapter 7 Hormones in Avian Eggs: Physiology, Ecology and Behavior*. Advances in the Study of Behavior.

Gil, D., Ninni, P., Lacroix, A., De Lope, F., Tirard, C., Marzal, A., *et al.* (2006). Yolk androgens in the barn swallow (Hirundo rustica): a test of some adaptive hypotheses. *J. Evol. Biol.*, 19, 123–131.

Godsave, S.F., Lohmann, R., Vloet, R.P.M. & Gahr, M. (2002). Androgen receptors in the embryonic zebra finch hindbrain suggest a function for maternal androgens in perihatching survival. *J. Comp. Neurol.*, 453, 57–70.

Grafen, A. (1989). The phylogenetic regression. *Philos. Trans. R. Soc. London. B, Biol. Sci.*, 326, 119–157.

Groothuis, T.G.. & Schwabl, H. (2008). Hormone-mediated maternal effects in birds: mechanisms matter but what do we know of them? *Philos. Trans. R. Soc. London B Biol.*

*Sci.*, 363.

Groothuis, T.G., Eising, C.M., Dijkstra, C. & Müller, W. (2005a). Balancing between costs and benefits of maternal hormone deposition in avian eggs. *Biol. Lett.*, 1, 78–81.

Groothuis, T.G., Kumar, N. & Hsu, B.-Y. (2020). Explaining discrepancies in the study of maternal effects: the role of context and embryo. *Curr. Opin. Behav. Sci.*, 36, 185–192.

Groothuis, T.G.G. & Von Engelhardt, N. (2005). Investigating maternal hormones in avian eggs: Measurement, manipulation, and interpretation. *Ann. N. Y. Acad. Sci.*, 1046, 168–180.

Groothuis, T.G.G., Hsu, B.-Y., Kumar, N. & Tschirren, B. (2019). Revisiting mechanisms and functions of prenatal hormone-mediated maternal effects using avian species as a model. *Philos. Trans. R. Soc. B Biol. Sci.*, 374, 20180115.

Groothuis, T.G.G., Müller, W., von Engelhardt, N., Carere, C. & Eising, C. (2005b). Maternal hormones as a tool to adjust offspring phenotype in avian species. *Neurosci. Biobehav. Rev.*, 29, 329–352.

Gurevitch, J., Koricheva, J., Nakagawa, S. & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, 555, 175–182.

Haussmann, M.F., Longenecker, A.S., Marchetto, N.M., Juliano, S.A. & Bowden, R.M. (2012). Embryonic exposure to corticosterone modifies the juvenile stress response, oxidative stress and telomere length. *Proc. R. Soc. London B Biol. Sci.*, 279.

Hegyi, G., Herényi, M., Szöllősi, E., Rosivall, B., Török, J. & Groothuis, T.G.G. (2011). Yolk androstenedione, but not testosterone, predicts offspring fate and reflects parental quality. *Behav. Ecol.*, 22, 29–38.

Hennessy, E.A., Acabchuk, R.L., Arnold, P.A., Dunn, A.G., Foo, Y.Z., Johnson, B.T., *et al.* (2022). Ensuring Prevention Science Research is Synthesis-Ready for Immediate and Lasting Scientific Impact. *Prev. Sci.*, 23, 809–820.

Hinchliff, C.E., Smith, S.A., Allman, J.F., Burleigh, J.G., Chaudhary, R., Coghill, L.M., *et al.* (2015). Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proc. Natl. Acad. Sci.*, 112, 12764–12769.

Hoyle, R.B. & Ezard, T.H.G. (2012). The benefits of maternal effects in novel and in stable environments. *J. R. Soc. Interface*, 9, 2403–2413.

Hsu, B.-Y., Dijkstra, C. & Groothuis, T.G.G. (2017). Organizing effects of adverse early-life condition on body mass, compensatory growth and reproduction: experimental studies in rock pigeons. *J. Avian Biol.*, 48, 1166–1176.

Hsu, B.-Y., Doligez, B., Gustafsson, L. & Ruuskanen, S. (2019). Transient growth-enhancing effects of elevated maternal thyroid hormones at no apparent oxidative cost during early

postnatal period. *J. Avian Biol.*, 50.

Hsu, B.-Y., Sarraude, T., Cossin-Sevrin, N., Crombecque, M., Stier, A. & Ruuskanen, S. (2020). Testing for context-dependent effects of prenatal thyroid hormones on offspring survival and physiology: an experimental temperature manipulation. *Sci. Rep.*, 10.

Ihle, M., Winney, I.S., Krystalli, A. & Croucher, M. (2017). Striving for transparent and credible research: practical guidelines for behavioral ecologists. *Behav. Ecol.*, 28, 348–354.

IntHout, J., Ioannidis, J.P.A., Rovers, M.M. & Goeman, J.J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, 6, e010247.

Jacobs, P. & Viechtbauer, W. (2017). Estimation of the biserial correlation and its sampling variance for use in meta-analysis. *Res. Synth. Methods*, 8, 161–180.

Krist, M. (2011). Egg size and offspring quality: a meta-analysis in birds. *Biol. Rev.*, 86, 692–716.

Kuijper, B. & Hoyle, R.B. (2015). When to rely on maternal effects and when on phenotypic plasticity? *Evolution (N. Y).*, 69, 950–968.

Kumar, N., van Faassen, M., Kema, I., Gahr, M. & Groothuis, T.G.G. (2018). Early embryonic modification of maternal hormones differs systematically among embryos of different laying order: A study in birds. *Gen. Comp. Endocrinol.*, 269, 53–59.

Kumar, N., Lohrentz, A., Gahr, M. & Groothuis, T.G.G. (2019). Steroid receptors and their regulation in avian extraembryonic membranes provide a novel substrate for hormone mediated maternal effects. *Sci. Rep.*, 9, 11501.

Lajeunesse, M.J. (2013). 13. Recovering Missing or Partial Data from Studies: A Survey of Conversions and Imputations for Meta-analysis. In: *Handbook of Meta-analysis in Ecology and Evolution* (eds. Koricheva, J., Gurevitch, J. & Mengersen, K.). Princeton University Press, Princeton, pp. 195–206.

López-Rull, I. & Gil, D. (2009). Elevated testosterone levels affect female breeding success and yolk androgen deposition in a passerine bird. *Behav. Processes*, 82, 312–318.

Love, O.P., Chin, E.H., Wynne-Edwards, K.E. & Williams, T.D. (2005). Stress hormones: a link between maternal condition and sex-biased reproductive investment. *Am. Nat.*, 166, 751–766.

Lovern, M.B. & Wade, J. (2001). Maternal plasma and egg yolk testosterone concentrations during embryonic development in green anoles (Anolis carolinensis). *Gen. Comp. Endocrinol.*, 124, 226–235.

Marshall, D.J. & Uller, T. (2007). When is a maternal effect adaptive? *Oikos*, 116, 1957–1963.

McCormick, M.I. (1999). Experimental test of the effect of maternal hormones on larval quality of

a coral reef fish. *Oecologia*, 118, 412–422.

Mentesana, L., Andersson, M.N., Casagrande, S., Goymann, W., Isaksson, C. & Hau, M. (2021a). Natural variation in yolk fatty acids, but not androgens, predicts offspring fitness in a wild bird. *Front. Zool.*, 18, 1–12.

Mentesana, L., Hau, M., Adreani, M.N., D'amelio, P. & Sánchez-Tójar, A. (2021b). Fitness consequences of egg hormones in the light of evidence synthesis. Retrieved from osf.io/st38j. *OSF*.

Mentesana, L., Isaksson, C., Goymann, W., Andersson, M.N., Trappschuh, M. & Hau, M. (2019). Female variation in allocation of steroid hormones, antioxidants and fatty acids: a multilevel analysis in a wild passerine bird. *J. Avian Biol.*, 50, 1–16.

Michonneau, F., Brown, J.W. & Winter, D.J. (2016). rotl: an R package to interact with the Open Tree of Life data. *Methods Ecol. Evol.*, 7, 1476–1481.

Moore, M.P., Whiteman, H.H. & Martin, R.A. (2019). A mother's legacy: the strength of maternal effects in animal populations. *Ecol. Lett.*, 22, 1620–1628.

Mousseau, T.A. & Dingle, H. (1991). Maternal effects in insect life histories. *Annu. Rev. Entomol. Vol. 36*, 511–534.

Mousseau, T.A. & Fox, C.W. (1998). The adaptive significance of maternal effects. *Trends Ecol. Evol.*

Müller, W., Deptuch, K., López-Rull, I. & Gil, D. (2007). Elevated yolk androgen levels benefit offspring development in a between-clutch context. *Behav. Ecol.*, 18, 929–936.

Müller, W., Dijkstra, C. & Groothuis, T.G.G. (2009). Maternal yolk androgens stimulate territorial behaviour in black-headed gull chicks. *Biol. Lett.*, 5, 586–588.

Müller, W., Groothuis, T.G.G., Kasprzik, A., Dijkstra, C., Alatalo, R. V & Siitari, H. (2005). Prenatal androgen exposure modulates cellular and humoral immune function of black-headed gull chicks. *Proc. R. Soc. B Biol. Sci.*, 272, 1971–1977.

Muriel, J., Pérez-Rodríguez, L., Puerta, M. & Gil, D. (2013). Differential effects of yolk testosterone and androstenedione in embryo development and nestling growth in the spotless starling (Sturnus unicolor). *Gen. Comp. Endocrinol.*, 194, 175–182.

Muriel, J., Pérez-Rodríguez, L., Puerta, M. & Gil, D. (2015a). Diverse dose-response effects of yolk androgens on embryo development and nestling growth in a wild passerine. *J. Exp. Biol.*, 218, 2241–9.

Muriel, J., Salmón, P., Nunez-Buiza, A., de Salas, F., Pérez-Rodríguez, L., Puerta, M., *et al.* (2015b). Context-dependent effects of yolk androgens on nestling growth and immune function in a multibrooded passerine. *J. Evol. Biol.*, 28, 1476–1488.

Nagy, G., Blázi, G., Hegyi, G. & Török, J. (2016). Side-specific effect of yolk testosterone elevation on second-to-fourth digit ratio in a wild passerine. *Sci. Nat.*, 103.

Nakagawa, S. & Cuthill, I.C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biol. Rev.*, 82, 591–605.

Nakagawa, S., Lagisz, M., Jennions, M.D., Koricheva, J., Noble, D.W.A., Parker, T.H., *et al.* (2022). Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods Ecol. Evol.*, 13, 4–21.

Nakagawa, S., Lagisz, M., O'Dea, R.E., Pottier, P., Rutkowska, J., Senior, A.M., *et al.* (2023). orchaRd 2.0: An R package for visualising meta-analyses with orchard plots. *Methods Ecol. Evol.*, 14, 2003–2010.

Nakagawa, S. & Schielzeth, H. (2013). A general and simple method for obtaining $R^2$ from generalized linear mixed-effects models. *Methods Ecol. Evol.*, 4, 133–142.

Navara, K.J., Hill, G.E. & Mendonça, M.T. (2005). Variable effects of yolk androgens on growth, survival, and immunity in eastern bluebird nestlings. *Physiol. Biochem. Zool.*, 78, 570–578.

Navara, K.J., Hill, G.E. & Mendonça, M.T. (2006). Yolk testosterone stimulates growth and immunity in house finch chicks. *Physiol. Biochem. Zool.*, 79, 550–555.

Noble, D.W.A., Lagisz, M., O'dea, R.E. & Nakagawa, S. (2017). Nonindependence and sensitivity analyses in ecological and evolutionary meta-analyses. *Mol. Ecol.*, 26, 2410–2425.

Noguera, J.C., Kim, S.-Y. & Velando, A. (2013). Maternal testosterone influences a begging component that makes fathers work harder in chick provisioning. *Horm. Behav.*, 64, 19–25.

Noguera, J.C., da Silva, A. & Velando, A. (2020). Egg corticosterone can stimulate telomerase activity and promote longer telomeres during embryo development. *Mol. Ecol.*

Norton, J.M. & Wira, C.R. (1977). Dose-related effects of the sex hormones and cortisol on the growth of the bursa of fabricius in chick embryos. *J. Steroid Biochem.*, 8, 985–987.

O'Dea, R.E., Lagisz, M., Jennions, M.D., Koricheva, J., Noble, D.W.A., Parker, T.H., *et al.* (2021). Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. *Biol. Rev.*, 96, 1695–1722.

Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Syst. Rev.*, 5, 1–10.

Paradis, E. & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526–528.

Parolini, M., Possenti, C.D., Romano, A., Caprioli, M., Rubolini, D. & Saino, N. (2018). Physiological increase of yolk testosterone level does not affect oxidative status and

telomere length in gull hatchlings. *PLoS One*, 13.

Partecke, J., Hegyi, G., Fitze, P.S., Gasparini, J. & Schwabl, H. (2020). Maternal effects and urbanization: Variation of yolk androgens and immunoglobulin in city and forest blackbirds. *Ecol. Evol.*

Pick, J.L., Nakagawa, S. & Noble, D.W.A. (2019). Reproducible, flexible and high-throughput data extraction from primary literature: The metaDigitise r package. *Methods Ecol. Evol.*, 10, 426–431.

Pilz, K.M., Quiroga, M., Schwabl, H. & Adkins-Regan, E. (2004). European starling chicks benefit from high yolk testosterone levels during a drought year. *Horm. Behav.*, 46, 179–192.

Pitala, N., Ruuskanen, S., Laaksonen, T., Doligez, B., Tschirren, B. & Gustafsson, L. (2009). The effects of experimentally manipulated yolk androgens on growth and immune function of male and female nestling collared flycatchers Ficedula albicollis. *J. Avian Biol.*, 40, 225–230.

Podlas, K., Helfenstein, F. & Richner, H. (2013). Brood Reduction via Intra-clutch Variation in Testosterone - An Experimental Test in the Great Tit. *PLoS One*, 8.

Podmokła, E., Drobniak, S.M. & Rutkowska, J. (2018). Chicken or egg? Outcomes of experimental manipulations of maternally transmitted hormones depend on administration method – a meta-analysis. *Biol. Rev.*, 93, 1499–1517.

Poisbleau, M., Müller, W., Carslake, D., Demongin, L., Groothuis, T.G.G., van Camp, J., *et al.* (2012). Penguin chicks benefit from elevated yolk androgen levels under sibling competition. *PLoS One*, 7.

Polich, R.L., Bodensteiner, B.L., Adams, C.I.M. & Janzen, F.J. (2018). Effects of augmented corticosterone in painted turtle eggs on offspring development and behavior. *Physiol. Behav.*, 183, 1–9.

Possenti, C.D., Romano, A., Caprioli, M., Rubolini, D., Spiezio, C., Saino, N., *et al.* (2016). Yolk testosterone affects growth and promotes individual-level consistency in behavioral lateralization of yellow-legged gull chicks. *Horm. Behav.*, 80, 58–67.

Rees, J.A. & Cranston, K. (2017). Automated assembly of a reference taxonomy for phylogenetic data synthesis. *Biodivers. data J.*, e12581.

Reinhold, K. (2002). Maternal Effects and the Evolution of Behavioral and Morphological Characters: A Literature Review Indicates the Importance of Extended Maternal Care. *J. Hered.*, 93, 400–405.

Remeš, V. (2011). Yolk androgens in great tit eggs are related to male attractiveness, breeding

density and territory quality. *Behav. Ecol. Sociobiol.*, 65, 1257–1266.

Rettenbacher, S., Groothuis, T.G., Henriksen, R. & Möstl, E. (2013). Corticosterone in bird eggs: The importance of analytical validation. *Wien. Tierarztl. Monatsschr.*, 100, 283–290.

Rettenbacher, S., Möstl, E. & Groothuis, T.G.G. (2009). Gestagens and glucocorticoids in chicken eggs. *Gen. Comp. Endocrinol.*, 164, 125–129.

Ronget, V., Gaillard, J.-M., Coulson, T., Garratt, M., Gueyffier, F., Lega, J.-C., *et al.* (2018). Causes and consequences of variation in offspring body mass: meta-analyses in birds and mammals. *Biol. Rev.*, 93, 1–27.

Rubolini, D., Romano, M., Boncoraglio, G., Ferrari, R.P., Martinelli, R., Galeotti, P., *et al.* (2005). Effects of elevated egg corticosterone levels on behavior, growth, and immunity of yellow-legged gull (Larus michahellis) chicks. *Horm. Behav.*, 47, 592–605.

Rubolini, D., Romano, M., Martinelli, R. & Saino, N. (2006). Effects of elevated yolk testosterone levels on survival, growth and immunity of male and female yellow-legged gull chicks. *Behav. Ecol. Sociobiol.*, 59, 344–352.

Ruuskanen, S., Darras, V.M., Visser, M.E. & Groothuis, T.G.G. (2016). Effects of experimentally manipulated yolk thyroid hormone levels on offspring development in a wild bird species. *Horm. Behav.*, 81, 38–44.

Ruuskanen, S., Doligez, B., Gustafsson, L. & Laaksonen, T. (2012a). Long-term effects of yolk androgens on phenotype and parental feeding behavior in a wild passerine. *Behav. Ecol. Sociobiol.*, 66, 1201–1211.

Ruuskanen, S., Doligez, B., Pitala, N., Gustafsson, L. & Laaksonen, T. (2012b). Long-term fitness consequences of high yolk androgen levels: Sons pay the costs. *Funct. Ecol.*, 26, 884–894.

Ruuskanen, S., Doligez, B., Tschirren, B., Pitala, N., Gustafsson, L., Groothuis, T.G.G., *et al.* (2009). Yolk androgens do not appear to mediate sexual conflict over parental investment in the collared flycatcher Ficedula albicollis. *Horm. Behav.*, 55, 514–519.

Ruuskanen, S., Groothuis, T.G.G., Baugh, A.T., Schaper, S. V, de Vries, B. & van Oers, K. (2018). Maternal egg hormones in the mating context: The effect of pair personality. *Funct. Ecol.*, 32, 439–449.

Ruuskanen, S. & Laaksonen, T. (2010). Yolk hormones have sex-specific long-term effects on behavior in the pied flycatcher (Ficedula hypoleuca). *Horm. Behav.*, 57, 119–127.

Ruuskanen, S. & Laaksonen, T. (2013). Sex-specific effects of yolk androgens on begging behavior and digestion in pied flycatchers. *J. Avian Biol.*, 44, 331–338.

Sánchez-Tójar, A., Lagisz, M., Moran, N.P., Nakagawa, S., Noble, D.W.A. & Reinhold, K.

(2020). The jury is still out regarding the generality of adaptive ‘transgenerational’ effects. *Ecol. Lett.*, 23, 1715–1718.

Sánchez-Tójar, A., Nakagawa, S., Sánchez-Fortún, M., Martin, D.A., Ramani, S., Girndt, A., *et al.* (2018). Meta-analysis challenges a textbook example of status signalling and demonstrates publication bias. *Elife*, 7, e37385.

Schwabl, H. (1993). Yolk is a source of maternal testosterone for developing birds. *Proc. Natl. Acad. Sci.*, 90, 11446–11450.

Schwabl, H. & Partecke, J. (2020). Developmental programming of the adrenocortical stress response by yolk testosterone depends on sex and life history stage. *J. Exp. Biol.*, 223.

Sockman, K.W. & Schwabl, H. (2000). Yolk androgens reduce offspring survival. *Proc. R. Soc. B Biol. Sci.*, 267, 1451–1456.

Sockman, K.W., Weiss, J., Webster, M.S., Talbott, V. & Schwabl, H. (2008). Sex-specific effects of yolk-androgens on growth of nestling American kestrels. *Behav. Ecol. Sociobiol.*, 62, 617–625.

Stier, K.S., Almasi, B., Gasparini, J., Piault, R., Roulin, A. & Jenni, L. (2009). Effects of corticosterone on innate and humoral immune functions and oxidative stress in barn owl nestlings. *J. Exp. Biol.*, 212, 2085–91.

Strange, M.S., Bowden, R.M., Thompson, C.F. & Sakaluk, S.K. (2016). Pre- and Postnatal Effects of Corticosterone on Fitness-Related Traits and the Timing of Endogenous Corticosterone Production in a Songbird. *J. Exp. Zool. Part A Ecol. Genet. Physiol.*, 325, 347–359.

Strasser, R. & Schwabl, H. (2004). Yolk testosterone organizes behavior and male plumage coloration in house sparrows (Passer domesticus). *Behav. Ecol. Sociobiol.*, 56, 491–497.

Tilgar, V., Mägi, M., Lind, M., Lodjak, J., Moks, K. & Mänd, R. (2016). Acute embryonic exposure to corticosterone alters physiology, behaviour and growth in nestlings of a wild passerine. *Horm. Behav.*, 84, 111–120.

Treidel, L.A., Whitley, B.N., Benowitz-Fredericks, Z.M. & Haussmann, M.F. (2013). Prenatal exposure to testosterone impairs oxidative damage repair efficiency in the domestic chicken (Gallus gallus). *Biol. Lett.*, 9.

Tschirren, B. (2015). Differential effects of maternal yolk androgens on male and female offspring: A role for sex-specific selection? *PLoS One*, 10.

Tschirren, B., Fitze, P.S. & Richner, H. (2007). Maternal modulation of natal dispersal in a passerine bird: An adaptive strategy to cope with parasitism? *Am. Nat.*, 169, 87–93.

Tschirren, B., Postma, E., Gustafsson, L., Groothuis, T.G.G. & Doligez, B. (2014). Natural

selection acts in opposite ways on correlated hormonal mediators of prenatal maternal effects in a wild bird population. *Ecol. Lett.*, 17, 1310–5.

Tschirren, B., Saladin, V., Fitze, P.S., Schwabl, H. & Richner, H. (2005). Maternal yolk testosterone does not modulate parasite susceptibility or immune function in great tit nestlings. *J. Anim. Ecol.*, 74, 675–682.

Uller, T. (2008). Developmental plasticity and the evolution of parental effects. *Trends Ecol. Evol.*, 23, 432–438.

Uller, T. (2012). The Evolution of Parental Care. In: *The Evolution of Parental Care* (ed. Royle NJ, Smiseth PT, K.M.). New York: Oxford University Press, pp. 247–267.

Valcu, C.M., Scheltema, R.A., Schweiggert, R.M., Valcu, M., Teltscher, K., Walther, D.M., *et al.* (2019). Life history shapes variation in egg composition in the blue tit Cyanistes caeruleus. *Commun. Biol.*, 2, 1–14.

Vassallo, B.G., Paitz, R.T., Fasanello, V.J. & Haussmann, M.F. (2014). Glucocorticoid metabolism in the in ovo environment modulates exposure to maternal corticosterone in Japanese quail embryos (Coturnix japonica). *Biol. Lett.*, 10.

Verboven, N., Monaghan, P., Evans, D.M., Schwabl, H., Evans, N., Whitelaw, C., *et al.* (2003). Maternal condition, yolk androgens and offspring performance: A supplemental feeding experiment in the lesser black-backed gull (Larus fuscus). *Proc. R. Soc. B Biol. Sci.*, 270, 2223–2232.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *J. Stat. Softw.*, 36, 1–48.

Weber, B.M., Bowers, E.K., Terrell, K.A., Falcone, J.F., Thompson, C.F. & Sakaluk, S.K. (2018). Pre- and postnatal effects of experimentally manipulated maternal corticosterone on growth, stress reactivity and survival of nestling house wrens. *Funct. Ecol.*, 32, 1995–2007.

Weiss, S.L., Kennedy, E.A., Safran, R.J. & Mcgraw, K.J. (2011). Pterin-based ornamental coloration predicts yolk antioxidant levels in female striped plateau lizards (Sceloporus virgatus). *J. Anim. Ecol.*, 80, 519–527.

Westgate, M.J. (2018). revtools: Tools to support evidence synthesis (0.3.0) [Computer software]., https://CRAN.R-project.org/package=revtools.

Williams, T.D. & Groothuis, T. (2015). *Egg quality, embryonic development, and post-hatching phenotype: an integrated perspective*.

Wolf, J.B., Leamy, L.J., Roseman, C.C. & Cheverud, J.M. (2011). Disentangling prenatal and postnatal maternal genetic effects reveals persistent prenatal effects on offspring growth in mice. *Genetics*, 189, 1069–1082.

Wolf, J.B. & Wade, M.J. (2001). On the assignment of fitness to parents and offspring: whose fitness is it and when does it matter? *J. Evol. Biol.*, 14, 347–356.

Wolf, J.B. & Wade, M.J. (2016). Evolutionary genetics of maternal effects. *Evolution (N. Y).*, 70, 827–839.

Yang, Y., Noble, D.W.A., Senior, A.M., Lagisz, M. & Nakagawa, S. (2024). Interpreting prediction intervals and distributions for decoding biological generality in meta-analyses. *bioRxiv*.

Yin, J., Zhou, M., Lin, Z., Li, Q.Q. & Zhang, Y. (2019). Transgenerational effects benefit offspring across diverse environments: a meta-analysis in plants and animals. *Ecol. Lett.*, 22, 1976–1986.