Facultative adjustment of the offspring sex ratio and male attractiveness: a systematic review and meta-analysis

Isobel Booksmythe[1,2,3,*], Brian Mautz[1,2], Jacqueline Davis[1,4], Shinichi Nakagawa[5,6] and Michael D. Jennions[1]

[1]*Division of Evolution, Ecology & Genetics, Research School of Biology, The Australian National University, Daley road, Canberra, Australian Capital Territory 2601, Australia*
[2]*Department of Animal Ecology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-75236, Uppsala, Sweden*
[3]*Institute of Evolutionary Biology and Environmental Studies, University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland*
[4]*Department of Psychology, University of Cambridge, Downing Street, CB2 3EB, Cambridge, U.K.*
[5]*Department of Zoology, University of Otago, PO Box 56, Dunedin 9054, New Zealand*
[6]*Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia*

ABSTRACT

Females can benefit from mate choice for male traits (e.g. sexual ornaments or body condition) that reliably signal the effect that mating will have on mean offspring fitness. These male-derived benefits can be due to material and/or genetic effects. The latter include an increase in the attractiveness, hence likely mating success, of sons. Females can potentially enhance any sex-biased benefits of mating with certain males by adjusting the offspring sex ratio depending on their mate's phenotype. One hypothesis is that females should produce mainly sons when mating with more attractive or higher quality males. Here we perform a meta-analysis of the empirical literature that has accumulated to test this hypothesis. The mean effect size was small ($r = 0.064$–0.095; i.e. explaining <1% of variation in offspring sex ratios) but statistically significant in the predicted direction. It was, however, not robust to correction for an apparent publication bias towards significantly positive results. We also examined the strength of the relationship using different indices of male attractiveness/quality that have been invoked by researchers (ornaments, behavioural displays, female preference scores, body condition, male age, body size, and whether a male is a within-pair or extra-pair mate). Only ornamentation and body size significantly predicted the proportion of sons produced. We obtained similar results regardless of whether we ran a standard random-effects meta-analysis, or a multi-level, Bayesian model that included a correction for phylogenetic non-independence. A moderate proportion of the variance in effect sizes (51.6–56.2%) was due to variation that was not attributable to sampling error (i.e. sample size). Much of this non-sampling error variance was not attributable to phylogenetic effects or high repeatability of effect sizes among species. It was approximately equally attributable to differences (occurring for unknown reasons) in effect sizes among and within studies (25.3, 22.9% of the total variance). There were no significant effects of year of publication or two aspects of study design (experimental/observational or field/laboratory) on reported effect sizes. We discuss various practical reasons and theoretical arguments as to why small effect sizes should be expected, and why there might be relatively high variation among studies. Currently, there are no species where replicated, experimental studies show that mothers adjust the offspring sex ratio in response to a generally preferred male phenotype. Ultimately, we need more experimental studies that test directly whether females produce more sons when mated to relatively more attractive males, *and* that provide the requisite evidence that their sons have higher mean fitness than their daughters.

Key words: female choice, mate choice, maternal effects, sex allocation, sex ratio adjustment, sexual selection.

*Address for correspondence (Tel: +41 044 6354804; E-mail: isobel.booksmythe@ieu.uzh.ch).

CONTENTS

I. INTRODUCTION

(1) Environmental condition-dependent sex allocation

The birth sex ratio in many species deviates from unity, despite selection on parents to produce equal numbers of sons and daughters due to negative frequency-dependent selection (Düsing, 1884; Fisher, 1930). Although the reproductive value of individuals of the rare sex is always greater than that of the other sex, thereby tending to select for equal investment into each sex, this does not constrain the offspring sex ratio to be even. First, if one sex is cheaper to produce and total sex allocation is equal, more offspring of the cheaper sex will be produced (West, 2009). Second, adaptive adjustment of sex allocation away from equal investment into both sexes is favoured by selection if the function relating investment to fitness returns differs for sons and daughters (i.e. if it is non-linear for at least one sex: Charnov, 1979) even if there is no variation in the resources mothers can allocate to reproduction (reviews: Frank, 1990, 1998, pp. 174–178). Third, facultative parental control of the offspring sex ratio is potentially advantageous when offspring fitness depends on some aspect of the breeding environment, which varies among parents (e.g. territory quality, female condition, adult size), if this condition dependence differently affects the absolute reproductive value of sons and daughters (Trivers & Willard, 1973). This is likely, as many environmental factors have sex-specific effects on offspring fitness (e.g. Jones,

Nakagawa & Sheldon, 2009; Bowers, Thompson & Sakaluk, 2014). More generally, it is clear that males and females tend to have different fitness functions (review: Connallon & Clark, 2014). Given variation among parents in the quality of their breeding environment, those breeding in a relatively better environment should increase allocation towards the sex that has greater reproductive value when investment is higher. One way to do this is to bias the offspring sex ratio towards that sex.

Trivers & Willard (1973) originally suggested that differences in maternal condition could lead to facultative sex ratio adjustment (henceforth, SRA) by parents. They noted that in polygynous mammals (e.g. many ungulates and pinnipeds) there is enormous variation in male mating success, and that larger, healthier males are disproportionately successful at acquiring mates when males compete. By contrast, female size has weak effects on fecundity (in many mammals the range in litter size is very low). In general, mothers in good physical condition have the ability to produce larger, healthier offspring. Trivers & Willard (1973) therefore argued that mothers in good condition should preferentially produce sons because of the relatively greater fitness gains they accrue from improving their sons' body condition. There is empirical support for this claim from studies of ungulates (meta-analysis: Sheldon & West, 2004).

Hypotheses of condition-dependent SRA have since been broadened to include many other environmental and social factors that affect 'breeding conditions' and generate

predictable differences in the relative fitness of sons and daughters (West, 2009, chapter 6). These scenarios include: production of daughters by high-ranking mothers when social rank is maternally inherited, as in primates (Schino, 2004); production of daughters by parasitoid insects laying eggs on larger hosts because of the stronger effect of size on female fecundity than on male mating success (meta-analysis: West & Sheldon, 2002); production of the helping sex if parents lack sufficient helpers-at-the-nest in cooperatively breeding birds and mammals (meta-analysis: West & Sheldon, 2002; but see Stubblefield & Orzack, 2013); seasonal or laying order effects if one sex requires more resources as a neonate, and food availability changes seasonally (e.g. for work on raptors see Rutz, 2012 and references therein). More generally, many factors have been posited to affect the optimal offspring sex ratio because they predict future resource availability for investment into offspring (e.g. territory quality, harem status; for a more complete list see West, 2009, p. 183).

(2) Genetic condition-dependent sex ratio adjustment

There is an important condition-dependent sex allocation scenario that is not based on parental ability to predict environmentally determined resource availability. It is the argument that a male's attractiveness or quality can differentially affect the fitness of his sons and daughters because of genetic effects (Burley, 1981, 1986). Males with larger ornaments and faster display rates, those in better body condition and/or those of larger body size are often preferred by females, gaining above-average mating success (Andersson, 1994). There is strong evidence for moderate heritability of sexually selected, preferred male traits (Prokuda & Roff, 2014; Wyman & Rowe, 2014). Burley (1981, 1986) therefore predicted that females mating with attractive males should preferentially produce sons because sons will benefit more than daughters by inheriting genes that increase their sexual attractiveness. Burley's idea is compelling because few traits are more intimately related to reproductive success and, by extension, to fitness, than those under directional sexual selection (Shuster & Wade, 2003). There are also plausible grounds for assuming that some of the benefits of mating with attractive males are sex-specific, or sex-biased, accruing only or mainly to sons. Still, there are several important caveats.

Preferred male traits can signal both material and genetic benefits (e.g. Hoelzer, 1989; meta-analysis: Møller & Jennions, 2001). If there are direct, fecundity-enhancing benefits of female choice due to material gains, then additional maternal resources might benefit daughters as much as, or even more than, sons (e.g. if the fecundity gains from increasing female size outweigh those of higher mating success for larger males or males who can invest more into costly ornaments). The adaptive prediction that females will produce sons when mated to attractive males implicitly assumes there are indirectly selected, genetic benefits of mate choice. Furthermore, for genetic benefits to favour SRA towards sons, the genes inherited from attractive sires must disproportionately advantage males (Burley, 1986). This is intuitively the case for arbitrarily attractive (Fisherian) male traits that are not inherited by daughters. The inheritance of genes that increase viability (so-called 'good genes') must, however, have a stronger effect on males than females to favour a male-biased offspring sex ratio. This might occur, for example, if sons benefit more than daughters from improved body condition because many sexual ornaments have highly condition-dependent expression (Cotton, Fowler & Pomiankowski, 2004). The benefits of an attractive father can also be greater for sons than daughters if attractiveness is similarly heritable for sons and daughters (i.e. both sexes are under sexual selection due to mate choice), but males experience greater variance in reproductive success (e.g. Du & Lu, 2010; Bowers et al., 2013), which is causally related to mate choice (i.e. sexual selection is stronger on males than females). In general, selection will favour a conditional sex allocation strategy when ornamentation imparts heritable benefits (be they due to the ornaments themselves or traits genetically correlated with ornamentation) that are greater for offspring of one sex than for the other (Fawcett et al., 2011).

If male-biased genetic benefits of mate choice exist, they can be enhanced by biasing the offspring sex ratio towards sons when mating with more attractive males, and towards daughters when mating with less attractive males (Burley, 1981, 1986). This form of conditional SRA has several potential fitness benefits for mothers that are driven by genetic effects. (i) It increases the production of 'sexy sons' (Weatherhead & Robertson, 1979; Firman, 2011; Prokop et al., 2012). There is, however, an implicit assumption that the benefits of a higher mating rate are not perfectly countered by greater naturally selected costs (e.g. higher mortality) (see Pen & Weissing, 2000), or countered by a trade-off with other sexually selected traits (e.g. reduced sperm competiveness; see Lüpold et al., 2014). (ii) Producing fewer daughters reduces the negative effects of sexually antagonistic genes when mated to an attractive male (if genes from more attractive sires make daughters of below-average fitness) (Rice & Chippindale, 2001; review: Rice, 2013). (iii) Even in the absence of any such sexual antagonism, producing more daughters lessens the disadvantage of mating with an unattractive male as the value of daughters increases due to their 'underproduction' by females mated to attractive mates. Consequently, the fitness difference between females mated to attractive and unattractive mates declines (Fawcett et al., 2007, 2011).

(3) Models of sex ratio adjustment (SRA) in response to mate attractiveness

Greater production of sons due to genes that confer male-specific fitness benefits is distinct from most other conditional sex allocation scenarios. In the latter, the factors that favour biased sex ratios (e.g. maternal condition, habitat quality) tend to do so because they predict the availability of resources that can be invested into offspring, and the sexes differ in the effect of investment on their reproductive value

(Frank, 1998). Consequently, most theoretical treatments of conditional sex allocation are not readily applicable to the mate attractiveness hypothesis. This is largely because covariances between genes for different traits (e.g. ornaments, preferences and sex allocation strategies), and changes in genetic variances have to be taken into account when modelling genetic benefits. Crucially, selection on condition-dependent SRA does not obviously affect variation in parental resource availability that favours SRA, and the sex-specific functions relating reproductive value to environmental quality remain the same. By contrast, selection on SRA in response to additive genetic variation in attractiveness could lead to a co-evolutionary feedback that changes the fitness benefits of being more attractive because it will affect phenotypic variation in male attractiveness and female mating preferences (Fawcett *et al.*, 2011).

To date, only four mathematical models have addressed the evolution of SRA in response to male attractiveness. Pen & Weissing (2000) provided the first formal model. They used a game theory approach with a discrete male trait (i.e. ornamented *versus* unornamented males) to show that an evolutionarily stable strategy for male-biased SRA when mated to ornamented males could evolve. The stability of such a strategy was, however, sensitive to the form of selection driving male trait exaggeration. Facultative SRA evolved under 'good genes' selection where the male trait signalled heritable viability (albeit in a simplified scenario where 'viability' genes did not benefit daughters). It did not evolve under a pure Fisherian process. In the latter case, all males are of equal fitness at equilibrium as natural selection perfectly balances sexual selection (ornamented males have greater mating success but higher mortality). There was therefore no fitness benefit from SRA towards sons when mated to an attractive male in this model.

Subsequently Fawcett *et al.* (2007) ran individual-based simulations that confirmed these initial results. Facultative SRA did not evolve when the male ornamental trait was purely Fisherian, but did evolve when it was a conditional indicator of viability, and when it was a Fisherian trait subject to a mutation bias (*sensu* Pomiankowski, Iwasa & Nee, 1991). Reassuringly, the results did not differ between two different proximate mechanisms of SRA. The findings were, however, slightly different if the male trait exhibited continuous rather than discrete variation. For a continuous male trait, the only scenario where facultative SRA unambiguously evolved was for a Fisherian trait subject to mutation bias. Although there was still a shift in sex allocation when the male trait was a conditional indicator, it was weak and inconsistent across simulation runs. The evolution of a stronger sex ratio bias for a Fisherian trait makes intuitive sense because such traits only benefit males. By contrast, conditional indicators signal the presence of viability genes that increase the fitness of both sexes. An additional, clear finding of Fawcett *et al.* (2007) was that the evolution of SRA in response to male attractiveness was in all cases very slow, suggesting that it is under very weak selection. This is readily seen by comparing the rate of evolution of facultative SRA to the rate of change in the female preference and expression of preferred male traits in the simulations (Fawcett *et al.*, 2007). Selection in the wild is likely to be even weaker than reported in the models, because (*i*) the models ignore the additional costs of actually adjusting the sex ratio (e.g. energetic costs if this requires strategic reabsorption of embryos, or time costs of delaying fertilization; Krackow, 1995); and (*ii*) in practice, females obtain imperfect information on male genotype due to both non-heritable environmental influences on male sexual trait expression (Holman & Kokko, 2013), and female perceptual errors when assessing the expression of sexual signals (e.g. Farris & Ryan, 2011).

More recently, Fawcett *et al.* (2011) noted that facultative SRA in response to heritable sire attractiveness reduces sexual selection on male traits. A co-evolutionary feedback occurs because females mated to less attractive males 'compensate' by producing daughters. This reduces the fitness difference between females mated to more and less attractive males. The strength of selection for costly female preferences for males with more elaborate sexual traits is therefore weakened. In addition, sexual selection declines because the mean strength of the female preference in the population is lower when choosier females (those mating with more attractive males) produce fewer daughters, because only daughters express mating preferences. Both processes select for reduced investment into costly sexual traits by males. Facultative SRA effectively reduces the heritability of attractiveness and mating preferences because of the lower fitness gain of ornamentation (Fawcett *et al.*, 2011). Fawcett *et al.* (2011) therefore conclude that species with elaborate sexual traits are the least likely to show facultative SRA in response to mate attractiveness. This is worrisome because empirical researchers testing for adaptive SRA have sought out species with strong sexual selection on males.

Booksmythe, Schwanz & Kokko (2013) suggested that the feedback between sexual selection and SRA based on attractiveness might be even more complicated than indicated by Fawcett *et al.* (2011). Specifically, as the expression of a preferred male trait declines, it is increasingly difficult for females to distinguish and discriminate between males due to perceptual errors while mate sampling. More generally, imperfect maternal control of sex allocation is expected. They also noted that given costs of adjusting the sex ratio, this strategy might be eliminated when male trait expression is low. In an extension of Fawcett *et al.*'s (2011) model they allowed the 'accuracy' of sex allocation to evolve (i.e. the extent to which there is a 'step-like' transition between the production of daughters and sons; see West, Reece & Sheldon, 2002, for comments on empirical evidence that sex allocation is imprecise and not 'step-like'). They also imposed a cost on females adjusting the sex ratio that increased with the precision of sex allocation. The simulation results were complex, suggesting there is not a simple linear relationship between the strength of SRA and male trait expression. In general, the greatest degree of SRA arose when sexual selection on male traits was intermediate in strength.
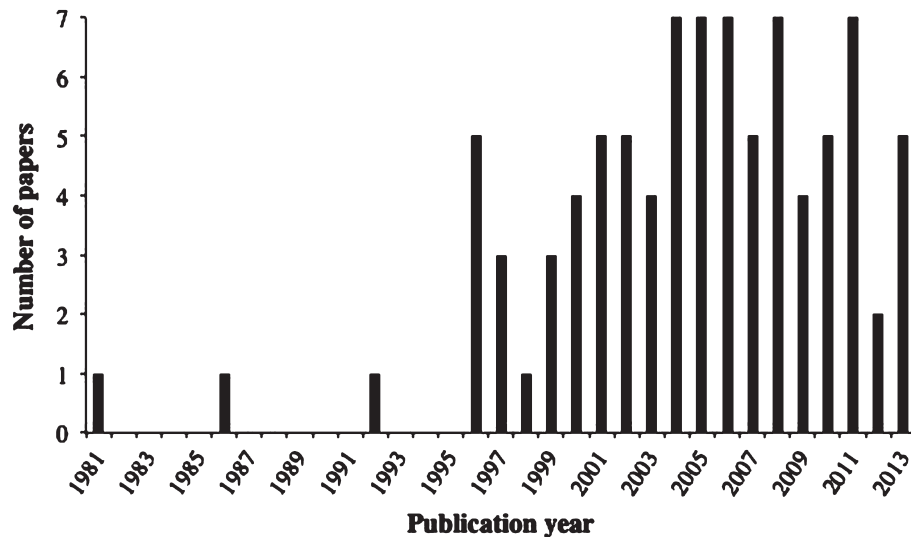
Fig. 1. Papers published per year since Burley (1981) on adjustment of offspring sex ratios in response to paternal traits. Only papers eligible for inclusion in our meta-analyses are shown; four included papers from 2014 are not shown as our search only encompassed half of this year.

(4) The empirical evidence

Many empirical studies testing Burley's original idea have accumulated in the last three decades (Fig. 1), but the available evidence for adjustment of the offspring sex ratio in response to mate attractiveness is still mixed. An early meta-analysis found a strong positive correlation between the proportion of sons in a brood and male attractiveness based on 11 studies of eight avian species ($r = 0.205$; West & Sheldon, 2002). This result seemed to spur a flurry of further research, but the extent to which the hypothesis is currently supported is less clear. Several studies have found a significant positive relationship between some measure of male attractiveness and the proportion of sons in a brood (e.g. Saino *et al.*, 2002; Sato & Karino, 2010; Correa *et al.*, 2011; Bowers *et al.*, 2013), others have found no such relationship (Ramsay *et al.*, 2003; Rutstein *et al.*, 2004; Parker, 2005; Limbourg, Mateman & Lessells, 2013), and still others have reported a significant relationship in the opposite direction (Rutstein *et al.*, 2005; Delhey *et al.*, 2007). In addition, studies over multiple years or from different populations of the same species have reported that the relationship varies in strength and/or direction (e.g. Griffith *et al.*, 2003; Rosivall *et al.*, 2004; see also Parker, 2013, for a meta-analysis of studies on blue tits). The extent to which intra-specific variation is due to sampling error or genuine biological differences is unknown.

Ultimately, the diverse empirical results obtained to date reflect the challenge of identifying male traits that: (*i*) have a sound theoretical reason to affect the offspring sex ratio (rather than other forms of differential allocation such as preferential provisioning of eggs or juveniles of one sex; see Horváthová, Nakagawa & Uller, 2012), (*ii*) are associated with male-biased genetic benefits, and (*iii*) are applicable to the study population/species. In combination with the added difficulties of controlling for confounding variables,

dealing with measurement error, and recurrent problems of low statistical power in behavioural ecology (Jennions & Møller, 2003) it is encouraging that any studies have obtained a significant result in the predicted direction. Even so, determining whether this has occurred more often than expected by chance remains a major challenge.

(5) The role of meta-analysis

Qualitative, narrative reviews have reported mixed evidence for adaptive sex allocation in birds and mammals, regardless of the source of selection on the offspring sex ratio (reviews: Clutton-Brock, 1986; Sheldon, 1998; Palmer, 2000; West, Herre & Sheldon, 2000; Komdeur & Pen, 2002; West *et al.*, 2002). By contrast, formal meta-analyses of hypotheses related to sex allocation in birds, mammals and insects have produced more clear-cut conclusions. Several meta-analyses have reported that the average relationship is significant in the predicted direction [e.g. helpers-at-the-nest: West & Sheldon, 2002; Griffin, Sheldon & West, 2005; but see Stubblefield & Orzack, 2013; mate attractiveness in birds: West & Sheldon, 2002; but see Parker, 2013; maternal condition in ungulates: Sheldon & West, 2004; host size in parasitoids: West & Sheldon, 2002; mate quality, laying date and season in birds: Cassey, Ewen & Møller, 2006 (correction of Ewen, Cassey & Møller, 2004); local mate competition: West, Shuker & Sheldon, 2005; split sex ratios in social insects: Meunier, West & Chapuisat, 2008; local resource competition in primates: Silk & Brown, 2008]. However, other meta-analyses, often using similar data sets, have reported that the mean relationship (effect size) is not significantly different from the null expectation (zero) (e.g. maternal rank in primates: Brown & Silk, 2002; Schino, 2004; Silk, Willoughby & Brown, 2005; maternal condition in mammals: Cameron, 2004). Meta-analysis is a

statistical technique designed to detect broad-scale patterns across empirical studies and quantitatively weigh studies by their accuracy in reporting the focal relationship. In short, estimates of a relationship are made with clearly stated confidence intervals. By contrast, narrative reviews are susceptible to vote counting (how many studies obtained a 'yes' or 'no' answer), which ignores the limitations of low power that plague most studies in evolutionary ecology. Meta-analysis has the additional advantage that it provides tools to gauge the extent to which moderating factors are associated with the observed relationships, and tools to assess the likely influence of publication biases (Jennions *et al.*, 2012). For example, although the mean effect size did not differ from zero in Ewen *et al.* (2004), moderator variables had a significant influence on offspring sex ratios (Cassey *et al.*, 2006). This suggests that, under certain conditions, there is still evidence that parents can skew sex ratios.

As new empirical data accumulate, meta-analyses need to be revised. This is especially important in fields where effect sizes tend to decline over time (see Jennions & Møller, 2002). For example, Pereira & Ioannidis (2011) have recently estimated – based on a decline in the reported mean effect sizes between meta-analyses conducted in 2005 and 2010 – that 37% of medical meta-analyses are likely to conclude falsely that there is evidence for a relationship. In addition, it is obvious to anyone who has conducted a meta-analysis that subjective decisions are made (e.g. how to extract effect sizes, how to search for papers). This can lead to researchers reporting different findings from meta-analyses on the same question, even when they use similar data sets. For a good example, see a recent dispute about the effect of the menstrual cycle on human mate choice between Wood *et al.* (2014) and Gildersleeve, Haselton & Fales (2014). For a topic closer to the current issue one can compare the findings of two meta-analyses on SRA in relation to helper number by Griffin *et al.* (2005) and Stubblefield & Orzack (2013).

The most recent meta-analysis of Burley's (1981) hypothesis was by West & Sheldon (2002). Since then, there have been many studies with data on SRA in relation to mate attractiveness/quality (Fig. 1). The accumulation of studies allows for a more robust test of whether there is a general trend, which seems especially relevant in the light of recent theory questioning the extent to which there will be detectable selection for SRA based on male attractiveness (Fawcett *et al.*, 2011). Here we use both standard meta-analytical and newer meta-regression approaches (including a control for phylogenetic relationships) to: (*i*) test whether there is a general trend for the offspring sex ratio to be male-biased when females are mated to attractive males. (*ii*) Investigate if the relationship between the offspring sex ratio and male attractiveness depends on the type of male trait or measurement used as a proxy for attractiveness. For example, are ornaments better predictors of SRA than behaviour? (*iii*) Test whether phylogenetic effects and differences in study design are associated with the strength of the relationship between the offspring sex ratio and male attractiveness.

II. METHODS

(1) Literature search

We used two complementary approaches to identify as many relevant studies as possible. First, we performed a key word search using the *ISI Web of Knowledge* search engine (for details of the databases covered see online Appendix S1). We cross-referenced a set of key words related to sex ratio (e.g. 'offspring sex ratio', 'proportion male') to two sets of key words relating to male attractiveness: (*i*) general terms (e.g. attractive, ornament, 'secondary sexual trait') and (*ii*) numerous specific secondary sexual characteristics (e.g. plumage, antler, horn, badge). The exact search term combinations used are given in Appendix S1. Second, we identified six empirical and theoretical papers that are highly cited in studies of sex allocation in response to mate attractiveness: Burley (1981), Ellegren, Gustafsson & Sheldon (1996), Frank (1990), Fawcett *et al.* (2007), Sheldon *et al.* (1999) and West & Sheldon (2002). At least one of these papers is likely to be cited by a study testing for SRA in response to male attractiveness. We performed a forward search to obtain papers citing any of these works. In total, the two search approaches yielded 4474 unique references (see online Appendix S2). To finalize the data-checking and analysis protocols we conducted our final search on 17 July 2014. We did not add papers encountered by other means because these are more likely to be drawn from higher impact journals that might preferentially publish studies with significant results (Murtaugh, 2002). We also did not solicit unpublished data sets from colleagues. This procedure reduces the risk of biasing our estimates of effect sizes (see Jennions *et al.*, 2013). As with an empirical study, our goal was to obtain an extensive, representative sample of studies rather than to locate every available data point.

The criteria we used to decide whether a study contained potentially usable data are provided in Appendix S1. In brief, we excluded studies of plants, humans and farmed animals. We looked for studies where there were likely to be data on both brood/litter offspring sex ratios (i.e. not just population-level ratios) and measurements of male traits that are generally assumed to be targets of female choice in many taxa (e.g. body condition, body size and, most importantly, male ornamentation).

(2) Data extraction and coding

We collected and transformed study results into a standard effect size measure ($\mathcal{Z}_r$; see below) to allow for comparisons between studies. Reported effects were used in the following order of preference: (*i*) direct reports of an effect size (e.g. correlations) ($N = 13$); (*ii*) effect measures with magnitude and direction (e.g. regression coefficients and standard errors, standardized mean differences, or descriptive statistics) ($N = 79$); (*iii*) raw numbers (e.g. contingency tables) ($N = 12$); (*iv*) inferential (test) statistics (often from multiple regressions) (e.g. t, z, F and χ^2 values, P values) ($N = 161$). In some cases we could only note that there was a non-significant effect

($N = 11$). All data were converted to correlation coefficients r, and then to Z_r, following the formulae listed in Lipsey & Wilson (2001) and Nakagawa & Cuthill (2007). The variance of Z_r is $1/(N - 3)$ (Lipsey & Wilson, 2001). We calculated separate effect sizes for each measure of attractiveness from every independent population and/or experiment (i.e. each unique group of parental individuals; hereafter referred to as a 'study'). Thus, if a paper tested the relationship between a male ornament and offspring sex ratio in two distinct populations, these were considered independent studies and the paper contributed two effect sizes to the meta-analysis. In some cases, where data were analysed by year, each year could reasonably be considered an independent sample of individuals and was treated as a separate study (e.g. Griffith *et al.*, 2003).

The included studies defined and measured male attractiveness in different ways. Namely, (*i*) researchers chose focal traits based on evidence from previous research that particular morphological or behavioural traits influenced male success in attracting and/or mating with females; (*ii*) researchers assumed that sexually dimorphic traits reflect sexual selection on males such that the trait will affect male reproductive success; (*iii*) researchers based their hypothesis on the general prediction that extra-pair males are more attractive than average; (*iv*) in a few studies researchers directly measured female preferences (e.g. association time with different males) to score or rank test males. We were broad in our criteria for what made an 'attractive' trait acceptable for inclusion as some studies cited strong evidence for the effect on male mating success of a certain trait (e.g. experimental manipulation of the trait affected female mate choice), while others would include this trait precisely because it had been shown to influence sex ratios in other species. To test whether relationships varied across different attractiveness proxies, including broad classes of male phenotypic traits, we grouped them into seven categories: Ornament, Behaviour, Size, Condition, Preference, Age, and Extra-pair paternity (EPP) (see Table 1). Each effect size was assigned to one of these categories. 'Ornament' included measures of the size or degree of elaboration of secondary sexual traits; for example, colour measures or the size of a particular body part such as a comb, badge or elongated tail feather. Studies that had manipulated male morphology (e.g. by adding coloured leg bands, or lengthening or shortening tail feathers) were also included in this category. Similarly, 'Behaviour' included any measures of male sexual behaviour that had been identified in the studies as attractive to females, such as song or call attributes, courtship displays, and nest decoration collected by males. 'Size' included measures of structural body size and body mass. 'Condition' included measures of moult timing, parasite infection, and the residuals of mass on body size. 'Preference' measured attractiveness using female preference scores for particular males (e.g. association time in choice tests). Measures of male age or breeding experience were grouped into the category 'Age'. Finally, EPP studies compared the sex ratio of within-pair and extra-pair offspring either within broods (extra-pair

young *versus* within-pair young), between broods (i.e. broods with *versus* without EPP), or across broods (extra-pair and within-pair young pooled across all broods) (see Table 1).

Studies could include multiple effect size estimates due to measurement of several male traits from the same category (e.g. two different ornaments) and/or from different categories (e.g. an ornament and a measure of condition). The presence of more than one effect size per study presents potential problems of non-independence of data for analyses. We dealt with this issue by running the meta-analysis in two ways. First, we used standard random-effects meta-analytic models, which require one effect size estimate per study (Borenstein *et al.*, 2009). Here we calculated a mean effect size for each study by conducting a within-study meta-analysis (i.e. weighting the estimates by the sample size, as this occasionally differed even within a study). This was done separately for analyses restricted to a single trait type, or by pooling effect sizes across all trait types for analyses of 'All' traits. Second, we ran multi-level, random-effects meta-analytic models, which allowed for the inclusion of multiple, non-independent effect size estimates per study by including the levels study and species (Nakagawa & Santos, 2012; Mengersen, Jennions & Schmid, 2013). In these analyses individual effect sizes were used and we included the trait type as a moderator.

The effect size calculated was the Fisher-transformed correlation coefficient, Z_r. By convention, researchers refer to $r = 0.1$, 0.3 and 0.5 as small, medium and large effect sizes (Cohen, 1988). We present Z_r in our results, which is almost identical to r at low values (e.g. 0.1003, 0.310 and 0.549 are the equivalent Z_r benchmarks). The effect size was coded as positive if an increase in the male attractiveness measure (i.e. greater ornamentation, more courtship, larger size, better condition, stronger preference score) was associated with a greater proportion of male offspring. We assumed that older males were more attractive as this is a general assumption in bird studies (Møller & Ninni, 1998; Brooks & Kemp, 2001), and that extra-pair males were more attractive than within-pair males (Møller & Birkhead, 1994; but see Dunn, Whittingham & Pitcher, 2001; reviews: Westneat & Stewart, 2003; Forstmeier *et al.*, 2014). The variance in Z_r was calculated based on the number of broods in the study. We used the number of broods rather than unique breeding pairs as the sample size because this was the information most widely available across studies. Using different sources of N among studies to weight them is problematic because it gives greater weighting to studies that used lower-level analyses (e.g. an analysis that uses $N = $ chicks). Where necessary, however, effect sizes were calculated using the source of N that was appropriate for the inferential statistic the authors presented. Notably, when we converted t or z values from mixed-effect models, we used $N = $ broods rather than $N = $ chicks (which potentially inflates the effect size; *cf*. Nakagawa & Cuthill, 2007; see also Section IV). We also coded information about the study that might have influenced the effect size, for use in a moderator analysis: study species, year of publication (*z*-transformed

Table 1. Summary of 93 papers coded for the meta-analyses

References	Species	N studies	Measures of attractiveness tested	N broods
Abroe *et al.* (2007)	Common yellowthroat, *Geothlypis trichas*	1	EPP: within broods; ornament: mask size, bib size, bib colour; size	69–138
Addison *et al.* (2008)	Rhinoceros auklet, *Cerorhinca monocerata*	1	Size	25
Aparicio & Cordero (2001)	Lesser kestrel, *Falco naumanni*	1	Condition; ornament: wing patch size	75
Bell, Owens & Lord (2014)	Blue tit, *Cyanistes caeruleus*	2	Condition	37, 41
Benito *et al.* (2013)	Common tern, *Sterna hirundo*	1	Behaviour: arrival date; size	81
Blanchard *et al.* (2007)	Wandering albatross, *Diomedea exulans*	1	Condition; size	63
Bowers *et al.* (2013)	Collared flycatcher, *Ficedula albicollis*	1	Age; EPP: across all broods; ornament: wing patch size, forehead patch size	160
Bradbury *et al.* (1997)	European starling, *Sturnus vulgaris*	1	Age; size	108
Burley (1981)	Zebra finch, *Taeniopygia guttata*	1	Ornament: leg band colour	20
Burley (1986)	Zebra finch, *Taeniopygia guttata*	1	Ornament: leg band colour	18
Calsbeek & Sinervo (2004)	Side-blotched lizard, *Uta stansburiana*	2	Size	33, 41
Cockburn & Double (2008)	Superb fairy wren, *Malurus cyaneus*	1	Condition; EPP: across all broods	993
Correa *et al.* (2011)	Japanese quail, *Coturnix japonica*	1	Behaviour: strutting, neck holding, mounting	26
Cotton *et al.* (2014)	Stalk-eyed fly, *Teleopsis dalmanni*	1	Ornament: eyespan, relative eyespan; size	130
Cox *et al.* (2011)	Brown anole, *Anolis sagrei*	1	Condition	67
Cunningham & Russell (2001)	Mallard, *Anas platyrhynchos*	1	Female preference	32
Debiak *et al.* (2014)	Common loon, *Gavius immer*	1	Size	14
Delhey *et al.* (2007)	Blue tit, *Cyanistes caeruleus*	4	Ornament: crown brightness, chroma, hue, UV chroma; size	20, 22, 40, 70
Delmore *et al.* (2008)	Tree swallow, *Tachycineta bicolor*	1	Age; condition; EPP: between broods; size	50
Dietrich-Bischoff *et al.* (2006)	Coal tit, *Parus ater*	1	Age; EPP: within broods, between broods; ornament: bib saturation, bib size; size	235–457
Dolan *et al.* (2009)	Eastern kingbird, *Tyrannus tyrannus*	1	Behaviour: song latency, song rate; EPP: between broods; size	42–72
Dowling & Mulder (2006)	Red-capped robin, *Petroica goodenovii*	1	EPP: within broods; ornament: breast & forehead patch brightness, hue, chroma; size	44
Dreiss *et al.* (2006)	Blue tit, *Cyanistes caeruleus*	1	Behaviour: song characteristics; EPP: within broods; ornament: crown brightness, hue, UV chroma; size	84
Drullion & Dubois (2011)	Zebra finch, *Taeniopygia guttata*	1	Female preference	6
Du & Lu (2010)	Tibetan ground tit, *Pseudopodoces humilis*	2	EPP: between broods	19, 58
Ellegren *et al.* (1996)	Collared flycatcher, *Ficedula albicollis*	1	Ornament: forehead patch size; size	79
Fargallo *et al.* (2004)	Chinstrap penguin, *Pygoscelis antarctica*	1	Behaviour: nest size	37–53
Ferree (2007)	Dark-eyed junco, *Junco hyemalis*	1	Ornament: tail patch size; size	32
Green (2002)	Brown thornbill, *Acanthiza pusilla*	1	Size	104
Griffith *et al.* (2003)	Blue tit, *Cyanistes caeruleus*	3	Age; ornament: plumage brightness, UV chroma	15–16, 22–23, 50–57
Grindstaff *et al.* (2001)	Dark-eyed junco, *Junco hyemalis*	1	Female preference	48
Helle *et al.* (2008)	Field vole, *Microtus agrestis*	1	Condition	28
Hjernquist *et al.* (2009)	Collared flycatcher, *Ficedula albicollis*	1	Condition; ornament: forehead patch size	74
House *et al.* (2011)	Dung beetle, *Onthophagus taurus*	1	Size	50
Husby *et al.* (2006)	House sparrow, *Passer domesticus*	1	Age; ornament: badge size; size	15
Johnson *et al.* (2009)	House wren, *Troglodytes aedon*	1	EPP: within broods, between broods	476

Table 1. Continued

References	Species	N studies	Measures of attractiveness tested	N broods
Karino, Kobayashi & Orita (2006)	Guppy, *Poecilia reticulata*	1	Ornament: tail length	74
Karino & Sato (2009)	Guppy, *Poecilia reticulata*	1	Ornament: orange spot brightness; Female preference	37
Katsuki *et al.* (2012)	Broad-horned flour beetle, *Gnathocerus cornutus*	1	Ornament: mandible length	50
Kempenaers, Verheyen & Dhondt (1997)	Blue tit, *Cyanistes caeruleus*	1	EPP: within broods, between broods	57–141
Kölliker *et al.* (1999)	Great tit, *Parus major*	1	Age; ornament: breast stripe size; size	57
Korpimäki *et al.* (2000)	Kestrel, *Falco tinnunculus*	1	Condition; size	80
Korsten *et al.* (2006)	Blue tit, *Cyanistes caeruleus*	2	Ornament: UV chroma	26, 44
Kraaijeveld *et al.* (2007)	Black swan, *Cygnus atratus*	1	EPP: across all broods; ornament: number of curled wing feathers; size	36
Lee & Hatchwell (2011)	Vinous-throated parrotbill, *Paradoxornis webbianus*	1	EPP: between broods; size	36
Leech *et al.* (2001)	Blue tit, *Cyanistes caeruleus*	1	Age; condition; EPP: within broods; size	103–154
Leitner *et al.* (2006)	Canary, *Serinus canaria*	1	Behaviour: song components	18
Leonard & Weatherhead (1996)	Red junglefowl (chicken), *Gallus gallus*	3	Behaviour: social dominance rank	31, 37, 116
Lessells, Mateman & Visser (1996)	Great tit, *Parus major*	1	Age; size	52–92
Limbourg *et al.* (2013)	Blue tit, *Cyanistes caeruleus*	1	Age; ornament: UV chroma	33
Long & Pischedda (2005)	Fruit fly, *Drosophila melanogaster*	2	Age	35–131, 384
Lu, Zeng & Du (2013)	White-rumped snowfinch, *Montifringilla taczanowskii*	2	Size	21
Magrath, Green & Komdeur (2002)	Fairy martin, *Petrochelidon ariel*	1	Age; condition; EPP: between broods; size	29–56
Neto, Hansson & Hasselquist (2011)	Savi's warbler, *Locustella luscinioides*	1	EPP: between broods; size	102
Nishiumi (1998)	Great reed warbler, *Acrocephalus arundinaceus*	1	Behaviour: arrival date	50
Oddie & Reim (2002)	Great tit, *Parus major*	1	Age; condition; size	143–173
Olsson, Wapstra & Uller (2005*b*)	Sand lizard, *Lacerta agilis*	1	Age; condition; size	58
Pariser *et al.* (2012)	Zebra finch, *Taeniopygia guttata*	1	Ornament: leg band colour	21
Parker (2005)	Red junglefowl (chicken), *Gallus gallus*	1	Ornament: comb size	76
Perlut *et al.* (2014)	Savannah sparrow, *Passerculus sandwichensis*	1	EPP: between broods; size	26–48
Pike & Petrie (2005)	Peafowl, *Pavo cristatus*	1	Ornament: number of eyespot feathers in tail	21
Polo *et al.* (2004)	Spotless starling, *Sturnus unicolor*	1	Behaviour: nest decoration	149
Potvin & MacDougall-Shackleton (2010)	Song sparrow, *Melospiza melodia*	1	Behaviour: song repertoire	55
Questiau *et al.* (2000)	Bluethroat, *Luscinia svecica*	1	Age; ornament: throat patch size; size	27–39
Radford & Blakey (2000*b*)	Great tit, *Parus major*	1	Age; condition; ornament: breast stripe size; size	84
Ramsay *et al.* (2003)	Black-capped chickadee, *Poecile atricapilla*	1	Behaviour: social dominance rank; EPP: between broods	150–174
Rathburn & Montgomery (2005)	White-winged fairy wren, *Malurus leucopterus*	1	Condition; ornament: nuptial plumage	79
Rosivall *et al.* (2004)	Collared flycatcher, *Ficedula albicollis*	1	Age; ornament: forehead patch size, wing patch size; size	54–57
Roulin *et al.* (2010)	Barn owl, *Tyto alba*	1	Ornament: spot size	44
Rutstein *et al.* (2004)	Zebra finch, *Taeniopygia guttata*	1	Ornament: leg band colour	32
Rutstein *et al.* (2005)	Zebra finch, *Taeniopygia guttata*	1	Ornament: leg band colour	101
Saino, Ellegren & Møller (1999)	Barn swallow, *Hirundo rustica*	1	Condition; EPP: across all broods; ornament: tail length; size	19–54
Saino *et al.* (2002)	Barn swallow, *Hirundo rustica*	1	Age; ornament: tail length	76
Saino *et al.* (2003)	Barn swallow, *Hirundo rustica*	1	Ornament: tail length	46
Sato & Karino (2010)	Guppy, *Poecilia reticulata*	1	Female preference	30

Table 1. Continued

References	Species	N studies	Measures of attractiveness tested	N broods
Schwarzova et al. (2008)	Red-backed shrike, *Lanius collurio*	1	EPP: across all broods	15
Sheldon & Ellegren (1996)	Collared flycatcher, *Ficedula albicollis*	1	EPP: within broods	19
Sheldon et al. (1999)	Blue tit, *Cyanistes caeruleus*	2	Age; ornament: plumage hue, brightness, UV chroma	20, 57
Soma & Okanoya (2013)	Bengalese finch, *Lonchura striata*	1	Behaviour: repertoire, song duration, song linearity	29
Stauss et al. (2005)	(a) Great tit, *Parus major* (b) Blue tit, *Cyanistes caeruleus*	2	Condition	(a) 95 (b) 88
Svensson & Nilsson (1996)	Blue tit, *Cyanistes caeruleus*	1	Age	33
Székely et al. (2004)	Kentish plover, *Charadrius alexandrines*	1	Age; size	59
Taff et al. (2011)	Common yellowthroat, *Geothlypis trichas*	2	Age; EPP: within broods; ornament: plumage yellow & UV brightness, carotenoid chroma, mask size, bib size; size	27–71, 44
Veiga, Polo & Cordero (2008)	Spotless starling, *Sturnus unicolor*	3	Behaviour: nest decoration	34, 40, 135
Watt, Shohet & Renshaw (2001)	Guppy, *Poecilia reticulata*	1	Ornament: orange spot size; female preference; size	17–24
Westerdahl et al. (1997)	Great reed warbler, *Acrocephalus arundinaceus*	1	Behaviour: song repertoire; EPP: within broods; size	9–63
Westneat et al. (2002)	House sparrow, *Passer domesticus*	1	Age; condition; ornament: bib size; size	93–142
Whittingham & Dunn (2000)	Tree swallow, *Tachycineta bicolor*	1	Condition	40
Wiebe & Bortolotti (1992)	American kestrel, *Falco sparverius*	1	Size	87
Wilk, Cichon & Wolff (2008)	Collared flycatcher, *Ficedula albicollis*	1	EPP: across all broods	78
Yamaguchi et al. (2004)	Varied tit, *Parus varius*	1	Ornament: forehead & breast patch size; size	59
Zann & Runciman (2003)	Zebra finch, *Taeniopygia guttata*	3	Condition; ornament: breast band size, leg band colour; size	20, 24, 68
Zielinska, Dubiec & Zielinski (2010)	House martin, *Delichon urbicum*	1	Condition; ornament: rump patch colour & size; size	129

N studies reports the number of independent populations/experiments for which the relationship between brood sex ratios and some measure of male attractiveness was tested. *N* broods gives the range of sample sizes for each study; even where a paper contained only one study, different numbers of broods may have contributed to the tests for different measured traits. For EPP (extra-pair paternity): 'within broods' compares the sex ratio of offspring sired by a female's social and extra-pair mate, within the same brood; 'between broods' compares the sex ratio of broods with and without any extra-pair offspring; 'across all broods' compares the probability of being male for individual offspring sired by a within- or extra-pair male.

for analysis), study design (experimental or observational), and study location (laboratory or field research). There was substantial variation among studies in the stage at which offspring sex ratio was measured; we therefore included as a moderator 'primary' (measured at birth/laying) *versus* 'secondary' (some juvenile mortality had occurred) sex ratio rather than excluding studies on this basis.

Seventeen papers reported insufficient information to calculate an effect size for one or more measured male traits (most commonly only reporting that the effect was 'not significant'). We obtained the relevant information directly from the authors of 10 papers, retaining 11 unknown effects from the remaining 7 papers. In addition, 28 papers contained one or more effects for which only model fit statistics or P values were reported, allowing us to extract the magnitude of an effect but not its direction. We obtained the direction of the effect from the authors of three of these papers. We retained 56 'directionless' effects from the remaining 25 papers. We attempted to contact all authors so 'incomplete' data reflects no response by an author or their inability to recalculate the effect size (i.e. data files unavailable).

In all analyses, we tried to account for the uncertainty introduced by effects of unknown magnitude and/or direction. First, we excluded all unknown and directionless effect sizes (dataset 1). Second, if this analysis showed a significant mean effect, we added the unknown effect sizes and assigned them a value of 0 (i.e. assumed that they were highly 'non-significant') (dataset 2). Third, if the analysis was still significant, we added the directionless effect sizes and also assigned them a value of 0 (dataset 3). Finally, if the analysis was still significant, we assigned all directionless effect sizes a negative direction (dataset 4). Each successive dataset represents increasingly conservative assumptions about the magnitude/direction of missing data. This allowed us to test the robustness of our findings to the value of unknown and directionless effects.

(3) Standard random-effects meta-analyses

We conducted random-effects meta-analysis for all trait categories combined (the 'All' analysis), as well as for each of the seven trait categories separately. These analyses assumed independence of effect sizes. Each study therefore contributed one effect size per trait category (and thus per meta-analysis). As noted above, if studies provided more than one effect size estimate per category, the weighted mean effect for the category was calculated. Likewise, for the 'All' analysis we calculated the weighted mean effect across all effect sizes for the study. We then assigned the variance to this combined effect size based on the study sample size (number of broods). This is the most conservative option as it assumes complete dependence between multiple estimates (e.g. that $r = 1$ for the correlation between within-study estimates: see p. 228 in Borenstein *et al.*, 2009).

For all random-effects meta-analytic models, we used a restricted maximum likelihood (REML) method to estimate τ^2 (the true between-study variance), which was combined with the within-study variance for each effect size (Borenstein *et al.*, 2009). For the 'All' data set, we used meta-regression to determine the effects of four study-level moderator variables: year of publication, design (experimental/observational), location (field/laboratory), and stage at which sex ratio was measured (primary/secondary) on the effect size. Based on the results, we did not include these moderators in the subsequent analyses for separate trait categories.

We also tested whether there was more heterogeneity in effect sizes among studies than could be explained by sampling error alone using Cochran's Q statistic. This formally tests whether there is greater variation in effect sizes among studies than expected if the true effect is identical for all studies (Hedges & Olkin, 1985). We also summarised heterogeneity using I^2, which is often defined as the ratio of true heterogeneity to the total variance among studies. Borenstein *et al.* (2009) note, however, that this is not always correct because the ratio assumes a constant within-study variance, which is not the case (sampling error varies due to studies differing in sample size). They suggest that I^2 be treated as a measure of 'inconsistency' in effect sizes among studies. Higgins *et al.* (2003) have suggested that benchmarks for I^2 of 25, 50 and 75% indicate low, moderate and high values of inconsistency among studies.

All statistical analyses for the above models were conducted in R statistical software using the *metafor* package (Viechtbauer, 2010).

(4) Multilevel meta-analyses including phylogenetic effects

While the standard meta-analyses described above included studies of vertebrate and invertebrate taxa, the multilevel meta-analyses with and without phylogenetic effects were limited to bird species, because there were few studies on other taxa in our data set (84.2% of species, 87.1% of included papers and 92.0% of effect sizes were for birds). The other taxa also had relatively few species sampled within each family. A tree including all species would be unreliable in terms of branch lengths (i.e. distance relationships between species). By contrast, we now have avian super-trees that include all extant species (Jetz *et al.*, 2012; see below). For birds, our data set had enough variation to conduct a phylogenetic meta-analysis. Studies could contribute multiple effect sizes because the multilevel model partially corrects for non-independence by modelling multiple random effects for both study and species. A complete correction of non-independence requires modelling the correlations among effect sizes. For example, effect sizes for ornament and body size are likely to be correlated when these traits were measured from the same group of individuals. Given the difficulty of determining actual correlations among related effect sizes (as such correlations are almost always not reported), the multilevel models we report in the main text assume correlations of zero. Additionally, as supplementary material (see online Table S2) we report on multilevel models that conservatively assume all correlations to be 0.5. Note that qualitatively

the results are very similar among these models. As for the standard random-effect meta-analyses we initially excluded all directionless and unknown effect sizes (avian dataset 1) and then added them to subsequent analyses if the preceding analysis yielded a significant mean effect size (avian datasets 2–4).

To estimate the overall effect size we used four Bayesian Markov Chain Monte Carlo (MCMC) generalised linear mixed-effects models, implemented in the MCMCglmm package of *R* (Hadfield, 2010; Hadfield & Nakagawa, 2010). First, the multilevel meta-analytic model (model A) included study and species as random effects. We did not include 'paper' as a random effect because it closely overlaps with study (99 and 81 levels, respectively, as most papers comprised a single study). This model was analogous to the standard random-effects meta-analysis described above, but used MCMC estimator rather than REML estimator and had two extra random effects (study and species). Second, we extended model A by running a phylogenetic meta-analysis that included a phylogenetic tree (model B). Third, we ran a multilevel meta-regression (model C) that extended model A by including moderators (fixed effects): year of publication, seven categories of attractiveness measure (Ornament, Behaviour, Size, Condition, Preference, Age, and EPP), design (experimental or observational), and location (laboratory or field study). This was to assess whether these study characteristics impacted the effect sizes, similar to the meta-regression used in the standard random-effects meta-analysis. Finally, for completeness, we ran a meta-regression model (model D) that extended model C by including phylogenetic effects. Models C and D were both run without the intercept. This ensures that the mean estimates for the seven trait categories (conditional on the other fixed effects) were directly obtained (for details, see Section III).

For both phylogenetic models (B and D) we used two different avian phylogenetic trees: (*i*) a version of the Ericson backbone trees from Jetz *et al.* (2012), referred to as the Ericson tree, and (*ii*) a version of the Hackett backbone trees also from Jetz *et al.* (2012), referred to as the Hackett tree. These phylogenies are presented in Fig. S1. Both trees were ultrametric with branch lengths. Each phylogenetic model produced quantitatively similar and qualitatively identical results using either tree. We therefore only present results for the Ericson tree (for results from the Hackett tree see online Table S1 and Fig. S2). As an index of phylogenetic signal we calculated phylogenetic heritability H^2, which can be defined as the proportion of phylogenetic variance in relation to the sum of all other variance components except sample error variance. Along with $I^2_{\text{phylogeny}}$ (see below), H^2 provides the magnitude of phylogenetic signal in the data. H^2 is equivalent to Pagel's λ when the unit of analysis is at the species level (Housworth, Martins & Lynch, 2004).

For all multilevel (phylogenetic) models, we used an inverse Gamma prior ($V = 0.002$ and $nu = 1$) for all random effects. Each model was run for 1100000 iterations with a thinning interval of 1000 after a burn-in of 100000 iterations. These settings resulted in posterior distributions consisting of 1000 samples for all parameters. We assessed chain mixing by examining autocorrelation among posterior samples; they were less than 0.1 for all estimated parameters (Hadfield, 2010). We report our point estimates from models based on posterior means rather than posterior modes. We considered fixed factor estimates statistically significant if the 95% credible interval (CI, also known as the highest posterior density, which can be considered a Bayesian equivalent of the frequentist confidence interval) did not overlap zero. To quantify heterogeneity in multilevel meta-analytic models, we calculated a modified version of the I^2 statistics, following Nakagawa & Santos (2012). Basically, this partitions the proportion of 'unknown' variance not attributable to sampling variance (i.e. traditional I^2) into the contribution of random factors. In the current case these are the variance in effect sizes due to phylogenetic relatedness, differences among species, differences among studies, and differences in within-study variation (also called 'residual' variation). The sum of the percentages of total variation due to these four sources equals the traditional I^2 of Higgins *et al.* (2003).

(5) Bias detection

A funnel plot reports the relationship between effect sizes and a measure of their variance (e.g. sampling variance, standard error or sample size). Variability in the observed effect size should decrease as sample size increases (i.e. as the estimate becomes more accurate due to lower sampling error). This yields a funnel-shaped plot. Almost all statistical tests for publication bias are based on the assumption that the funnel plot should be symmetrical (Jennions *et al.*, 2013). Asymmetry is assumed to reflect publication bias. We used funnel plots for visual inspection of funnel asymmetry, and then conducted Egger's regression (Egger *et al.*, 1997) to test statistically for publication bias (i.e. funnel plot asymmetry) for each category of attractiveness measure. We then conducted trim-and-fill tests (the L0 model; Duval & Tweedie, 2000*a*,*b*) as another test for publication bias. Asymmetry due to publication bias is only predicted to appear as missing studies on the left of the plot (i.e. fewer significant results in the direction opposite to a true non-zero effect), but trim-and-fill can detect asymmetry on either side. We present results when studies were missing on the left or right. We do so because, in a broader context, the proportion of tests that identify missing studies on the left *versus* right can be viewed as an indicator of the extent to which asymmetry is due to publication bias (left only) rather than other factors that might generate asymmetry on either side. Small sample sizes limit the detection of significant asymmetry using Egger's regression (Sterne & Egger, 2005), and funnel asymmetry can occur by chance especially when the sample size is low. Consequently, care is required when interpreting these publication bias analyses.

Tests for bias were only run on 'Data 1' data sets (i.e. excluding unknown and directionless effect size estimates). For the multi-level models (i.e. the avian-only data sets) we conducted analyses of bias on meta-analytic residuals
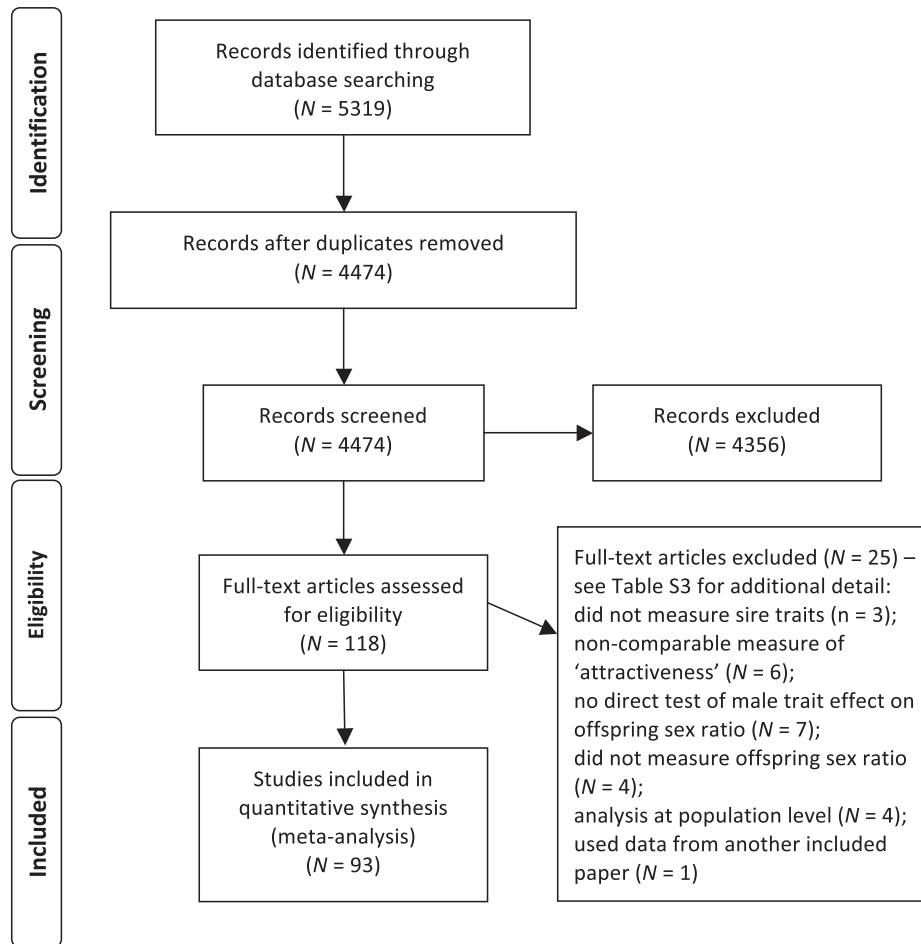
Fig. 2. Preferred reporting items for systematic reviews and meta analyses (PRISMA) flow chart of our literature search and data collection process (see http://www.prisma-statement.org/).

(*sensu* Nakagawa & Santos, 2012). This ensures that we meet assumptions of independence, and that we reduce the effect of heterogeneity due to moderators that could induce funnel asymmetry. All publication bias analyses were conducted using functions in the R package *metafor*.

III. RESULTS

(1) Attrition of studies

We scanned the title and abstract of all 4474 papers generated by our search protocol to identify studies potentially suitable for inclusion in the meta-analysis. We then read and assessed the methods and results sections of short-listed studies to determine their eligibility for inclusion. We identified 118 eligible papers. After full-text screening, 25 of these papers were excluded (Fig. 2, see online Table S3). In total, 276 effect sizes for the meta-analysis were computed from 113 independent studies within 93 published papers. Of these, 254 effect sizes were from studies of birds (48 species). The remaining studies contributed nine effect sizes from four

insect species, eight effect sizes from one species of fish, four effect sizes from three reptile species, and one effect size from a mammal. The full data set, and the data set reduced to one effect size per trait category per study are provided in Appendix S3.

(2) Standard random-effects meta-analyses

The mean effect of male attractiveness on offspring sex ratio for traits pooled across all categories was small, positive and statistically significant for 'dataset 1' (Fig. 3A, Table 2). Females mated to attractive males produced more sons. This result was robust to the implementation of the successively more conservative datasets 2, 3 and 4 (all $P < 0.005$; Table 2). There were, however, reasons for concern about the effect of publication bias. The Egger's regression test revealed significant asymmetry in the funnel plot ($t_{79} = 3.45$, $P = 0.001$; Table 3), and trim-and-fill estimated 17 (± 5.9 S.E.) studies 'missing' from the left-hand side of the plot (see Fig. 4A). Correcting for this bias substantially reduced the mean effect size estimate, which no longer differed from zero ($P = 0.149$; Table 3).
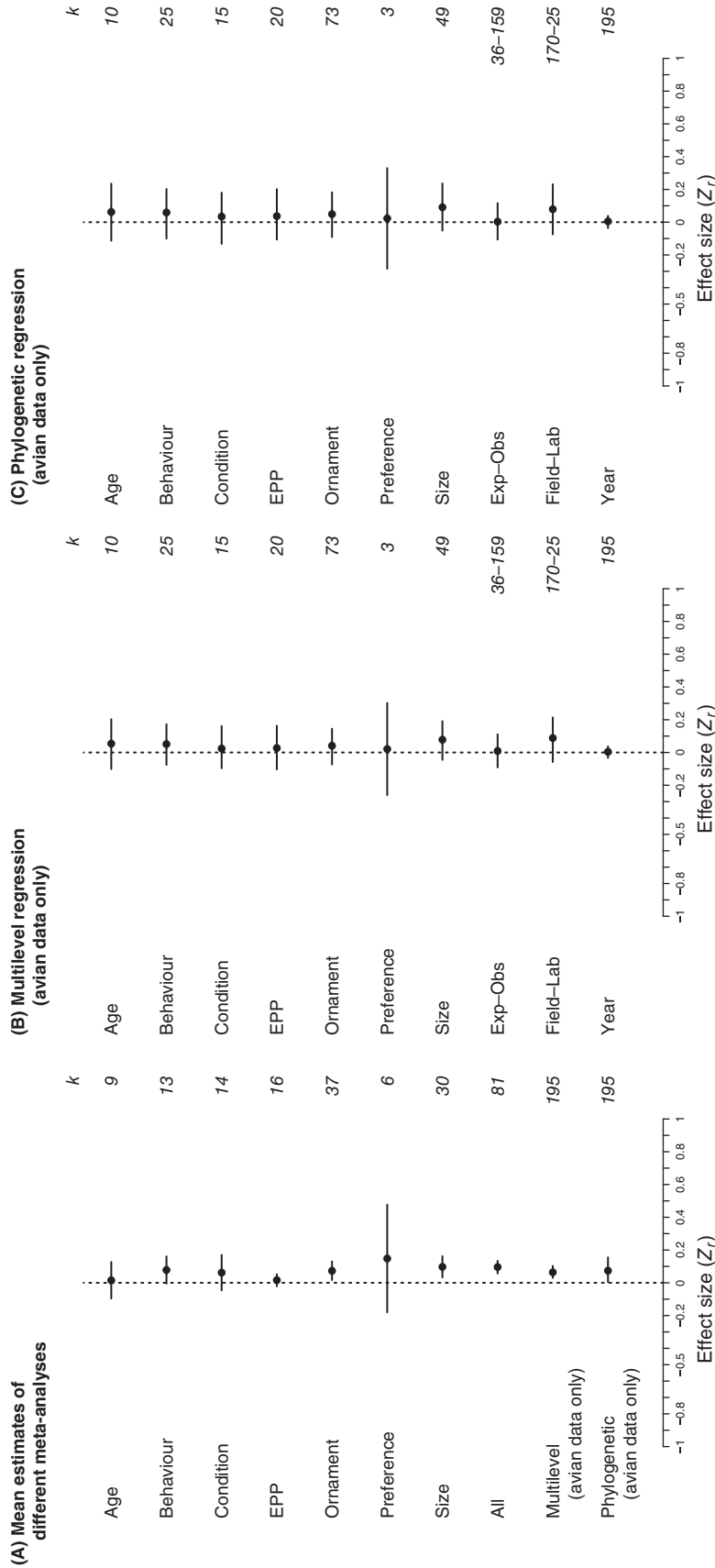
Fig. 3. Forest plots of (A) the meta-analytic means of the eight random-effects meta-analyses, the multilevel meta-analysis, and the phylogenetic meta-analysis (the latter two used the avian-only data set; the phylogenetic analysis used the Ericson backbone tree), (B) the regression coefficients of the multilevel meta-regression, and (C) of the phylogenetic meta-regression. EPP, extra-pair paternity; *k*, number of effect sizes.

Table 2. Results from the random-effects meta-analyses for the eight different trait categories

Trait	Dataset	k	m	n	Mean ($\bar{Z_r}$)	Lower CI (2.5%)	Upper CI (97.5%)	z value	$P(z)$	Q	d.f. (Q)	$P(Q)$	I^2 (%)
All	1	81	49	7512	0.096	0.057	0.134	4.815	<0.001	177.40	80	<0.001	56.20
	2	84	50	7807	0.072	0.040	0.104	4.432	<0.001	136.70	83	<0.001	37.07
	3	94	57	8788	0.061	0.033	0.089	4.238	<0.001	138.35	93	0.002	30.02
	4	94	57	8788	0.039	0.012	0.067	2.817	0.005	135.46	93	0.003	26.77
Age	1	9	6	1053	0.015	−0.096	0.127	0.273	0.785	19.60	8	0.012	58.81
Behaviour	1	13	11	1159	0.078	−0.005	0.161	1.849	0.064	23.72	12	0.022	43.21
Condition	1	14	12	1908	0.062	−0.046	0.170	1.118	0.264	38.59	13	<0.001	73.29
EPP	1	16	14	3159	0.016	−0.021	0.052	0.851	0.395	15.26	15	0.433	0.000
Ornament	1	37	18	2394	0.073	0.017	0.130	2.534	0.011	61.86	36	0.005	42.56
	2	38	18	2418	0.072	0.016	0.128	2.528	0.012	61.94	37	0.006	41.31
	3	43	21	2911	0.060	0.012	0.109	2.451	0.014	63.61	42	0.017	34.68
	4	43	21	2911	0.042	−0.006	0.089	1.708	0.088	64.32	42	0.015	33.26
Preference	1	6	4	172	0.148	−0.180	0.477	0.885	0.376	17.26	5	0.004	73.43
Size	1	30	27	1757	0.097	0.033	0.162	2.950	0.003	49.02	29	0.011	42.67
	2	31	28	1865	0.093	0.031	0.155	2.926	0.003	50.10	30	0.012	41.86
	3	44	36	3207	0.060	0.018	0.102	2.778	0.005	57.30	43	0.071	25.75
	4	44	36	3207	0.025	−0.020	0.069	1.092	0.275	61.98	43	0.030	32.41

k = number of effect sizes, m = number of species, n = number of broods in each data set used for these meta-analyses (for the descriptions of Dataset, Q and I^2, see text Section II); EPP, extra-pair paternity.

Table 3. Results of Egger's regression tests and trim-and-fill (TAF) tests (for the descriptions of Q and I^2, see text Section II)

| Trait	t: Egger	d.f. (t)	$P(t)$	Missing k: TAF	Mean ($\bar{Z_r}$)	Lower CI (2.5%)	Upper CI (97.5%)	z value	$P(z)$	Q	d.f. (Q)	$P(Q)$	I^2 (%)
All	3.450	79	0.001	17	0.034	−0.012	0.080	1.443	0.149	287.08	97	<0.001	72.99
Age	3.650	7	0.008	4	−0.102	−0.239	0.036	−1.447	0.148	44.40	12	<0.001	80.17
Behaviour	0.918	11	0.378	0	—	—	—	—	—	—	—	—	—
Condition	0.949	12	0.361	3	−0.009	−0.129	0.111	−0.143	0.886	63.89	16	<0.001	81.71
EPP	0.293	14	0.774	0	—	—	—	—	—	—	—	—	—
Ornament	1.515	35	0.139	7	0.029	−0.031	0.089	0.935	0.350	92.54	43	<0.001	53.75
Preference	−0.873	4	0.432	2*	0.304	−0.046	0.654	1.701	0.089	27.09	7	<0.001	78.40
Size	−0.984	28	0.334	7*	0.151	0.087	0.215	4.618	< 0.001	74.02	36	<0.001	52.33

EPP, extra-pair paternity.
*Missing effect sizes in the unexpected direction.

The heterogeneity due to variation between studies that could not be attributed to sampling error, I^2, was moderate at 56.20% (*sensu* Higgins *et al.*, 2003). Given the wide range of taxa, traits and study designs, one might expect a higher level of heterogeneity. To try to explain some of this heterogeneity, we ran a meta-regression model including the moderators year of publication, design, location and stage at which sex ratio was measured. They did not, however, explain a significant proportion of the observed heterogeneity ($Q_4 = 2.848$, $P = 0.584$; $b_{[year]} = 0.014$, 95% confidence interval, CI [−0.324, 0.061], $b_{[Experiment–Observation]} = 0.026$ [−0.080, 0.133], $b_{[Field–Laboratory]} = 0.082$ [−0.022, 0.185], $b_{[Primary–Secondary]} = 0.014$ [−0.082, 0.109]). Given this result, we did not run meta-regression models for the other trait categories.

When we examined the effects for each trait category separately, we found a significantly positive, albeit small, effect for both male ornamentation and male size (Fig. 3A, Table 2). Females mated to more ornamented males or to larger males produced a higher proportion of sons than those mating with less ornamented or smaller mates. For both categories, this result held across datasets 1–3, and was only statistically non-significant in the most conservative 'dataset 4' analysis (where directionless effect sizes were assigned negative values) (Table 2). For ornaments, correction for a potential publication bias identified using trim-and-fill resulted in a mean effect size estimate that was non-significant ($P = 0.350$; Table 3). By contrast, for male size, trim-and-fill identified 7 (± 3.6 S.E.) studies 'missing' from the right of the funnel plot. Correcting for this asymmetry increased the mean effect size estimate to 0.15 (± 0.03 S.E.; $P < 0.001$). The mean effect size estimate did not differ from zero for the other five trait types (behaviour, body condition, female preference score, male age, and male extra-pair or within-pair status) (Fig. 3A, Table 2). For male age, the Egger's test suggested a publication bias, and trim-and-fill analysis suggested there were four missing studies. Even so, the addition of putative effect sizes still did not yield a
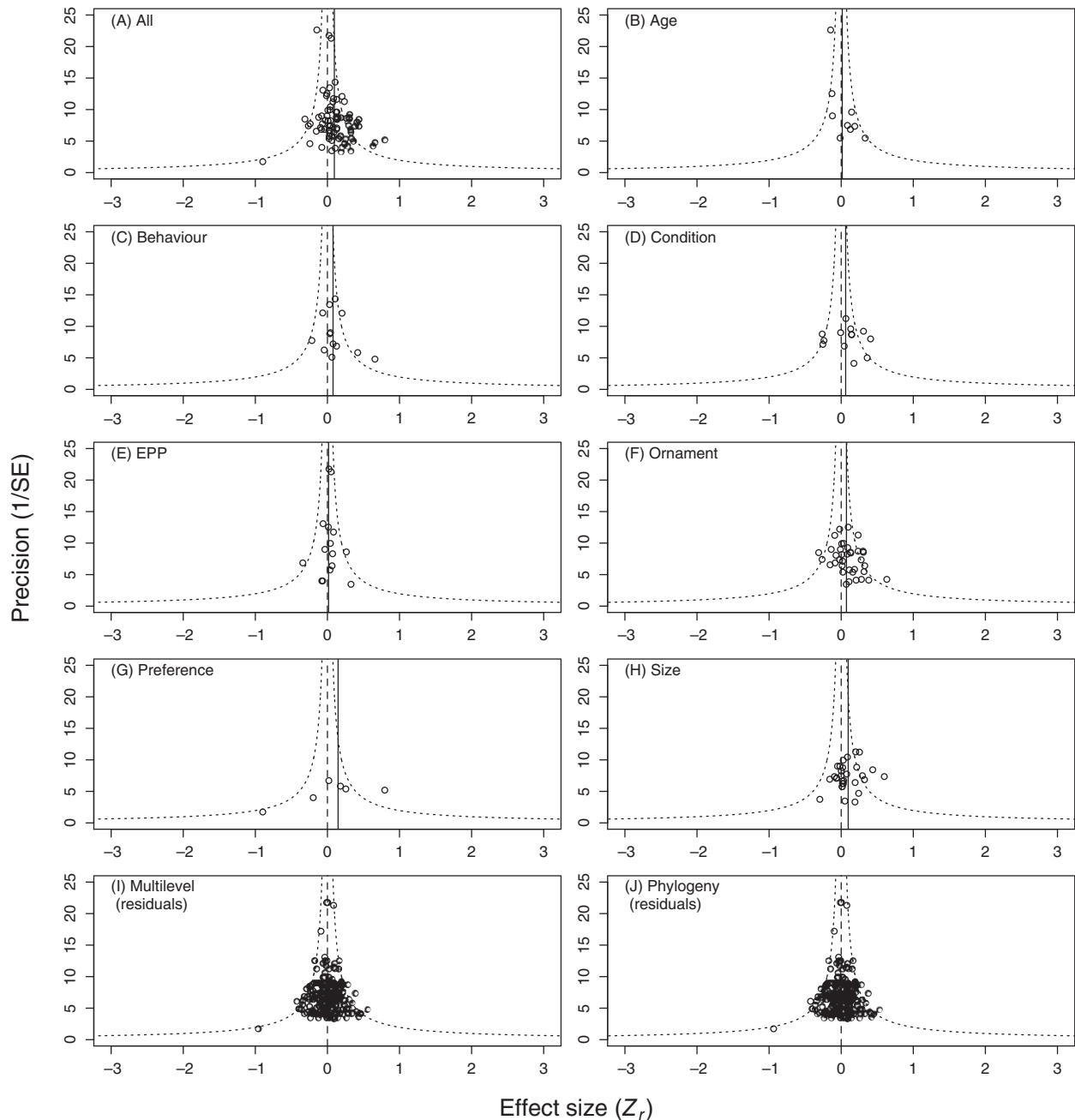
Fig. 4. Funnel plots (effect sizes plotted against their corresponding precision, the inverse of the standard error) of (A–H) the eight random-effects meta-analyses, (I) the multilevel meta-analysis and (J) the phylogenetic multilevel meta-analysis (Ericson backbone tree). Plots for the multilevel analyses (I, J) use the meta-analytic residuals (see Nakagawa & Santos, 2012); these analyses used the avian-only data set. EPP, extra-pair paternity.

significant effect size (Table 3). Trim-and-fill analyses also suggested there were three missing studies on the left for body condition, and two on the right for preference scores. Again, the addition of these studies did not yield significant mean effect size estimates (both $P > 0.089$; Table 3). Funnel plots are shown in Fig. 4.

The estimates of heterogeneity for the individual trait categories (dataset 1 analyses) ranged from $I^2 = 42.56$ to 73.43% (Table 2). This is moderate to high (*sensu* Higgins *et al.*, 2003), but it should be noted that uncertainty in estimates of I^2 is high when sample sizes are small. These results, along with those for all trait categories combined, suggest that over half the variation in effect sizes was due to systematic between-study differences. The sources of this variation are explored in Section IV. Notably, however, there was no heterogeneity in the effect size estimates for

Table 4. Results of multilevel and phylogenetic (random-effects) meta-analyses

Type	Dataset (avian)	k	m	n	Mean ($\bar{\mathcal{Z}}$)	Lower CI (2.5%)	Upper CI (97.5%)	I^2_{study} (%)	$I^2_{species}$ (%)	$I^2_{effect\ size}$ (%)	$I^2_{phylogeny}$ (%)	I^2_{total} (%)	H^2 (%)
Multilevel (non-phylogenetic)	1	195	42	85	0.064	0.030	0.103	21.03	6.03	24.50	—	51.56	—
	2	205	43	88	0.058	0.026	0.093	17.53	5.84	23.79	—	47.16	—
	3	254	48	99	0.041	0.017	0.066	11.04	5.59	12.01	—	28.64	—
	4	254	48	99	0.025	−0.002	0.052	12.37	6.08	18.84	—	37.29	—
Phylogenetic	1	195	42	85	0.074	0.005	0.155	16.26	5.45	23.21	10.88	55.79	9.82
	2	205	43	88	0.069	−0.009	0.142	15.34	5.06	21.40	10.47	52.26	9.77

k = number of effect sizes, m = number of species and n = number of experiments/populations (or studies) in each data set used for these meta-analyses (for the descriptions of Dataset, H^2 and I^2, see the text Section II).

extra-pair mating status ($I^2 = 0$), indicating all variation was due to sampling error, although the number of studies in this analysis was rather limited.

(3) Multilevel meta-analyses including phylogenetic effects

For both the simple and phylogenetic multilevel meta-analyses we found small but significantly positive mean effect sizes for 'avian dataset 1' analyses (Table 4; Fig. 3A–C). These estimates were very similar to those from the standard random-effects meta-analysis for all traits pooled. The simple multilevel meta-analytic model (model A) had 95% credible intervals that did not span zero for avian datasets 2 and 3, but did for the most conservative dataset 4 (but even then, the lower CI was only just below zero at −0.002; Table 4). For the phylogenetic multilevel meta-analysis (model B), the mean effect had a 95% credible interval that included zero for avian dataset 2, so we did not run datasets 3 or 4.

In both multilevel meta-analytic models we can partition the variance not due to sampling error (i.e. heterogeneity). Here, I^2_{study} is the variance due to differences among studies while $I^2_{species}$ is that among species. In the phylogenetic models, $I^2_{species}$ is the variation due to the non-phylogenetic portion of species-level differences (associated with repeatability of effect sizes among species), while $I^2_{phylogeny}$ is the variance due to phylogenetic relatedness. $I^2_{effect\ size}$ is the residual variance that is not due to the sampling error. I^2_{total} is the sum of these values. In the simple multilevel meta-analysis, the largest component of variance for 'avian dataset 1' (effects of known direction and magnitude) was at the effect-size level (24.5%) then the study level (21.0%). Species had little influence (6.0%) (Table 4), suggesting that repeatability is very low, although low replication within species (only 13 of 57 species were represented in more than one study, and only six species in more than two studies) likely contributes to this result. For the phylogenetic, multilevel meta-analysis of avian dataset 1, the largest component was also at the effect-size level (23.2%) followed by the study level (16.3%), while 10.9% of heterogeneity was due to phylogeny (Table 4).

The meta-regression models (C and D) used trait category as one of four moderators. The effect size estimates for each trait type were comparable to those

from the standard random-effects models (compare Fig. 3A with 3B or C). Again, the other moderators had no significant effect ($b_{[year]} = 0.005$, 95% credible interval, CI [−0.036, 0.039], $b_{[Experiment–Observation]} = 0.003$ [−0.106, 0.116], $b_{[Field–Laboratory]} = 0.079$ [−0.074, 0.231]).

Egger's regression test did not indicate any funnel asymmetry based on meta-analytic residuals from the multilevel meta-regression models (simple: $t_{193} = 1.036$, $P = 0.302$; phylogenetic: $t_{193} = 0.994$, $P = 0.321$; see funnel plots, Fig. 4). However, the trim-and-fill tests indicated that $8 (\pm 8.6\ SE$; multilevel) and $5 (\pm 8.4\ SE$; phylogenetic) effect sizes were missing, and that we might have overestimated the simple and phylogenetic model means by 0.0093 and 0.0054, respectively. Shifting the CIs for 'avian dataset 1' by these values would result in a credible interval that included zero for the phylogenetic meta-analysis, but not for the simple multilevel meta-analysis. Given this result, and the importance of controlling for phylogeny, our analyses did not provide robust evidence for an offspring sex-ratio bias related to male attractiveness.

IV. DISCUSSION

(1) General findings

We used meta-analysis to test whether females mated to attractive males have a more male-biased offspring sex ratio (Burley, 1981). There was a small but statistically significant relationship in the predicted direction. However, male attractiveness accounted for less than 1% of the variation in offspring sex ratios based on estimates from either standard random-effects meta-analysis or multilevel meta-analysis. The estimated mean effect sizes were $r = 0.096$ and 0.064, respectively. Similarly, when we divided the data into different trait categories we found small, but significant, effects of male attractiveness on the sex ratio for two trait types: male ornamentation and male body size. The means for all seven traits were positive, but there were no significant effects for male behavioural displays, body condition, age, extra-pair status or female preference score. Studies measuring ornaments and body size were the most common (these categories contained 37 and 30 studies,

respectively, for dataset 1), while there were relatively few studies for any of the other five trait types (6–16 studies; Table 2). The absence of a significant effect for these five trait types might reflect low statistical power. This is plausible given the estimated magnitude of the effects: no matter how the data were analysed, and even when significant, the mean r was always <0.2. Alternatively, the lack of significantly positive effect size estimates for some traits might reflect a true null relationship, and/or that unmeasured variation obscured a true relationship. In general though, it is clear that the offspring sex ratio is, on average, only weakly associated with various measures of male attractiveness.

The significant mean effect size observed when combining data from all trait types could be seen as strong evidence for the hypothesis, first proposed by Burley (1981), that there is facultative maternal SRA in response to mate attractiveness/quality. It is, however, prudent to consider the robustness of meta-analytic results to narrow-sense publication bias (i.e. unpublished studies) and/or selective reporting of data (i.e. failure to report non-significant results, or to do so in a way that hinders the ability to calculate effect sizes; see Cassey *et al.*, 2004; review: Jennions *et al.*, 2013). We used increasingly conservative data sets (datasets 2–4) to include unknown effect sizes in our analyses, treating them as neutral or opposite to the test hypothesis. Depending on the analysis, the significance of the mean effect tended to vary in its sensitivity to the inclusion of these effects. Specifically, for the standard random effects meta-analysis the estimated mean effect stayed significantly positive for all traits combined across datasets 2, 3 and 4; for ornaments across datasets 2 and 3; and for body size across datasets 2 and 3. For the simple multi-level analysis the estimated mean effect remained significantly positive for all traits combined across avian datasets 1–3.

Crucially, the mean effect was not significant for avian dataset 2 when phylogenetic relatedness between species was taken into account (multilevel model B). This analysis is arguably the most important we present. Although it is based only on bird studies, these represent 92% of the available effect sizes. It is appropriate to take phylogeny into account when conducting a meta-analysis that involves reasonably closely related taxa (Chamberlain *et al.*, 2012). There is therefore justifiable concern about the robustness of the main conclusion to selective reporting. In addition, a publication bias test for 'missing' studies indicated a strong asymmetry in the funnel plot for all traits. If the putative missing studies were included, the estimated mean effect size was only $r = 0.034$, and did not differ significantly from zero. Again, this suggests that the finding of a significant mean effect size is sensitive to narrow-sense publication bias.

Another perspective is that our most important analysis is that based on ornamental male traits. A strict interpretation of the mate attractiveness hypothesis relies on the existence of paternal traits whose heritable benefits are partly male-limited (which is most obviously applicable to ornaments). The initial 'dataset 1' analysis indicated a significant effect of ornaments on SRA. At face value, this

might reflect the use of traits that are likely to be under the strongest sex-specific selection of any of our categories. A stronger sex bias in the heritable benefits of male traits theoretically leads to stronger SRA results: for example, SRA in response to variation in Fisherian traits (given mutation bias) is predicted to result in more strongly biased offspring sex ratios than is SRA in response to viability indicator traits signalling genes that tend to benefit both sexes (Pen & Weissing, 2000; Fawcett *et al.*, 2007). Again, however, there is cause for concern about publication bias. The significant effect for ornaments was robust to selective reporting (i.e. datasets 2 and 3 still yielded a significant mean effect size), but not to narrow-sense publication bias. A trim-and-fill analysis identified seven 'missing' studies that, when imputed, yielded a non-significant mean effect size estimate of $r = 0.029$.

The only other trait type with a significant mean effect size was male body size. Why this might be the case is unclear. In birds, which comprise most of our data, there is little experimental evidence that females choose larger males as mates. It is noteworthy, however, that body size frequently correlates with major components of male fitness (e.g. success when competing for mates and/or resources, survival, and EPP) in birds (e.g. Kölliker *et al.*, 1999; Kissner, Weatherhead & Francis, 2003). More generally, body size often predicts male attractiveness and mating success in other taxa (Fairbairn, Blanckenhorn & Székely, 2007). If sons benefit more than daughters from increased body size then, whether or not females prefer larger males, it could still be adaptive for females to engage in SRA in response to the size of their mate. This is a subtly different explanation from one based on female assessment of mate attractiveness, but only if females do not prefer larger males. It seems possible that the role of male size in attractiveness has been underestimated in birds – or should at least be tested for experimentally. Interestingly, the results for body size are robust to publication bias. The significant effect was not sensitive to selective reporting (i.e. datasets 2 and 3 still yielded a significant mean effect size). It was also robust to narrow-sense publication bias. A trim-and-fill analysis identified seven 'missing' studies but their putative values were positive: when imputed, the mean effect size was $r = 0.151$ ($P < 0.001$).

(2) Reasons for a small effect: sex allocation decisions are influenced by many factors

There were moderate levels of among-study variation not attributable to sampling error (i.e. heterogeneity) in our data sets (Table 2). This indicates that fairly large differences among studies were unaccounted for. None of our three main moderator variables (year of publication, laboratory *versus* field studies, or observational *versus* experimental studies) had a significant influence on the relationship between male attractiveness and the offspring sex ratio. This was true for both the standard random-effects model and the multilevel analyses. The inclusion of phylogenetic information did little to explain heterogeneity, which remained around

Table 5. Factors proposed to affect the relative fitness of sons and daughters differentially, leading to adaptive sex ratio adjustment strategies

Influencing factor		References
Maternal condition		Trivers & Willard (1973), Clutton-Brock *et al.* (1984) and Kruuk *et al.* (1999); meta-analyses: Cameron (2004) and Sheldon & West (2004)
Mate attractiveness/quality		Burley (1981, 1986) (and see Table 1, this study); meta-analyses: West & Sheldon (2002), Ewen *et al.* (2004) and Cassey *et al.* (2006)
Local resource enhancement (cooperative breeders: number of helpers)		Taylor (1981) and Gowaty & Lennartz (1985); meta-analyses: West & Sheldon (2002), Griffin *et al.* (2005) and Stubblefield & Orzack (2013)
Local resource competition		Taylor (1981) and Cockburn *et al.* (1985); meta-analysis: Silk & Brown (2008)
Local mate competition		Hamilton (1967), Werren (1983) and Shuker & West (2004); meta-analysis: West *et al.* (2005)
Resource quality/availability:	Diet quality, composition	Warner, Lovern & Shine (2008), Pryke & Rollins (2012) and Rutz (2012)
	Habitat/territory quality	Romano *et al.* (2012) and Bell *et al.* (2014)
	Host size for parasitoid larval development	Godfray (1994); meta-analysis: West & Sheldon (2002)
Seasonal effects:	Time of breeding/laying date	Saino *et al.* (2008), Graham *et al.* (2011) and Barclay (2012); meta-analysis: Ewen *et al.* (2004) and Cassey *et al.* (2006)
	Rainfall	Berkeley & Linklater (2010)
Laying/hatch order		Cichon, Dubiec & Stoczko (2003), Lezalova *et al.* (2005) and Lislevand *et al.* (2005)
Clutch size		Benito *et al.* (2013) and Bowers *et al.* (2014)
Social dominance rank (maternal)		Nevison (1997) and Maestripieri (2002); meta-analyses: Brown & Silk (2002), Schino (2004) and Silk *et al.* (2005)
Parental relatedness/compatibility		Pryke & Griffith (2009) and Sardell & DuVal (2014)
Parental age/breeding experience/pair bond duration		Green (2002) and Benito *et al.* (2013)
Female mating status/extent of polygyny (sole/primary/secondary female, harem size)		Weatherhead (1983), Nishiumi (1998) and Trnka *et al.* (2012)

Examples are given of empirical studies that report a significant effect of the factor; note that for all factors we also located studies that find no relationship. References to meta-analyses are provided where possible.

50% (Table 4). The unexplained variance among studies highlights the difficulty of testing specific hypotheses for SRA. Many selective factors have been suggested to affect SRA and, for some of them, meta-analyses have formally shown that they explain a significant amount of variation in offspring sex ratios (review: West, 2009) (Table 5). These putative forces of selection on SRA are not mutually exclusive; nor, more importantly here, do they act independently of any influence of mate attractiveness on SRA. If multiple sources of selection on SRA interact they could create extremely complex outcomes. This makes it difficult to predict the optimal offspring sex ratio. For example, the benefits of SRA based on mate attractiveness appear to be seasonally dependent (Griffith *et al.*, 2003; Korsten *et al.*, 2006; Delhey *et al.*, 2007), context-dependent (Pryke *et al.*, 2011), or only reliably indicated by specific age classes of males (Taff *et al.*, 2011).

There are numerous other obstacles to testing SRA hypotheses directly. One is the ubiquitous problem of inferring causality from observational studies where there are potential confounding factors that actually generate the observed pattern. For example, assortative mating such that females in good condition tend to mate with more attractive males seems likely in many seasonally breeding bird species. If this occurs, then a male-biased offspring sex ratio might be partly (or entirely) driven by maternal condition (i.e. Trivers & Willard, 1973), rather than by male-biased genetic benefits. Similarly, males might vary in their ability to provide parental care, which could affect SRA (e.g. Bowers *et al.*, 2013). If parental ability is correlated with male attractiveness and sons and daughters differ in the benefits of increased investment, this could lead to adaptive SRA that is correlated with, but not causally related to, attractiveness. Some observational studies in our data set do, however, try to control statistically

for confounding maternal and environmental variables (e.g. Addison, Kitaysky & Hipfner, 2008; Taff *et al.*, 2011).

Another potentially common problem arises if the benefits of SRA depend on a male–female interaction. For instance, one of the clearest examples of adaptive SRA – mainly because it involves several different experimental manipulations of male colouration – is in Gouldian finches (*Erythrura gouldiae*). In this species there is severe genetic incompatibility between red and black colour morphs, which has a more detrimental effect on daughters than sons. Females express strong mating preferences for males of the same colour morph. When paired with a male bearing the phenotype of the other morph they exhibit adaptive SRA towards sons (Pryke & Griffith, 2009; Pryke *et al.*, 2011). This species was excluded from our meta-analysis because we only investigated directional female preferences for male traits (i.e. we excluded individual-specific preferences). A type of male–female interaction that might arise in many species and select for SRA based on a non-directional preference is associated with inbreeding. There is evidence that pair relatedness can affect adaptive SRA (e.g. Sardell & DuVal, 2014). If widespread, this could confound patterns of SRA predicted solely on a male's general attractiveness.

Even when genetic incompatibility is absent, SRA might still depend on both female and male qualities. Studies of zebra finches (*Taeniopygia guttata*) suggest that assortative mating accounts for the strong sex ratio biases based on artificial leg colour bands originally detected by Burley (Rutstein *et al.*, 2005). If higher quality females can better bear the costs of producing the more attractive/expensive sex (Katsuki *et al.*, 2012), assortative mating between high-quality females and attractive males could amplify adaptive SRA based solely on male identity. One caveat to this is that the influence of sexually antagonistic genes on adaptive SRA has been poorly explored (but see Calsbeek & Sinervo, 2004; Alonzo & Sinervo, 2007; Blackburn, Albert & Otto, 2010). Given sexual antagonism, it is plausible that assortative mating might eliminate SRA. If genes affecting female quality have more beneficial effects on daughters' fitness (e.g. Leimar, 1996), while genes that affect male attractiveness more strongly elevate sons' fitness (Pischedda & Chippindale, 2006), then assortative mating could lead to sex-specific maternal and paternal effects on offspring that nullify the advantages of SRA. If the relative fitness of offspring of each sex remains similar due to 'symmetrical' effects of maternal and paternal genes on offspring of the opposite sex (but see Connallon & Jakubowski, 2007; Cox & Calsbeek, 2010), a more equitable sex ratio is predicted. This would occur even if attractive males sire sons of above-average fitness when randomly assigned a mate (i.e. even if there is still a general male-biased benefit of sire attractiveness).

Finally, in addition to SRA, there are other, non-mutually exclusive sex allocation responses to the sources of selection described in Table 5, including mate attractiveness. SRA is not the only form of adaptive sex allocation based on changes in investment patterns. There is no evidence in birds that mothers can detect the sex of an embryo and then differentially invest into eggs on this basis (meta-analysis: Rutkowska, Dubiec & Nakagawa, 2014). Mothers can, however, differentially invest into total offspring production (i.e. produce larger broods and/or bigger eggs when mated to an attractive male; Burley, 1986; review: Sheldon, 2000; meta-analysis: Horváthová *et al.*, 2012) and differentially allocate resources towards one sex without any change in the sex ratio or number of offspring per breeding attempt (e.g. Saino *et al.*, 2003). These varied strategic responses are forms of sex allocation that are not always distinguished in the literature, either from each other or from SRA (Carranza, 2002).

(3) Reasons for a small effect: measuring sexual selection is tricky

Although there are clear predictions for adaptive SRA based on male attractiveness (West & Sheldon, 2002), it is often unclear (*i*) what traits actually determine male attractiveness or quality, and (*ii*) if these traits actually affect offspring fitness through male-biased genetic benefits.

Identifying targets of mate choice is a perennial problem in comparative analyses or meta-analyses of sexual selection. What makes a male attractive? Many researchers measure several different male traits that are putatively preferred by females. Thus, even within a single empirical study, there can be several reported relationships between 'attractiveness' and the offspring sex ratio (e.g. Griffith *et al.*, 2003; Bowers *et al.*, 2013). This could be statistically problematic if several male traits are included in a multiple regression as potential predictors of the offspring sex ratio. This was often the case for studies in our data set: many of our effect sizes were taken from generalized linear models (GLMs) or generalized linear mixed-effects models (GLMMs) with several predictors (exclusion of such data would have halved our sample size). The resultant effect size estimates are therefore the effect of the focal trait controlling for its relationships with other measured traits, and for the relationship between those traits and the offspring sex ratio (i.e. we reported partial correlations). For a review of the potential statistical biases that occur when using partial rather than bivariate correlations see Aloe (2015), and references therein. The more strongly the measured traits are correlated the greater the discrepancy between partial and bivariate relationships. In general, however, collinearity between measured traits is modest (otherwise they would not be included as independent predictors in a model). A lack of correlation among measured traits does, however, highlight the likelihood that some of them might contribute little if anything to net attractiveness. More generally, female use of multiple cues for mate choice (review: Candolin, 2003) means that effect sizes for individual traits imperfectly capture the true relationship being tested: that between *net male attractiveness* and the offspring sex ratio.

Interpretation of what constitutes an attractive male trait has also expanded since Burley's original hypothesis was formulated. For example, it has been argued that whether a male is a within-pair or extra-pair mate is informative about his relative attractiveness to females

(reviews: Westneat & Stewart, 2003; but see Forstmeier *et al.*, 2014). In addition, traits used by females for mate choice are not necessarily those that directly influence adaptive SRA, even if there are male-biased genetic benefits. For example, male attractiveness and competiveness are often imperfectly correlated (see Kelly, 2008), or even negatively correlated (Hunt *et al.*, 2009). Even if researchers correctly identify heritable traits that determine male attractiveness, these are not necessarily the ones that females assess for the largest potential SRA fitness gains (e.g. if the marginal fitness gains from being more 'dominant' during fights for territories outweigh those of being more attractive to females when holding a territory, and both attractiveness and dominance are heritable).

Determining the relationship between fitness and attractiveness is a major challenge for two reasons. First, it is difficult to measure fitness, and it is often simply assumed that attractiveness predicts net fitness. But do males bearing attractive traits and experiencing higher mating rates actually have greater net fitness than less attractive males? Studies often report an association between male attractiveness/quality and an increase in some components of reproductive success (e.g. number of mates in a single breeding season) without testing whether natural selection (Hine, McGuigan & Blows, 2011) or even other forms of sexual selection (e.g. sperm competitiveness; Lüpold *et al.*, 2014) negate the sexually selected benefits of traits that elevate attractiveness. Second, the mate attractiveness hypothesis for adaptive SRA assumes that male-derived fitness benefits exist, and are to some extent sex-limited. In the context of genetic benefits this basically means that male attractiveness is heritable. Many studies have reported significant additive genetic variation for male sexual traits (Prokuda & Roff, 2014), but it remains unclear if net attractiveness (and, ideally, net fitness) is heritable. Genetic covariance among traits affecting fitness can result in the absence of genetic variation in the direction of selection on net attractiveness and, ultimately, fitness (Walsh & Blows, 2009).

(4) Updated results and future improvements

Our updated meta-analyses indicate a substantially weaker relationship between measures of mate attractiveness and offspring sex ratios than the $r = 0.205$ reported by West & Sheldon (2002). A decline in effect size over time is a common phenomenon in many research fields, including ecology and evolution (Jennions & Møller, 2002). There are many explanations for a temporal decline, including an early publication bias against weaker effect sizes (i.e. non-significant results) and the inclusion of less appropriate study systems over time (review: Koricheva, Jennions & Lau, 2013). It should be noted, however, that the effect of publication year was not significant in our analyses. Below we discuss other factors potentially contributing to our finding of a smaller effect than previously reported for the relationship between mate attractiveness and SRA.

In general, we argue that the observed small effect sizes are expected for many reasons. Six factors that we discussed are: (*i*) poor control for other selective factors that affect SRA (see Table 5); (*ii*) other forms of conditional sex allocation (e.g. sex-biased provisioning); (*iii*) assortative mating by mate quality; (*iv*) difficulty identifying preferred male traits that directly influence sex allocation – ideally researchers should use measures of *net male attractiveness*; (*v*) uncertainty that attractiveness increases net male fitness; (*vi*) uncertainty that the benefits of mate choice are genetic and greater for sons than daughters. In addition, there are at least four other factors that might lower effect sizes: (*vii*) high measurement error due to logistic challenges (e.g. measuring behavioural traits accurately); (*viii*) the inclusion of inappropriate study species (e.g. species with high fecundity are less likely to skew the sex ratio; for a review and explanation see Frank, 1990); (*ix*) difficulty in obtaining appropriate effect size estimates – we had to use data from multiple regression models that included several prospective measures of sire attractiveness because simple bivariate correlations were rarely available; and (*x*) stochastic noise (e.g. female perceptual errors when assessing males, and constraints on sex-biasing mechanisms).

Aside from considering the problems described above, future studies should broaden their taxonomic scope. Approximately 84% of the studies in our meta-analysis were on birds. This limits the generality of our conclusions, and evaluation of the wider relevance of the SRA mate attractiveness hypothesis. In general, specific sex allocation hypotheses have been tested with a focus on taxa that seem especially amenable (e.g. local mate competition in fig wasps, or helper production in birds and mammals). The disproportionate use of birds to test the SRA mate attractiveness hypothesis might be due to high sexual dimorphism and extravagance of male ornamental traits in many species that suggest strong sexual selection through female choice. More importantly, recent theory suggests that SRA due to mate attractiveness is least likely to occur when there is very strong sexual selection (Fawcett *et al.*, 2011; Booksmythe *et al.*, 2013). Finally, many avian life histories do not conform to key assumptions of mathematical models of sex allocation/adjustment, such as non-overlapping generations, semelparity (or a fixed total investment at every reproductive event), a short and discrete bout of parental investment and investment (potential for control) by one parent only, and no cost of adjustment (review: Komdeur & Pen, 2002). Although some of these limitations will apply to non-avian taxa as well, researchers might wish to reconsider their choice of study species and look to taxa other than birds to test the SRA mate attractiveness hypothesis.

During our attempts to extract effect sizes from papers it became obvious that many studies are guilty of *post hoc* analyses, and fail to clearly specify study goals prior to analysis. Many statistical tests are better seen as exploratory, rather than acting to test a hypothesis with the stated likelihood of type I error (for an approachable case study see Simmons, Nelson & Simonsohn, 2011). There was often an excess of male traits being examined, the inclusion of covariates seemed arbitrary, and the actual statistical test

used (e.g. GLM or contingency table) and whether data were pooled or analysed at the brood level was often not justified. All these problems increase the danger of *P*-hacking (selective collection, manipulation, and/or reporting of data to achieve statistical significance) and erroneous reporting of significant results, which inflates effect size estimates (see Simonsohn, Nelson & Simmons, 2014; Head *et al.*, 2015). They also made it difficult to locate the information needed to extract effect sizes. The SRA hypothesis is straightforward so it is reasonable to expect greater consistency in how researchers test for a relationship between male attractiveness and offspring sex ratio. We mention this not to criticise colleagues (we are guilty of the same 'misconduct' *sensu* Simmons *et al.*, 2011, in some of our own papers), but simply to highlight wider calls by many others (e.g. Garamszegi *et al.*, 2009) for the need to improve standards of reporting in evolutionary ecology studies. It is very possible, for example, that the conclusions of our meta-analysis might have differed if researchers had always used composite measures of male attractiveness (rather than analysing multiple traits) or integrated fitness measures such as female responses, had only controlled for covariates with well-supported biological relevance to their study system and for which they could provide evidence of effects on sex ratios (e.g. level of local mate competition in parasitoids), and had fully reported the direction and magnitude of all tests, regardless of whether or not the effect was significant.

Finally, we should note that there is at least one species where well-replicated experimental studies confirm that mothers adjust the offspring sex ratio in response to male phenotype: head colour in Gouldian finches (Pryke & Griffith, 2009; Pryke *et al.*, 2011). As such, there is evidence that SRA in response to male attractiveness has evolved. In our meta-analysis, however, the small mean effect size and its sensitivity to narrow-sense publication bias (the 'file drawer' problem, where a greater proportion of non-significant than significant results go unpublished; Rosenthal, 1979; Jennions *et al.*, 2013) and/or selective reporting, indicates little support for the hypothesis that such SRA is widespread. Ultimately, the strongest evidence has (and will continue to) come from experiments that involve direct manipulation of male attractiveness and/or random assignment of males to females. To date there are only a handful of appropriate studies (e.g. Burley, 1981, 1986; Sheldon *et al.*, 1999; Pike & Petrie, 2005; Ferree, 2007). In species where it is feasible, the study design of Sato & Karino (2010), whereby a female's social environment is manipulated so that the same male is assessed to be either relatively attractive or unattractive, seems a particularly powerful approach. Once species in which SRA has been shown robustly to occur are identified, the next step is to test whether this is because sons sired by relatively more attractive males are of higher mean fitness than are daughters. Both lines of evidence are needed to test the SRA mate attractiveness hypothesis fully.

Given the many difficulties with the interpretation of observational studies (i.e. inability to predict the sex ratio

with certainty due to potentially confounding factors) we suggest that the most fruitful way forward is to conduct experiments that directly rank males according to their mean attractiveness to females, and then randomly assign females a mate of low or high attractiveness. If these experiments are conducted on species where there is little evidence for direct material assistance by males (i.e. no male parental care or mate provisioning) it should be possible to test the hypothesis that genetic benefits that accrue more strongly to sons than daughters can drive greater sex allocation to males.

V. CONCLUSIONS

(1) The mean effect size for the relationship between male attractiveness and offspring sex ratio based on all trait types was small ($r < 0.10$) but significant. Assuming no publication bias this result was consistent using standard random-effects meta-analysis or multilevel meta-analysis models (including a control for phylogenetic relatedness in a data set restricted to birds).

(2) These significant mean effect size estimates from all traits combined were, however, sensitive to publication bias – due to a possible 'file drawer' effect and/or selective reporting of significant results within studies. Correction for possible bias reduced effect size estimates to non-significant values.

(3) The mean effect sizes for seven specific male trait types were significant only for ornamental male traits and measures of male body size. The mean effect remained significant for body size even after correcting for potential publication bias.

(4) There was moderate heterogeneity in effect size estimates. The unexplained variation between studies indicates that there are moderating variables that have yet not been identified. We detected no effect of our three test moderators: year of publication, experimental *versus* observational study, and laboratory *versus* field study. There was a very weak phylogenetic signal for birds (i.e. related species did not have similar effect sizes).

(5) There was a strong taxonomic bias in our data set. Most studies were on bird species (∼84%). Given the lack of robust evidence in birds, future studies on other groups (invertebrates, fish, reptiles) are needed to shed light on whether the mate attractiveness SRA hypothesis will have any greater support in other taxa.

(6) In sum, our results provide very weak support for the hypothesis that females facultatively adjust the sex ratio of their offspring in response to the attractiveness of their mate.

VII. REFERENCES

Studies marked with * contributed effect sizes to the meta-analyses; studies marked with † are key papers from which forward searches were conducted

References marked with ‡ have been cited within the supporting information

*Abroe, B., Garvin, J. C., Pedersen, M. C., Whittingham, L. A. & Dunn, P. O. (2007). Brood sex ratios are related to male size but not to attractiveness in common yellowthroats (*Geothlypis trichas*). *Auk* **124**, 176–184.

*Addison, B., Kitaysky, A. S. & Hipfner, J. M. (2008). Sex allocation in a monomorphic seabird with a single-egg clutch: test of the environment, mate quality, and female condition hypotheses. *Behavioral Ecology and Sociobiology* **63**, 135–141.

Aloe, A. M. (2015). Inaccuracy of regression results in replacing bivariate correlations. *Research Synthesis Methods* **6**, 21–27.

Alonzo, S. H. & Sinervo, B. (2007). The effect of sexually antagonistic selection on adaptive sex ratio allocation. *Evolutionary Ecology Research* **9**, 1097–1117.

Andersson, M. (1994). *Sexual Selection*. Princeton University Press, Princeton.

*Aparicio, J. M. & Cordero, P. J. (2001). The effects of the minimum threshold condition for breeding on offspring sex-ratio adjustment in the lesser kestrel. *Evolution* **55**, 1188–1197.

Barclay, R. M. R. (2012). Variable variation: annual and seasonal changes in offspring sex ratio in a bat. *PLoS ONE* **7**, e36344.

*Bell, S. C., Owens, I. P. F. & Lord, A. M. (2014). Quality of breeding territory mediates the influence of paternal quality on sex ratio bias in a free-living bird population. *Behavioral Ecology* **25**, 352–358.

*Benito, M. M., Schielzeth, H., González-Sólis, J. & Becker, P. H. (2013). Sex ratio adjustments in common terns: influence of mate condition and maternal experience. *Journal of Avian Biology* **44**, 179–188.

Berkeley, E. V. & Linklater, W. L. (2010). Annual and seasonal rainfall may influence progeny sex ratio in the black rhinoceros. *South African Journal of Wildlife Research* **40**, 53–57.

Blackburn, G. S., Albert, A. Y. K. & Otto, S. P. (2010). The evolution of sex ratio adjustment in the presence of sexually antagonistic selection. *American Naturalist* **176**, 264–275.

*Blanchard, P., Hanuise, N., Dano, S. & Weimerskirch, H. (2007). Offspring sex ratio in relation to parental structural size and body condition in the long-lived wandering albatross (*Diomedea exulans*). *Behavioral Ecology and Sociobiology* **61**, 767–773.

Booksmythe, I., Schwanz, L. E. & Kokko, H. (2013). The complex interplay of sex allocation and sexual selection. *Evolution* **67**, 673–678.

Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. John Wiley & Sons, Chichester.

*Bowers, E. K., Munclinger, P., Bures, S., Kucerova, L., Nadvornik, P. & Krist, M. (2013). Cross-fostering eggs reveals that female collared flycatchers adjust clutch sex ratios according to parental ability to invest in offspring. *Molecular Ecology* **22**, 215–228.

Bowers, E. K., Thompson, C. F. & Sakaluk, S. K. (2014). Offspring sex ratio varies with clutch size for female house wrens induced to lay supernumerary eggs. *Behavioral Ecology* **25**, 165–171.

*Bradbury, R. B., Cotton, P. A., Wright, J. & Griffiths, R. (1997). Nestling sex ratio in the European Starling *Sturnus vulgaris*. *Journal of Avian Biology* **28**, 255–258.

Brooks, R. & Kemp, D. J. (2001). Can older males deliver the good genes? *Trends in Ecology & Evolution* **16**, 308–316.

Brown, G. R. & Silk, J. B. (2002). Reconsidering the null hypothesis: is maternal rank associated with birth sex ratios in primate groups? *Proceedings of the National Academy of Sciences of the United States of America* **99**, 11252–11255.

†Burley, N. (1981). Sex-ratio manipulation and selection for attractiveness. *Science* **211**, 721–722.

*Burley, N. (1986). Sex-ratio manipulation in color-banded populations of zebra finches. *Evolution* **40**, 1191–1206.

*Calsbeek, R. & Sinervo, B. (2004). Within-clutch variation in offspring sex determined by differences in sire body size: cryptic mate choice in the wild. *Journal of Evolutionary Biology* **17**, 464–470.

Cameron, E. Z. (2004). Facultative adjustment of mammalian sex ratios in support of the Trivers-Willard hypothesis: evidence for a mechanism. *Proceedings of the Royal Society of London Series B* **271**, 1723–1728.

Candolin, U. (2003). The use of multiple cues in mate choice. *Biological Reviews* **78**, 575–595.

Carranza, J. (2002). What did Trivers and Willard really predict? *Animal Behaviour* **63**, F1–F3.

Cassey, P., Ewen, J. G., Blackburn, T. M. & Møller, A. P. (2004). A survey of publication bias within evolutionary ecology. *Proceedings of the Royal Society of London Series B* **271**, S451–S454.

Cassey, P., Ewen, J. G. & Møller, A. P. (2006). Revised evidence for facultative sex ratio adjustment in birds: a correction. *Proceedings of the Royal Society of London Series B* **273**, 3129–3130.

Chamberlain, S. A., Hovick, S. M., Dibble, C. J., Rasmussen, N. L., Van Allen, B. G., Maitner, B. S., Ahern, J. R., Bell-Dereske, L. P., Roy, C. L., Meza-Lopez, M., Carrillo, J., Siemann, E., Lajeunesse, M. J. & Whitney, K. D. (2012). Does phylogeny matter? Assessing the impact of phylogenetic information in ecological meta-analysis. *Ecology Letters* **15**, 627–636.

Charnov, E. L. (1979). The genetical evolution of patterns of sexuality: Darwinian fitness. *American Naturalist* **113**, 465–480.

Cichon, M., Dubiec, A. & Stoczko, M. (2003). Laying order and offspring sex in blue tits *Parus caeruleus*. *Journal of Avian Biology* **34**, 355–359.

Clutton-Brock, T. H. (1986). Sex ratio variation in birds. *Ibis* **128**, 317–329.

Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. (1984). Maternal dominance, breeding success and birth sex-ratios in red deer. *Nature* **308**, 358–360.

*Cockburn, A. & Double, M. C. (2008). Cooperatively breeding superb fairy-wrens show no facultative manipulation of offspring sex ratio despite plausible benefits. *Behavioral Ecology and Sociobiology* **62**, 681–688.

Cockburn, A., Scott, M. P. & Dickman, C. R. (1985). Sex-ratio and intrasexual kin competition in mammals. *Oecologia* **66**, 427–429.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Second Edition. Lawrence Erlbaum, Hillsdale.

Connallon, T. & Clark, A. G. (2014). Evolutionary inevitability of sexual antagonism. *Proceedings of the Royal Society of London Series B* **281**, 20132013.

Connallon, T. & Jakubowski, E. (2007). Association between sex ratio distortion and sexually antagonistic fitness consequences of female choice. *Evolution* **63**, 2179–2183.

*Correa, S. M., Horan, C. M., Johnson, P. A. & Adkins-Regan, E. (2011). Copulatory behaviors and body condition predict post-mating female hormone concentrations, fertilization success, and primary sex ratios in Japanese quail. *Hormones and Behavior* **59**, 556–564.

*Cotton, A. J., Földvari, M., Cotton, S. & Pomiankowski, A. (2014). Male eyespan size is associated with meiotic drive in wild stalk-eyed flies (*Teleopsis dalmanni*). *Heredity* **112**, 363–369.

Cotton, S., Fowler, K. & Pomiankowski, A. (2004). Do sexual ornaments exhibit heightened condition-dependent expression as predicted by the handicap hypothesis? *Proceedings of the Royal Society of London Series B* **271**, 771–783.

Cox, R. M. & Calsbeek, R. (2010). Cryptic sex-ratio bias provides indirect genetic benefits despite sexual conflict. *Science* **328**, 92–94.

*Cox, R. M., Duryea, M. C., Najarro, M. & Calsbeek, R. (2011). Paternal condition drives progeny sex-ratio bias in a lizard that lacks parental care. *Evolution* **65**, 220–230.

*‡Cruickshank, T. & Wade, M. J. (2012). Maternal adjustment of the sex ratio in broods of the broad-horned flour beetle, *Gnathocerus cornutus*. *Integrative and Comparative Biology* **52**, 100–107.

*Cunningham, E. J. A. & Russell, A. F. (2001). Sex differences in avian yolk hormone levels – Reply. *Nature* **412**, 498–499.

*Debiak, A. L., McCormick, D. L., Kaplan, J. D., Tischler, K. B. & Lindsay, A. R. (2014). A molecular genetic assessment of sex ratios from pre-fledged juvenile and migrating adult common loons (*Gavia immer*). *Waterbirds* **37**, 6–15.

*Delhey, K., Peters, A., Johnsen, A. & Kempenaers, B. (2007). Brood sex ratio and male UV ornamentation in blue tits (*Cyanistes caeruleus*): correlational evidence and an experimental test. *Behavioral Ecology and Sociobiology* **61**, 853–862.

*Delmore, K. E., Kleven, O., Laskemoen, T., Crowe, S. A., Lifjeld, J. T. & Robertson, R. J. (2008). Sex allocation and parental quality in tree swallows. *Behavioral Ecology* **19**, 1243–1249.

*Dietrich-Bischoff, V., Schmoll, T., Winkel, W., Krackow, S. & Lubjuhn, T. (2006). Extra-pair paternity, offspring mortality and offspring sex ratio in the socially monogamous coal tit (*Parus ater*). *Behavioral Ecology and Sociobiology* **60**, 563–571.

*Dolan, A. C., Murphy, M. T., Redmond, L. J. & Duffield, D. (2009). Maternal characteristics and the production and recruitment of sons in the eastern kingbird (*Tyrannus tyrannus*). *Behavioral Ecology and Sociobiology* **63**, 1527–1537.

*Dowling, D. K. & Mulder, R. A. (2006). Combined influence of maternal and paternal quality on sex allocation in red-capped robins. *Journal of Evolutionary Biology* **19**, 440–449.

*Dreiss, A., Richard, M., Moyen, F., White, J., Møller, A. P. & Danchin, E. (2006). Sex ratio and male sexual characters in a population of blue tits, *Parus caeruleus*. *Behavioral Ecology* **17**, 13–19.

*Drullion, D. & Dubois, F. (2011). Neighbours' breeding success and the sex ratio of their offspring affect the mate preferences of female zebra finches. *PLoS ONE* **6**, e29737.

*Du, B. & Lu, X. (2010). Sex allocation and paternity in a cooperatively breeding passerine: evidence for the male attractiveness hypothesis? *Behavioral Ecology and Sociobiology* **64**, 1631–1639.

‡Dubois, N. S., Kennedy, E. D. & Getty, T. (2006). Surplus nest boxes and the potential for polygyny affect clutch size and offspring sex ratio in house wrens. *Proceedings of the Royal Society of London Series B* **273**, 1751–1757.

Dunn, P. O., Whittingham, L. A. & Pitcher, T. E. (2001). Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution* **55**, 161–175.

Düsing, C. (1884). *Die Regulierung des Geschlechtsverhältnisses bei der Vermehrung der Menschen, Tiere und Pflanzen*. Fischer, Jena.

Duval, S. & Tweedie, R. (2000*a*). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association* **95**, 89–98.

Duval, S. & Tweedie, R. (2000*b*). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463.

‡Edvardsson, M. & Arnqvist, G. (2005). The effects of copulatory courtship on differential allocation in the red flour beetle *Tribolium castaneum*. *Journal of Insect Behavior* **18**, 313–322.

Egger, M., Davey Smith, G., Schneider, M. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.

*†Ellegren, H., Gustafsson, L. & Sheldon, B. C. (1996). Sex ratio adjustment in relation to paternal attractiveness in a wild bird population. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 11723–11728.

Ewen, J. G., Cassey, P. & Møller, A. P. (2004). Facultative primary sex ratio variation: a lack of evidence in birds? *Proceedings of the Royal Society of London Series B* **271**, 1277–1282.

Fairbairn, D. J., Blanckenhorn, W. U. & Székely, T. (eds) (2007). *Sex, Size and Gender Roles: Evolutionary Studies of Sexual Size Dimorphism*. Oxford University Press, Oxford.

*Fargallo, J. A., Dávila, J. A., Potti, J., de León, A. & Polo, V. (2004). Nest size and hatchling sex ratio in chinstrap penguins. *Polar Biology* **27**, 339–343.

Farris, H. E. & Ryan, M. J. (2011). Relative comparisons of call parameters enable auditory grouping in frogs. *Nature Communications* **2**, 410.

†Fawcett, T. W., Kuijper, B., Pen, I. & Weissing, F. J. (2007). Should attractive males have more sons? *Behavioral Ecology* **18**, 71–80.

Fawcett, T. W., Kuijper, B., Weissing, F. J. & Pen, I. (2011). Sex-ratio control erodes sexual selection, revealing evolutionary feedback from adaptive plasticity. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 15925–15930.

*Ferree, E. D. (2007). White tail plumage and brood sex ratio in dark-eyed juncos (*Junco hyemalis thurberi*). *Behavioral Ecology and Sociobiology* **62**, 109–117.

Firman, R. C. (2011). Polyandrous females benefit by producing sons that achieve high reproductive success in a competitive environment. *Proceedings of the Royal Society of London Series B* **278**, 2823–2831.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Forstmeier, W., Nakagawa, S., Griffith, S. C. & Kempenaers, B. (2014). Female extra-pair mating: adaptation or genetic constraint? *Trends in Ecology & Evolution* **29**, 456–464.

†Frank, S. A. (1990). Sex allocation theory for birds and mammals. *Annual Review of Ecology and Systematics* **21**, 13–55.

Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton University Press, Princeton.

Garamszegi, L. Z., Calhim, S., Dochtermann, N. A., Hegyi, G., Hurd, P. L., Jørgensen, C., Kutsukake, N., Lajeunesse, M. J., Pollatд, K. A., Schielzeth, H., Symonds, M. R. E. & Nakagawa, S. (2009). Changing philosophies and tools for statistical inferences in behavioral ecology. *Behavioral Ecology* **20**, 1363–1375.

Gildersleeve, K., Haselton, M. G. & Fales, M. R. (2014). Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychological Bulletin* **140**, 1205–1259.

Godfray, H. C. J. (1994). *Parasitoids: Behavioural and Evolutionary Ecology*, pp. 151–210. Princeton University Press, Princeton.

Gowaty, P. A. & Lennartz, M. R. (1985). Sex-ratios of nestling and fledgling red-cockaded woodpeckers (*Picoides borealis*) favor males. *The American Naturalist* **126**, 347–353.

Graham, E. B., Caro, S. P. & Sockman, K. W. (2011). Change in offspring sex ratio over a very short season in Lincoln's sparrows: the potential role of bill development. *Journal of Field Ornithology* **82**, 44–51.

‡Grana, S. C., Sakaluk, S. K., Bowden, R. M., Doellman, M. A., Vogel, L. A. & Thompson, C. F. (2012). Reproductive allocation in female house wrens is not influenced by experimentally altered male attractiveness. *Behavioral Ecology and Sociobiology* **66**, 1247–1258.

*Green, D. J. (2002). Pair bond duration influences paternal provisioning and the primary sex ratio of brown thornbill broods. *Animal Behaviour* **64**, 791–800.

Griffin, A. S., Sheldon, B. C. & West, S. A. (2005). Cooperative breeders adjust offspring sex ratios to produce helpful helpers. *American Naturalist* **166**, 628–632.

*Griffith, S. C., Örnborg, J., Russell, A. F., Andersson, S. & Sheldon, B. C. (2003). Correlations between ultraviolet coloration, overwinter survival and offspring sex ratio in the blue tit. *Journal of Evolutionary Biology* **16**, 1045–1054.

*Grindstaff, J. L., Buerkle, C. A., Casto, J. M., Nolan, V. Jr. & Ketterson, E. D. (2001). Offspring sex ratio is unrelated to male attractiveness in dark-eyed juncos (*Junco hyemalis*). *Behavioral Ecology and Sociobiology* **50**, 312–316.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* **33**, 1–22.

Hadfield, J. D. & Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology* **23**, 494–508.

Hamilton, W. D. (1967). Extraordinary sex ratios. *Science* **156**, 477–488.

Head, M.L., Holman, L., Lanfear, R., Kahn, A. T. & Jennions, M D. (2015). The extent and consequences of P-hacking in science. *Public Library of Science, Biology* **13**: e1002106.

Hedges, L. & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Academic Press, New York.

*Helle, S., Laaksonen, T., Adamsson, A., Paranko, J. & Huitu, O. (2008). Female field voles with high testosterone and glucose levels produce male-biased litters. *Animal Behaviour* **75**, 1031–1039.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal* **327**, 557–560.

Hine, E., McGuigan, K. & Blows, M. W. (2011). Natural selection stops the evolution of male attractiveness. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 3659–3664.

*Hjernquist, M. B., Hjernquist, K. A. T., Forsman, J. T. & Gustafsson, L. (2009). Sex allocation in response to local resource competition over breeding territories. *Behavioral Ecology* **20**, 335–339.

Hoelzer, G. A. (1989). The good parent process of sexual selection. *Animal Behaviour* **38**, 1067–1078.

Holman, L. & Kokko, H. (2013). The consequences of polyandry for population viability, extinction risk and conservation. *Philosophical Transactions of the Royal Society of London, Series B* **368**, 20120053.

Horváthová, T., Nakagawa, S. & Uller, T. (2012). Strategic female reproductive investment in response to male attractiveness in birds. *Proceedings of the Royal Society of London Series B* **279**, 163–170.

*House, C. M., Simmons, L. W., Kotiaho, J. S., Tomkins, J. L. & Hunt, J. (2011). Sex ratio bias in the dung beetle *Onthophagus taurus*: adaptive allocation or sex-specific offspring mortality? *Evolutionary Ecology* **25**, 363–372.

Housworth, E. A., Martins, E. P. & Lynch, M. (2004). The phylogenetic mixed model. *The American Naturalist* **163**, 84–96.

Hunt, J., Breuker, C. J., Sadowski, J. A. & Moore, A. J. (2009). Male-male competition, female mate choice and their interaction: determining total sexual selection. *Journal of Evolutionary Biology* **22**, 13–26.

*Husby, A., Saether, B. E., Jensen, H. & Ringsby, T. H. (2006). Causes and consequences of adaptive seasonal sex ratio variation in house sparrows. *Journal of Animal Ecology* **75**, 1128–1139.

‡Janssen, K., Erikstad, K. E. & Bensch, S. (2006). Offspring sex ratio allocation in the parasitic jaeger: selection for pale females and melanic males? *Behavioral Ecology* **17**, 236–245.

Jennions, M. D., Kahn, A. T., Kelly, C. D. & Kokko, H. (2012). Meta-analysis and sexual selection: past studies and future possibilities. *Evolutionary Ecology* **26**, 1119–1151.

Jennions, M. D., Lortie, C. J., Rosenberg, M. S. & Rothstein, H. R. (2013). Publication and related biases. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. Koricheva, J. Gurevitch and K. Mengersen), pp. 207–236. Princeton University Press, Princeton.

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

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

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature* **491**, 444–448.

*Johnson, L. S., Thompson, C. F., Sakaluk, S. K., Neuhäuser, M., Johnson, B. G. P., Soukup, S. S., Forsythe, S. J. & Masters, B. S. (2009). Extra-pair young in house wren broods are more likely to be male than female. *Proceedings of the Royal Society of London Series B* **276**, 2285–2289.

Jones, K. S., Nakagawa, S. & Sheldon, B. C. (2009). Environmental sensitivity in relation to size and sex in birds: meta-regression analysis. *American Naturalist* **174**, 122–133.

*Karino, K., Kobayashi, M. & Orita, K. (2006). Adaptive offspring sex ratio depends on male tail length in the guppy. *Ethology* **112**, 1050–1055.

*Karino, K. & Sato, A. (2009). Male-biased sex ratios in offspring of attractive males in the guppy. *Ethology* **115**, 682–690.

*Katsuki, M., Harano, T., Miyatake, T., Okada, K. & Hosken, D. (2012). Intralocus sexual conflict and offspring sex ratio. *Ecology Letters* **15**, 193–197.

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

*Kempenaers, B., Verheyen, G. R. & Dhondt, A. A. (1997). Extrapair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics, and offspring quality. *Behavioral Ecology* **8**, 481–492.

‡Kie, J. G., Johnson, B. K., Noyes, J. H., Williams, C. L., Dick, B. L., Rhodes, O. E., Stussy, R. J. & Bowyer, R. T. (2013). Reproduction in North American elk *Cervus elaphus*: paternity of calves sired by males of mixed age classes. *Wildlife Biology* **19**, 302–310.

Kissner, K. J., Weatherhead, P. J. & Francis, C. M. (2003). Sexual size dimorphism and timing of spring migration in birds. *Journal of Evolutionary Biology* **16**, 154–162.

*Kölliker, M., Heeb, P., Werner, I., Mateman, A. C., Lessells, C. M. & Richner, H. (1999). Offspring sex ratio is related to male body size in the great tit (*Parus major*). *Behavioral Ecology* **10**, 68–72.

Komdeur, J. & Pen, I. (2002). Adaptive sex allocation in birds: the complexities of linking theory and practice. *Philosophical Transactions of the Royal Society of London, Series B* **357**, 373–380.

Koricheva, J., Jennions, M. D. & Lau, J. (2013). Temporal trends in effect sizes: causes, detection and implication. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. Koricheva, J. Gurevitch and K. Mengersen), pp. 237–254. Princeton University Press, Princeton.

*Korpimäki, E., May, C. A., Parkin, D. T., Wetton, J. H. & Wiehn, J. (2000). Environmental- and parental condition-related variation in sex ratio of kestrel broods. *Journal of Avian Biology* **31**, 128–134.

*Korsten, P., Lessells, C. M., Mateman, A. C., van der Velde, M. & Komdeur, J. (2006). Primary sex ratio adjustment to experimentally reduced male UV attractiveness in blue tits. *Behavioral Ecology* **17**, 539–546.

*Kraaijeveld, K., Ming, M., Komdeur, J. & Mulder, R. A. (2007). Offspring sex ratios in relation to mutual ornamentation and extra-pair paternity in the Black Swan *Cygnus atratus*. *Ibis* **149**, 79–85.

Krackow, S. (1995). Potential mechanisms for sex ratio adjustment in mammals and birds. *Biological Reviews* **70**, 225–241.

Kruuk, L. E. B., Clutton-Brock, T. H., Albon, S. D., Pemberton, J. M. & Guinness, F. E. (1999). Population density affects sex ratio variation in red deer. *Nature* **399**, 459–461.

*Lee, J.-W. & Hatchwell, B. J. (2011). Do females adjust brood sex ratio according to males' genetic structures in a gregarious passerine, the vinous-throated parrotbill *Paradoxornis webbianus*? *Journal of Avian Biology* **42**, 447–453.

*Leech, D. I., Hartley, I. R., Stewart, I. R. K., Griffith, S. C. & Burke, T. (2001). No effect of parental quality or extrapair paternity on brood sex ratio in the blue tit (*Parus caeruleus*). *Behavioral Ecology* **12**, 674–680.

Leimar, O. (1996). Life-history analysis of the Trivers and Willard sex-ratio problem. *Behavioral Ecology* **7**, 316–325.

*Leitner, S., Marshall, R. C., Leisler, B. & Catchpole, C. K. (2006). Male song quality, egg size and offspring sex in captive canaries (*Serinus canaria*). *Ethology* **112**, 554–563.

*Leonard, M. L. & Weatherhead, P. J. (1996). Dominance rank and offspring sex ratios in domestic fowl. *Animal Behaviour* **51**, 725–731.

*Lessells, C. M., Mateman, A. C. & Visser, J. (1996). Great tit hatchling sex ratios. *Journal of Avian Biology* **27**, 135–142.

Lezalova, R., Tkadlek, E., Obornik, M., Simek, J. & Honza, M. (2005). Should males come first? The relationship between offspring hatching order and sex in the black-headed gull *Larus ridibundus*. *Journal of Avian Biology* **36**, 478–483.

‡Limbourg, T., Mateman, A. C., Andersson, S. & Lessells, C. M. (2004). Female blue tits adjust parental effort to manipulated male UV attractiveness. *Proceedings of the Royal Society of London Series B* **271**, 1903–1908.

*Limbourg, T., Mateman, A. C. & Lessells, C. M. (2013). Parental care and UV coloration in blue tits: opposite correlations in males and females between provisioning rate and mate's coloration. *Journal of Avian Biology* **44**, 17–26.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical Meta-Analysis*. Sage Publishing, Beverly Hills.

Lislevand, T., Byrkjedal, I., Borge, T. & Saetre, G. P. (2005). Egg size in relation to sex of embryo, brood sex ratios and laying sequence in northern lapwings (*Vanellus vanellus*). *Journal of Zoology* **267**, 81–87.

*Long, T. A. F. & Pischedda, A. (2005). Do female *Drosophila melanogaster* adaptively bias offspring sex ratios in relation to the age of their mate? *Proceedings of the Royal Society of London Series B* **272**, 1781–1787.

‡Loyau, A. & Lacroix, F. (2010). Watching sexy displays improves hatching success and offspring growth through maternal allocation. *Proceedings of the Royal Society of London Series B* **277**, 3453–3460.

‡Loyau, A., Saint Jalme, M., Mauget, R. & Sorci, G. (2007). Male sexual attractiveness affects the investment of maternal resources into the eggs in peafowl (*Pavo cristatus*). *Behavioral Ecology and Sociobiology* **61**, 1043–1052.

*Lu, X., Zeng, X. & Du, B. (2013). Body attributes of both parents jointly affect offspring sex allocation in a socially monogamous, size-monomorphic passerine. *Current Zoology* **59**, 271–277.

Lüpold, S., Tomkins, J. L., Simmons, L. W. & Fitzpatrick, J. L. (2014). Female monopolization mediates the relationship between pre- and post-copulatory sexual traits. *Nature Communications* **5**, 3184.

Maestripieri, D. (2002). Maternal dominance rank and age affect offspring sex ratio in pigtail macaques. *Journal of Mammalogy* **83**, 563–568.

*Magrath, M. J. L., Green, D. J. & Komdeur, J. (2002). Sex allocation in the sexually monomorphic fairy martin. *Journal of Avian Biology* **33**, 260–268.

Mengersen, K., Jennions, M. D. & Schmid, C. H. (2013). Statistical models for the meta-analysis of nonindependent data. In *Handbook of Meta-Analysis for Ecology and Evolution* (eds J. Koricheva, J. Gurevitch and K. Mengersen), pp. 255–283. Princeton University Press, Princeton.

Meunier, J., West, S. A. & Chapuisat, M. (2008). Split sex ratios in the social Hymenoptera: a meta-analysis. *Behavioral Ecology* **19**, 382–390.

Møller, A. P. & Birkhead, T. (1994). The evolution of plumage brightness in birds is related to extrapair paternity. *Evolution* **48**, 1089–1100.

Møller, A. P. & Jennions, M. D. (2001). How important are direct fitness benefits of sexual selection? *Naturwissenschaften* **88**, 401–415.

Møller, A. P. & Ninni, P. (1998). Sperm competition and sexual selection: a meta-analysis of paternity studies of birds. *Behavioral Ecology and Sociobiology* **43**, 345–358.

‡Moore, A. J., Gowaty, P. A., Wallin, W. G. & Moore, P. J. (2001). Sexual conflict and the evolution of female mate choice and male social dominance. *Proceedings of the Royal Society of London Series B* **268**, 517–523.

Murtaugh, P. A. (2002). Journal quality, effect size, and publication bias in meta-analysis. *Ecology* **83**, 1162–1166.

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

Nakagawa, S. & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology* **26**, 1253–1274.

*Neto, J. M., Hansson, B. & Hasselquist, D. (2011). Sex allocation in Savi's warblers *Locustella luscinioides*: multiple factors affect seasonal trends in brood sex ratios. *Behavioral Ecology and Sociobiology* **65**, 297–304.

‡Neville, K. J., Vallender, R. & Robertson, R. J. (2008). Nestling sex ratio of golden-winged warblers *Vermivora chrysoptera* in an introgressed population. *Journal of Avian Biology* **39**, 599–604.

Nevison, C. (1997). Social rank and birth sex ratios in rhesus monkeys (*Macaca mulatta*). *Applied Animal Behaviour Science* **51**, 285–291.

*Nishiumi, I. (1998). Brood sex ratio is dependent on female mating status in polygynous great reed warblers. *Behavioral Ecology and Sociobiology* **44**, 9–14.

*Oddie, K. R. & Reim, C. (2002). Egg sex ratio and paternal traits: using within-individual comparisons. *Behavioral Ecology* **13**, 503–510.

‡Olsson, M., Madsen, T., Uller, T., Wapstra, E. & Ujvari, B. (2005a). The role of Haldane's rule in sex allocation. *Evolution* **59**, 221–225.

*Olsson, M., Wapstra, E. & Uller, T. (2005b). Differential sex allocation in sand lizards: bright males induce daughter production in a species with heteromorphic sex chromosomes. *Biology Letters* **1**, 378–380.

Palmer, A. R. (2000). Quasireplication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annual Review of Ecology and Systematics* **31**, 441–480.

*Pariser, E. C., Gilbert, L., Hazon, N., Arnold, K. E. & Graves, J. E. (2012). Mind the gap: the ratio of yolk androgens and antioxidants varies between sons and daughters dependent on paternal attractiveness. *Behavioral Ecology and Sociobiology* **66**, 519–527.

*Parker, T. H. (2005). No evidence for adaptive differential sex allocation in red junglefowl (*Gallus gallus*). *The Auk* **122**, 1161–1168.

Parker, T. H. (2013). What do we really know about the signaling role of plumage colour in blue tits? A case study of impediments to progress in evolutionary biology. *Biological Reviews* **88**, 511–536.

Pen, I. & Weissing, F. J. (2000). Sexual selection and the sex ratio: an ESS analysis. *Selection* **1**, 111–121.

Pereira, T. V. & Ioannidis, J. P. A. (2011). Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. *Journal of Clinical Epidemiology* **64**, 1060–1069.

*Perlut, N. G., Travis, S. E., Dunbar, C. A., Strong, A. M. & Wright, D. M. (2014). Nestling sex ratios do not support long-term parity in two species with different life-history strategies. *Auk* **131**, 224–234.

‡Petrie, M., Schwabl, H., Brande-Lavridsen, N. & Burke, T. (2001). Sex differences in avian yolk hormone levels. *Nature* **412**, 498.

*Pike, T. W. & Petrie, M. (2005). Offspring sex ratio is related to paternal train elaboration and yolk corticosterone in peafowl. *Biology Letters* **1**, 204–207.

Pischedda, A. & Chippindale, A. (2006). Intra-locus sexual conflict diminishes the benefits of sexual selection. *Public Library of Science Biology* **4**, 2099–2103.

*Polo, V., Veiga, J. P., Cordero, P. J., Viñuela, J. & Monaghan, P. (2004). Female starlings adjust primary sex ratio in response to aromatic plants in the nest. *Proceedings of the Royal Society of London Series B* **271**, 1929–1933.

Pomiankowski, A., Iwasa, Y. & Nee, S. (1991). The evolution of costly mate preferences I. Fisher and biased mutation. *Evolution* **45**, 1422–1430.

*Potvin, D. A. & MacDougall-Shackleton, E. A. (2010). Paternal song complexity predicts offspring sex ratios close to fledging, but not hatching, in song sparrows. *Wilson Journal of Ornithology* **122**, 146–152.

Prokop, Z. M., Michalczyk, L., Drobniak, S. M., Herdegen, M. & Radwan, J. (2012). Meta-analysis suggests choosy females get sexy sons more than "good genes". *Evolution* **66**, 2665–2673.

Prokuda, A. & Roff, D. (2014). The quantitative genetics of sexually selected traits, preferred traits and preference: a meta-analysis. *Journal of Evolutionary Biology* **27**, 2283–2296.

Pryke, S. R. & Griffith, S. C. (2009). Genetic incompatibility drives sex allocation and maternal investment in a polymorphic finch. *Science* **323**, 1605–1607.

Pryke, S. R. & Rollins, L. A. (2012). Mothers adjust offspring sex to match the quality of the rearing environment. *Proceedings of the Royal Society of London Series B* **279**, 4051–4057.

Pryke, S. R., Rollins, L. A., Buttemer, W. A. & Griffith, S. C. (2011). Maternal stress to partner quality is linked to adaptive offspring sex ratio adjustment. *Behavioral Ecology* **22**, 717–722.

*Questiau, S., Escaravage, N., Eybert, M.-C. & Taberlet, P. (2000). Nestling sex ratios in a population of bluethroats *Luscinia svecica* inferred from AFLP (TM) analysis. *Journal of Avian Biology* **31**, 8–14.

‡Radford, A. N. & Blakey, J. K. (2000a). Intensity of nest defence is related to offspring sex ratio in the great tit *Parus major*. *Proceedings of the Royal Society of London Series B* **267**, 535–538.

*Radford, A. N. & Blakey, J. K. (2000b). Is variation in brood sex ratios adaptive in the great tit (*Parus major*)? *Behavioral Ecology* **11**, 294–298.

*Ramsay, S. M., Mennill, D. J., Otter, K. A., Ratcliffe, L. M. & Boag, P. T. (2003). Sex allocation in black-capped chickadees *Poecile atricapilla*. *Journal of Avian Biology* **34**, 134–139.

*Rathburn, M. K. & Montgomerie, R. (2005). Offspring sex ratios correlate with pair-male condition in a cooperatively breeding fairy-wren. *Behavioral Ecology* **16**, 41–47.

Rice, W. R. (2013). Nothing in genetics makes sense except in light of genomic conflict. *Annual Review of Ecology, Evolution, and Systematics* **44**, 217–237.

Rice, W. R. & Chippindale, A. K. (2001). Intersexual ontogenetic conflict. *Journal of Evolutionary Biology* **14**, 685–693.

‡Røed, K. H., Holand, Ø., Mysterud, A., Tverdal, A., Kumpula, J. & Nieminen, M. (2007). Male phenotypic quality influences offspring sex ratio in a polygynous ungulate. *Proceedings of the Royal Society of London Series B* **274**, 727–733.

Romano, A., Ambrosini, R., Caprioli, M., Bonisoli-Alquati, A. & Saino, N. (2012). Secondary sex ratio covaries with demographic trends and ecological conditions in the barn swallow. *Evolutionary Ecology* **26**, 1041–1053.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin* **86**, 638–641.

*Rosivall, B., Török, J., Hasselquist, D. & Bensch, S. (2004). Brood sex ratio adjustment in collared flycatchers (*Ficedula albicollis*): results differ between populations. *Behavioral Ecology and Sociobiology* **56**, 346–351.

*Roulin, A., Altwegg, R., Jensen, H., Steinsland, I. & Schaub, M. (2010). Sex-dependent selection on an autosomal melanic female ornament promotes the evolution of sex ratio bias. *Ecology Letters* **13**, 616–626.

Rutkowska, J., Dubiec, A. & Nakagawa, S. (2014). All eggs are made equal: meta-analysis of egg sexual size dimorphism in birds. *Journal of Evolutionary Biology* **27**, 153–160.

*Rutstein, A. N., Gilbert, L., Slater, P. J. B. & Graves, J. A. (2004). Mate attractiveness and primary resource allocation in the zebra finch. *Animal Behaviour* **68**, 1087–1094.

*Rutstein, A. N., Gorman, H. E., Arnold, K. E., Gilbert, L., Orr, K. J., Adam, A., Nager, R. & Graves, J. A. (2005). Sex allocation in response to paternal attractiveness in the zebra finch. *Behavioral Ecology* **16**, 763–769.

Rutz, C. (2012). Brood sex ratio varies with diet composition in a generalist raptor. *Biological Journal of the Linnean Society* **105**, 937–951.

‡Saether, B.-E., Solberg, E. J., Heim, M., Stacy, J. E., Jakobsen, K. S. & Olstad, R. (2004). Offspring sex ratio in moose *Alces alces* in relation to paternal age: an experiment. *Wildlife Biology* **10**, 51–57.

*Saino, N., Ambrosini, R., Martinelli, R., Calza, S., Møller, A. P. & Pilastro, A. (2002). Offspring sexual dimorphism and sex-allocation in relation to parental age and paternal ornamentation in the barn swallow. *Molecular Ecology* **11**, 1533–1544.

*Saino, N., Ellegren, H. & Møller, A. P. (1999). No evidence for adjustment of sex allocation in relation to paternal ornamentation and paternity in barn swallows. *Molecular Ecology* **8**, 399–406.

Saino, N., Martinelli, R. & Romano, M. (2008). Ecological and phenological covariates of offspring sex ratio in barn swallows. *Evolutionary Ecology* **22**, 659–674.

*Saino, N., Romano, M., Ferrari, R. P., Martinelli, R. & Møller, A. P. (2003). Maternal antibodies but not carotenoids in barn swallow eggs covary with embryo sex. *Journal of Evolutionary Biology* **16**, 516–522.

Sardell, R. J. & DuVal, E. H. (2014). Differential allocation in a lekking bird: females lay larger eggs and are more likely to have male chicks when they mate with less related males. *Proceedings of the Royal Society of London Series B* **281**, 20132386.

*Sato, A. & Karino, K. (2010). Female control of offspring sex ratios based on male attractiveness in the guppy. *Ethology* **116**, 524–534.

‡Sato, A. & Karino, K. (2011). Offspring sex ratio produced by female guppies in the wild correlates with sexual ornaments of their sons. *Journal of Ethology* **29**, 221–226.

Schino, G. (2004). Birth sex ratio and social rank: consistency and variability between primate groups. *Behavioral Ecology* **15**, 850–856.

*Schwarzova, L., Simek, J., Coppack, T. & Tryjanowski, P. (2008). Male-biased sex of extra pair young in the socially monogamous Red-backed Shrike *Lanius collurio*. *Acta Ornithologica* **43**, 235–240.

Sheldon, B. C. (1998). Recent studies of avian sex ratios. *Heredity* **80**, 397–402.

Sheldon, B. C. (2000). Differential allocation: tests, mechanisms and implications. *Trends in Ecology & Evolution* **15**, 397–402.

*†Sheldon, B. C., Andersson, S., Griffith, S. C., Örnborg, J. & Sendecka, J. (1999). Ultraviolet colour variation influences blue tit sex ratios. *Nature* **402**, 874–877.

*Sheldon, B. C. & Ellegren, H. (1996). Offspring sex and paternity in the collared flycatcher. *Proceedings of the Royal Society of London Series B* **263**, 1017–1021.

Sheldon, B. C. & West, S. A. (2004). Maternal dominance, maternal condition and offspring sex ratio in ungulate mammals. *American Naturalist* **163**, 40–54.

Shuker, D. M. & West, S. A. (2004). Information constraints and the precision of adaptation: sex ratio manipulation in wasps. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10363–10367.

Shuster, S. M. & Wade, M. J. (2003). *Mating Systems and Strategies*. Princeton University Press, Princeton.

Silk, J. B. & Brown, G. R. (2008). Local resource competition and local resource enhancement shape primate birth sex ratios. *Proceedings of the Royal Society of London Series B* **275**, 1761–1765.

Silk, J. B., Willoughby, E. & Brown, J. R. (2005). Maternal rank and local resource competition do not predict birth sex ratios in wild baboons. *Proceedings of the Royal Society of London Series B* **272**, 859–864.

Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* **22**, 1359–1366.

Simonsohn, U., Nelson, L. D. & Simmons, J. P. (2014). P-curve: a key to the file drawer. *Journal of Experimental Psychology* **143**, 534–547.

*Soma, M. & Okanoya, K. (2013). Differential allocation in relation to mate song quality in the Bengalese finch. *Behaviour* **150**, 1491–1508.

*Stauss, M., Segelbacher, G., Tomiuk, J. & Bachmann, L. (2005). Sex ratio of *Parus major* and *P. caeruleus* broods depends on parental condition and habitat quality. *Oikos* **109**, 367–373.

Sterne, J. A. C. & Egger, M. (2005). Regression method to detect publication and other bias in meta-analysis. In *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments* (eds H. Rothstein, A. J. Sutton and M. Borenstein), pp. 99–125. John Wiley & Sons, Chichester.

Stubblefield, J. W. & Orzack, S. H. (2013). Resource transfers and evolution: helpful offspring and sex allocation. *Theoretical Population Biology* **83**, 64–81.

*Svensson, E. & Nilsson, J.-A. (1996). Mate quality affects offspring sex ratio in blue tits. *Proceedings of the Royal Society of London Series B* **263**, 357–361.

*Székely, T., Cuthill, I. C., Yezerinac, S., Griffiths, R. & Kis, J. (2004). Brood sex ratio in the Kentish plover. *Behavioral Ecology* **15**, 58–62.

*Taff, C. C., Freeman-Gallant, C. R., Dunn, P. O. & Whittingham, L. A. (2011). Relationship between brood sex ratio and male ornaments depends on male age in a warbler. *Animal Behaviour* **81**, 619–625.

Taylor, P. D. (1981). Intra-sex and inter-sex sibling interactions as sex-ratio determinants. *Nature* **291**, 64–66.

‡Tölü, C., Savas, T., Pala, A. & Thomsen, H. (2007). Effects of goat social rank on kid gender. *Czech Journal of Animal Sciences* **52**, 77–82.

Trivers, R. L. & Willard, D. E. (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science* **179**, 90–92.

Trnka, A., Prokop, P., Kasova, M., Sobekova, K. & Kocian, L. (2012). Hatchling sex ratio and female mating status in the great reed warbler, *Acrocephalus arundinaceus* (Aves, Passeriformes): further evidence for offspring sex ratio manipulation. *Italian Journal of Zoology* **79**, 212–217.

‡Ungerfeld, R., González-Sierra, U. T. & Piaggio, J. (2008). Reproduction in a semi-captive herd of pampas deer *Ozotocerus bezoarticus*. *Wildlife Biology* **14**, 350–357.

‡Veen, T., Träff, J., Weissing, F. J. & Sheldon, B. C. (2009). Reduced costs of mixed-species pairings in flycatchers: by-product or female strategy? *Behavioral Ecology and Sociobiology* **63**, 329–337.

*Veiga, J. P., Polo, V. & Cordero, P. J. (2008). Age dependent sex allocation in the polygynous spotless starling. *Evolutionary Ecology* **22**, 167–176.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**, 1–48.

Walsh, B. & Blows, M. W. (2009). Abundant genetic variation plus strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annual Review of Ecology, Evolution, and Systematics* **40**, 41–59.

WARNER, D. A., LOVERN, M. B. & SHINE, R. (2008). Maternal influences on offspring phenotypes and sex ratios in a multi-clutching lizard with environmental sex determination. *Biological Journal of the Linnean Society* **95**, 256–266.

*WATT, P. J., SHOHET, A. J. & RENSHAW, K. (2001). Female choice for good genes and sex-biased broods in guppies. *Journal of Fish Biology* **59**, 843–850.

WEATHERHEAD, P. J. (1983). Secondary sex-ratio adjustment in red-winged blackbirds (*Agelaius phoeniceus*). *Behavioral Ecology and Sociobiology* **12**, 57–61.

WEATHERHEAD, P. J. & ROBERTSON, R. J. (1979). Offspring quality and the polygyny threshold – sexy son hypothesis. *American Naturalist* **113**, 201–208.

WERREN, J. H. (1983). Sex-ratio evolution under local mate competition in a parasitic wasp. *Evolution* **37**, 116–124.

WEST, S. A. (2009). *Sex Allocation*. Princeton University Press, Princeton.

WEST, S. A., HERRE, E. A. & SHELDON, B. C. (2000). The benefits of allocating sex. *Science* **290**, 288–290.

WEST, S. A., REECE, S. E. & SHELDON, B. C. (2002). Sex ratios. *Heredity* **88**, 117–124.

†WEST, S. A. & SHELDON, B. C. (2002). Constraints in the evolution of sex ratio adjustment. *Science* **295**, 1685–1688.

WEST, S. A., SHUKER, D. M. & SHELDON, B. C. (2005). Sex-ratio adjustment when relatives interact: a test of constraints on adaptation. *Evolution* **59**, 1211–1228.

*WESTERDAHL, H., BENSCH, S., HANSSON, B., HASSELQUIST, D. & VON SCHANTZ, T. (1997). Sex ratio variation among broods of great reed warblers *Acrocephalus arundinaceus*. *Molecular Ecology* **6**, 543–548.

WESTNEAT, D. F. & STEWART, I. R. K. (2003). Extra-pair paternity in birds: causes, correlates, and conflict. *Annual Review of Ecology, Evolution, and Systematics* **34**, 365–396.

*WESTNEAT, D. F., STEWART, I. R. K., WOESTE, E. H., GIPSON, J., ABDULKADIR, L. & POSTON, J. P. (2002). Patterns of sex ratio variation in house sparrows. *Condor* **104**, 598–609.

‡WHEELWRIGHT, N. T. & SEABURY, R. E. (2003). Fifty:fifty offspring sex ratios in savannah sparrows (*Passerculus sandwichensis*). *Auk* **120**, 171–179.

*WHITTINGHAM, L. A. & DUNN, P. O. (2000). Offspring sex ratios in tree swallows: females in better condition produce more sons. *Molecular Ecology* **9**, 1123–1129.

*WIEBE, K. L. & BORTOLOTTI, G. R. (1992). Facultative sex-ratio manipulation in American kestrels. *Behavioral Ecology and Sociobiology* **30**, 379–386.

*WILK, T., CICHON, M. & WOLFF, K. (2008). Lack of evidence for improved immune response of extra-pair nestlings in collared flycatcher *Ficedula albicollis*. *Journal of Avian Biology* **39**, 546–552.

WOOD, W., KRESSEL, L., JOSHI, P. D. & LOUIE, B. (2014). Meta-analysis of menstrual cycle effects on women's mate preferences. *Emotion Review* **6**, 229–249.

WYMAN, M. J. & ROWE, L. (2014). Male bias in distributions of additive genetic, residual, and phenotypic variances of shared traits. *The American Naturalist* **184**, 326–337.

*YAMAGUCHI, N., KAWANO, K. K., EGUCHI, K. & YAHARA, T. (2004). Facultative sex ratio adjustment in response to male tarsus length in the varied tit *Parus varius*. *Ibis* **146**, 108–113.

*ZANN, R. & RUNCIMAN, D. (2003). Primary sex ratios in zebra finches: no evidence for adaptive manipulation in wild and semi-domesticated populations. *Behavioral Ecology and Sociobiology* **54**, 294–302.

*ZIELINSKA, M., DUBIEC, A. & ZIELINSKI, P. (2010). Offspring sex ratio skew in the sexually monomorphic house martin *Delichon urbicum*. *Journal of Avian Biology* **41**, 591–596.

VIII. SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

Fig. S1. The phylogenetic trees used for the phylogenetic meta-analysis and meta-regression analysis: (A) a tree with Ericson backbone and (B) a tree with Hackett backbone.

Fig. S2. Forest plot of the regression coefficients of the phylogenetic meta-regression with the Hackett tree.

Table S1. Results of phylogenetic (random-effects) meta-analyses using the Hackett tree.

Table S2. Results of multilevel and phylogenetic (random-effects) meta-analyses, incorporating correlations among effect sizes that come from the same set of individuals.

Table S3. Summary of the 25 studies excluded on full-text screening.

Appendix S1. Literature search and selection procedure.

Appendix S2. EndNote library (SexRatioAdjustment.enlp): all unique records obtained from the *Web of Science* literature searches outlined in Appendix S1.

Appendix S3. Raw data file (SexRatioAdjustment-Data.xlsx): all effect sizes included in the meta-analyses.