**No meta-analytic support for greater male variability in animal personalities**

**Abstract**

Human and animal studies alike often report ‘greater male variability’; males have greater representation at the extremes of most traits by virtue of their sex chromosomes. Men have greater variability for personality traits that attract women, yet it is unclear whether animal personalities show similar patterns. We conducted a meta-analysis of animal personality behaviours (*n*=2,248 effect sizes) across five taxonomic groups (*n*=231 species) to look for sex differences in personalities and the influence of sexual selection (degree of sexual size dimorphism, SSD) on trait means and variability. We found some male- and female-biased sex differences in personality means, but differences in trait variance always favoured greater female, not male, variability. Finally, SSD had significant interactions only with mean personality for mammals and invertebrates, which could reflect sex differences in life-histories rather than sexual selection. We suggest future studies take caution when discussing greater male variability for shared traits.

**Introduction**

*These first two paragraphs just won’t work no matter what I try to write – help!*

That males are more variable than females is a trend observed across the animal kingdom, consistent with the ‘greater male variability’ hypothesis. Male-biased variability than females across a range of behavioural, physiological and morphological traits (Pomiankowski & Møller 1995). In humans, male-biased sex differences in variability have been reported for traits like intelligence (Arden & Plomin 2006; Johnson *et al.* 2008), birth weights and several blood parameters (Lehre *et al.* 2009), and creativity (Ju *et al.* 2015; Karwowski *et al.* 2016). Furthermore, there are often more men at the extreme ends of aggressive personalities, leading to their overrepresentation in the prison system (Deary *et al.* 2003). Yet the mechanisms behind such widespread greater male variability remain unclear.

In the realm of human personality, our behavioural quirks can be categorised into one of five factors: Extraversion, Neuroticism, Openness, Conscientiousness and Agreeableness (Costa & McCrae 1992). Like most biological traits, these ‘Big Five’ factors are continuous and there are costs and benefits associated with personalities that fall at their extremes. For example, … Like human personalities, consistent animal behaviours, or personalities, can be grouped into trait types that resemble the human ‘big five’: Activity, Aggression, Boldness, Exploration and Sociality/Sociability (Reale *et al*. 2007). For organisms living in the wild, consistent and predictable (i.e. repeatable) behaviour across time and contexts seems maladaptive, especially when there is such large inter-individual variation in response to the same environmental conditions (Dingemanse *et al.* 2010). However, the consistency of personality within individuals, and its persistence across generations, suggests that low intra-individual variation can be adaptive (Smith & Blumstein 2008; Dingemanse & Wolf 2010). On the surface, human and animal personality dimensions seem to share behavioural traits, so it’s possible that the mechanisms behind such widespread, sex-biased variation in human personalities might also drive sex differences in animal personalities.

Firstly, the ‘sex chromosome’ hypothesis (James 1973) suggests that greater male variability is the result of sex chromosomal arrangement. For mammals, females have two copies of the X chromosome (homogamety), so the effects of any genes found on the X chromosome become averaged out. The averaging of gene expression can occur either through epigenetic inactivation of some regions during development (Amos-Landgraf *et al.* 2006), or through mosaicism where the ratio of genes on either X chromosome are variably expressed (Lyon 1961). Male mammals, however, only have one copy of the X chromosome (heterogamety). As such, all genes present on the male copy of the X chromosome are fully expressed, leading to more extreme phenotypes and, consequently, higher variances for shared traits (i.e. dosage compensation; see review by Charlesworth 1996). However, not all males are heterogametic – male birds have ZZ sex chromosomal arrangement while female birds have ZW, making females the heterogametic sex (reviewed in Graves & Shetty 2001). If sex chromosomal arrangement is the mechanism driving sex-biased variability, species with heterogametic females should have greater trait variability than males. Indeed, in a meta-analysis covering a broad range of taxonomic groups, Reinhold & Engqvist (2013) found that for species with heterogametic males, males had greater variability in body size than did females, and in species with heterogametic females, females had greater variability in body size than males. Sex chromosomal arrangement and greater male variability can also extend to behaviours, as they too have a genetic underpinning (Dochtermann *et al.* 2014). For example, men have greater representation at the tail ends of intelligence distributions, even when there are no mean differences in intelligence between men and women (Feingold 1992; Halpern & LaMay 2000; Arden & Plomin 2006; Johnson *et al.* 2008).

Second, under sexual selection, the sex with the greatest variation in their mating success (often males) should also have greater variation in traits used to obtain matings (e.g. weapons, ornaments) (Darwin 1871). In order for sexual selection to operate, traits important for male reproduction should have increased variance, compared to the same traits in females or other non-sexually selected traits (Pomiankowski & Møller 1995). For example, a meta-analysis comparing phenotypic trait variances for males and females found that there was a significant male-bias for coefficients of phenotypic variance for reproductive and non-reproductive traits (Wyman & Rowe 2014). However, traits important for reproduction showed more male-biased phenotypic variance than traits unrelated to reproduction (Wyman & Rowe 2014). While it is reasonable to expect male sexual traits to have greater variability than non-sexually selected traits, traits shared by both sexes are also frequently expected to show male-biased variability. For example, male-male competition plays an important role in male reproduction in many animal systems, thus aggression is a personality trait where males might have greater variability than females. In human antisocial personality types (low Agreeableness and high Neuroticism; similar to the Aggression dimension in animals) males show greater variability than females (Budaev 1999, Archer & Mehdikhani 2003). Greater male variability in human aggression is thought to be maintained by negative frequency-dependent selection whereby aggressive tendencies have both costs (risk of physical injuries) and benefits (more mating opportunities via social dominance) for male reproduction (Buss 1996, Budaev 1999).

Non-random mate choice might also contribute to sex differences in personalities. Females might prefer males that consistently behave in a certain way if such behaviour signals that they can provide direct or indirect benefits to her (Jennions & Petrie 1997; Schuett *et al.* 2010; Munson *et al.* 2020). For example, boldness might function as an honest signal of a potential partner’s quality; consistently bolder males might be better able to provision offspring with food compared to shy males, thus providing direct benefits to the female for her choice (Munson *et al*. 2020). Where the choosy sex (usually females) shows a strong preference for partners with a particular trait, that trait will have increased variance because it is under sexual selection (Pomiankowski & Møller 1995). Finally, different sex roles can also drive sex differences in behaviour. Both mating system and parental care can change the strength of sexual selection on one sex (usually males), producing larger variances in male reproductive success (Janicke *et al.* 2016). For example, species with monogamous males that share parental care might have no sex differences in their personalities, but where males mate multiply and do not care for offspring there might be more male-biased personality traits. Regardless of the underlying mechanism, most hypotheses surrounding sex differences in all traits tend to focus on greater male variability, as males tend to benefit more, or incur greater costs, for having greater representation at the tail-ends of trait distributions.

To investigate the prevalence of male-biased sex differences in variability for shared behavioural traits, we chose to compare animal personality traits across five broad taxonomic groups. Personalities, as part of their definition, should be consistent over time (i.e. low intra-individual variation), making personality a useful trait for comparing sex differences in behaviours. Additionally, tests used to measure personality are often similar enough to enable comparison over many studies, allowing us to conduct a systematic review and meta-analysis of sex differences in personalities. Although sex differences in animal personalities have been the focus of a recent meta-analysis (Tarka *et al*. 2018), it’s unclear whether the mechanisms behind greater male variability in human personalities are similarly responsible for sex differences in animal personalities. Therefore, we chose to include SSD as an additional moderator in our meta-analysis because it is a good proxy for the strength of sexual selection (i.e. larger male size reflects a condition-dependent trait, polygynous mating system, and/or different sex roles; Reiss 1986). Additionally, including SSD as a moderator can help disentangle the expected roles of sexual selection and sex chromosomal arrangement on sex-biased variability. First, if males do have greater variability than females in their personality traits, because of sexual selection, then SSD should be a significant moderator of such differences. Second, if sex chromosomal arrangement drives sex differences in behaviour, we should see greater male variability in taxonomic groups where males are the homogametic sex (i.e. mammals) and greater female variability in taxonomic groups where females are the homogametic sex (i.e. birds).

We had three main questions around which we formed our three meta-analytic models: 1) do males and females differ in either their mean trait expression, or variability in personality, 2) are there sex differences in mean or variability for the different personality types, and 3) does the degree of sexual size dimorphism explain sex differences in personality means or variability?

**Results**

*Summary of Effect Size Dataset*

Our final dataset contained n=2,248 effect sizes across five taxonomic groups: Mammals, Birds, Reptilia (amphibians and reptiles combined), Fish and Invertebrates. Birds were the most diverse group in our meta-analysis (n=107 unique species), while Mammals had effect sizes from more studies (n=61 studies) than any other group (Table 1). Reptilia had the least representation with n=10 species and n=95 effect sizes. Boldness was the most well-studied personality type and Sociality the least-studied personality type, with n=825 and n=169 effect sizes, respectively (Table 2).

*No Sex Differences in Personality*

Across all traits measured, males and females did not differ in personality for all taxonomic groups (birds: = -0.13, 95% CIs: -0.36, 0.10; fish: = -0.04, 95% CIs: -0.35, 0.28; invertebrates: = 0.24, 95% CIs: -0.03, 0.51; mammals: = 0.08, 95% CIs: -0.28, 0.45; reptiles: =0.07, 95% CIs: -0.52, 0.67). Variability in behavioural traits was similar for males and females across all traits and taxonomic groups (birds: = -0.15, 95% CIs: -0.66, 0.36; fish: = -0.04, 95% CIs: -0.09, 0.01; invertebrates: = -0.02, 95% CIs: -0.09, 0.06; mammals: = 0.07, 95% CIs: -0.20, 0.35; reptiles: =0.05, 95% CIs: -0.04, 0.14) (see Table 1 and Supplementary Table S1).

*‘Big Five’ Personality Axes*

When personality traits were broken down into the Big Five (Activity, Aggression, Boldness, Exploration and Sociality/Sociability) we found significant sex differences in both the means and the variability for some personality traits within some taxonomic groups.

Female birds were more sociable than males (=-0.68, 95% CIs: -1.16, -0.21, *t* = -2.83, *p*=0.005), male invertebrates were more aggressive than females (=0.45, 95% CIs: 0.00, 0.90, *t* = 1.98, *p*=0.05), and were bolder than females (=0.32, 95% CIs: 0.01, 0.62, *t* = 2.04, *p*=0.04), and male reptiles were more explorative than females (= 0.25, 95% CIs: 0.05, 0.45, *t* = 2.44, *p*=0.02). In contrast, there were no significant sex differences in mean personality for any of the Big Five traits for mammals or fish (see Table 2, Figures 1 & 2).

Female birds had greater variability in their exploratory behaviour than did males (=-0.26, 95% CIs: -0.52, 0.00, *t* = -1.97, *p*=0.05), and female fish had greater variability in aggressive behaviour compared to males (=-0.13, 95% CIs: -0.25, -0.01, *t* = -2.10, *p*=0.04). There were no sex differences in variability for any other personality trait for birds and fish, and we found no significant sex differences in the variability of invertebrates, mammals or reptiles for any of the five personality types (see Table 2, Figures 1 & 2 and Supplementary Table S2).

*Degree of Sexual Size Dimorphism Impacts Personality Differences between Sexes*

Sexual size dimorphism (SSD) alone did not significantly moderate sex differences in personality for any taxonomic group (Supplementary Tables S5 & S6), but SSD did interact significantly with some personality traits (Table 3). The degree of sexual size dimorphism between males and females did not significantly influence boldness behaviour in birds (intercept: *β* =-0.25, 95% CIs: -0.85, 0.34, *t* = -0.84, *p*=0.40; SSD: *β* =-0.24, 95% CIs: -1.45, 0.97, *t* = -0.39, *p*=0.70), nor did it significantly moderate aggressive or boldness behaviour in fish (aggression: *β* =-0.09, 95% CIs: -0.66, 0.47, *t* = -0.32, *p*=0.75; boldness: *β* =-0.13, 95% CIs: -0.70, 0.44, *t* = -0.44, *p*=0.66; SSD x aggression: *β* =0.16, 95% CIs: -0.85, 1.17, *t* = 0.31, *p*=0.76; SSD x boldness: *β* =-0.38, 95% CIs: -1.45, 0.69, *t* = -0.70, *p*=0.48)*.* For invertebrates, SSD significantly interacted with boldness; as males became larger than females (SSD became more male-biased) males were much bolder than females (boldness: *β* =0.37, 95% CIs: -0.02, 0.76, *t* = 1.89, *p*=0.06; SSD x boldness: *β* =1.29, 95% CIs: 0.22, 2.36, *t* = 2.37, *p*=0.02). SSD did not interact significantly with activity within invertebrates; as SSD became more male-biased there was a non-significant trend for females to be more active than males (activity: *β* =0.03, 95% CIs: -0.37, 0.42, *t* = 0.13, *p*=0.90; SSD x activity: *β* =-0.87, 95% CIs: -2.15, 0.42, *t* = -1.32, *p*=0.19). Finally, SSD interacted strongly and significantly with all personality traits (except sociality for which there was not enough data) for mammals. As SSD became more male-biased, females were more active than males (SSD x activity: *β* =-2.02, 95% CIs: -2.94, -1.10, *t* = -4.31, *p*<0.0001), but males were more aggressive (SSD x aggression: *β* =3.45, 95% CIs: 2.05, 4.85, *t* = 4.85, *p*<0.0001), were bolder (SSD x boldness: *β* =1.96, 95% CIs: 0.96, 2.96, *t* = 3.86, *p*=0.0001), and more explorative than females (SSD x exploration: *β* =1.92, 95% CIs: 0.91, 2.92, *t* = 3.75, *p=*0.0002). There were non-significant interactions with SSD for all personality trait types for all taxonomic groups for trait variability (see Table 3 and Supplementary Table S4).

*Mating system, age, population, study environment and study type can influence personality*

Monogamous and multiple mating systems were not significantly different from each other for means or variability for any of the taxonomic groups (Supplementary Table S7). Mean personality effect sizes for adults were marginally significantly different to juveniles for invertebrates (intercept: =0.24, 95% CIs: -0.03, 0.51, *t* = 1.74, *p*=0.08; juvenile: =-0.03, 95% CIs: -0.34, 0.28, *t* = -0.18, *p*=0.86), but not any other taxonomic group, and not for variability (Supplementary Table S8). Fish from the wild had greater differences in variability than fish from lab populations (intercept: =-0.09, 95% CIs: -0.18, -0.01, *t* = -2.11, *p*=0.04; lab: =0.08, 95% CIs: -0.02, 0.19, *t* = 1.55, *p*=0.12), but not for mean personality differences (Supplementary Table S9). Studies conducted in the lab were significantly different to field studies for mammals (intercept: =0.24, 95% CIs: -0.09, 0.57, *t* = 1.44, *p*=0.15; lab: = -0.31, 95% CIs: -0.56, -0.05, *t* = -2.34, *p*=0.02), but not for variability, and not for any other taxonomic group (Supplementary Table S10). Finally, effect sizes from observational studies were significantly different from experimental studies for mammals (intercept: =0.00, 95% CIs: -0.21, 0.22, *t* = 0.03, *p*=0.98; observation: = 0.39, 95% CIs: 0.13, 0.64, *t* = 2.39, *p*=0.004), but not for variability, and not for any other taxonomic group (Supplementary Table S11).

*Publication bias*

We found evidence of publication bias for means reported in invertebrate studies, but not for variability, and not for any other taxonomic group (Supplementary Table S12).

**Discussion**

Surprisingly, our meta-analysis failed to find significant sex differences in either trait means or trait variability for personality-like behaviours. When personalities were divided into their respective trait type (the Big Five) we found several significant sex differences in means for some personality types within some taxonomic groups (birds: Sociality; reptilia: Exploration; invertebrates: Aggression, Boldness). However, only exploratory behaviour for birds and aggressive behaviour for fish had significant sex differences in variability, both skewing towards females. Finally, personality and SSD interacted significantly to show strong sex-bias for mean personality, yet only within mammals (male-bias for all personality types except activity, which had a female-bias) and invertebrates (male-biased aggression), and did not moderate sex differences in variability for any personality type or taxonomic group.

The sex chromosome hypothesis suggests that variability for shared traits should reflect sex chromosomal arrangement. While we did observe a tendency for mean variability estimates to skew towards the heterogametic sex (especially mammals and birds), most estimates were nonsignificant. This was surprising considering the number of studies that suggest heterogamety is the mechanism driving greater male variability. For example, greater male than female variability has been reported for traits like personality types in humans (Archer & Mehdikhani 2003; Karwowski *et al.* 2016), cognition and intelligence (Halpern & LaMay 2000; Jones *et al.* 2003; Arden & Plomin 2006; Johnson *et al.* 2008, 2009), and for morphological traits like body size (Reinhold & Engqvist 2013), and brain structure (Arnold 2004; van der Linden *et al.* 2017; DeCasien *et al.* 2020). Importantly, these traits are also considered important for reproduction; in men, extraversion and creative personality types attract more mating opportunities (Buss 1995; Nettle 2006), while body size is important for male-male competition (Darwin 1871). Therefore, we should expect traits related to reproduction to have greater male variability, not just because of sex chromosomal arrangement.

Despite suggestions that most shared traits should have greater male variability, our meta-analysis found no evidence for male-biased variability for any personality trait. We expected aggression to show strong male bias within all taxonomic groups for mean differences and variability, yet only invertebrates had significant sex differences in mean aggression, and only fish had significant sex differences in variability for aggression (female-biased). These results were surprising given that male aggression should correlate with male-male competition and sexual selection (e.g. Schuett *et al*. 2010). However, female aggression is common for establishing social hierarchies and for defending offspring (e.g. Campbell 1999). Although sex roles can vary across the animal kingdom, females generally have the greatest investment in offspring (via anisogamy, Janicke *et al.* 2016), so females who provide parental care should be more aggressive than males in this context. Parental care could also potentially explain why only invertebrate males were significantly more aggressive than females; in many species, females don’t often provide care beyond egg-laying. However, we lacked suitable data to test this idea.

An additional explanation for the absence of sex differences comes from the human personality literature. Creativity and aggressive personality types have greater male variability and are important for female mate choice. Yet other personality traits frequently show no sex differences in either trait means or their variability. Archer and Mehdikhani (2003) compared means and variances for males and females for a range of personality-like traits and found that traits related to sexual selection (directly like physical aggression, or indirectly through a change in reproductive strategy) consistently had significant greater male variability, while traits unrelated to sexual selection had no sex differences in variability (Archer & Mehdikhani 2003). Our chosen sexual selection moderators (SSD and mating system) did not significantly change either mean trait expression or variability; while SSD did interact with personality trait type for mean trait expression in mammals, SSD had no significant effect on variability for any taxonomic group. If sexual selection operates to maintain inter-individual variation in animal personalities (Schuett *et al.* 2010), we would then expect greater variability for those traits in the sex under sexual selection. Since we did not observe significant sex differences in variability, animal personalities could instead reflect behaviours that don’t differentially affect reproductive outcomes.

Personality traits are frequently measured under the context of survival; personality encompasses antipredator responses (‘risky’ behaviour, or Boldness), foraging (Activity/Exploration), and other behaviours related to survival that are important for both sexes. Indeed, we found no sex differences in activity and only one instance of male-biased boldness behaviour for invertebrates. Where we did find sex differences in personality, the traits and their direction tended to reflect life-history differences between the sexes. For example, male reptiles were more explorative than females which might indicate male-biased dispersal (e.g. salamanders *Plethodon cinereus*; Liebgold *et al.* 2011), while female birds were more sociable than males that reflects different social group behaviour (e.g. zebra finches *Taeniopygia guttata*; Schuett & Dall 2009). Additionally, because there are costs (e.g. energetic investment) and benefits (e.g. faster growth, more feeding opportunities) associated with the expression of personality types (Wolf *et al.* 2007), both sexes likely experience trade-offs that lead to similar trait means and variabilities. Previous meta-analyses have found evidence of trade-offs between personality and survival (Smith & Blumstein 2008; Moiron *et al.* 2020), and we found some evidence that the degree of SSD might impact trade-offs between personality and survival for the sexes differently. For example, strong interactions between SSD and personality trait types for mammals, and between boldness and SSD for invertebrates, show that species with larger males trade-off ‘shy’ behaviour for larger body size. Larger animals are less likely to suffer mortality from predation, but also require more food to fuel their bigger bodies, thus need to feed more often and cannot play it safe. In mammals, as males became larger than females, males became bolder, more explorative and more aggressive than females. Male mammals are quite often the larger sex, due to sexual selection, so as males become larger they likely face more intense male-male competition (Darwin 1871)). However, SSD encompasses several different forms of selection, not just sexual selection (Blanckenhorn 2005). Because SSD did not interact significantly with trait variability, the sex differences in mean personality expression could instead be explained by adaptive, life-history differences.

Finally, phylogenetic variance (*I*2phylo) explained a large proportion of variance for personality variability (*lnCVR*) in both birds and mammals. This would suggest that variability in personality traits are heritable in endotherms and not ectotherms (Supplementary Table S1). Personality behaviours are heritable (Dochtermann *et al.* 2014), so assortative mating or mate choice for similar personality types could potentially allow variation in personality traits to persist, while keeping means and their variances relatively similar for both males and females (Schuett *et al.* 2010; Munson *et al.* 2020).

**Materials and Methods**

*Literature search and data collection*

We conducted a systematic review and meta-analysis by systematically searching the literature for empirical studies quantifying animal personality behavioural traits in males and females (as classified by Sih *et al.* 2004; Réale *et al.* 2007). We collected 9,698 records from both ISI Web of Science and Scopus databases on 11 December 2018, using primary and secondary keyword searches (see PRISMA in Figure 3, Table 4). A secondary search was used to find and include studies that measured ‘personality-like behaviours’; behaviours not explicitly defined as personalities but were measures of activity, aggression, boldness, exploration or sociality (Table 4). We also conducted a forwards citation search in both databases to include every record that had cited one of five key animal personality reviews: (Dall *et al.* 2004; Sih *et al.* 2004; Réale *et al.* 2007; Schuett *et al.* 2010; Dingemanse & Wolf 2010).

After removing duplicates, 3,739 records were left for title and abstract screening by LH. A total of 942 papers were excluded at the title/abstract screening stage because they were: 1) duplicates (*n* = 130), 2) non-animal studies (*n* = 256), 3) reviews or non-empirical studies (*n* = 516), 4) were not published in English (*n* = 8), or 5) were inaccessible (*n* = 32). This left 2,797 papers for full-text screening. We included papers in the final dataset only if they provided raw means and some measurement of error for both males and females because raw summary statistics are required for the derivation of variance-based effect size measures and provide greater opportunities to control for sources of non-independence (Nakagawa *et al.* 2015; Noble *et al.* 2017). Studies that only report PCAs and factor loadings were excluded as their interpretation can be challenging (e.g. Tarka *et al.* 2018). Briefly, we excluded papers at the full-text stage if they met the following exclusion criteria: 1) were not personality studies (i.e. studied a behaviour that could not fit into any of the ‘Big Five’ categories, like social node position), 2) studies used domestic or agricultural animals where individuals were neutered, 3) studies did not compare males and females (i.e. studies where animals were hermaphroditic, clonal or gynandromorphs, or where both sexes were analysed together, or only one sex was studied), or 4) data were unsuitable/missing (as shown in Figure 3). By definition personality traits must be repeatable (Sih *et al*. 2004), yet many studies did not report repeatability, nor refer to previous studies that found those traits to be repeatable. We therefore decided to include a broad range of behavioural types that match up with personality studies.

In total, this provided 211 eligible studies with suitable data that could be included in our meta-analysis (see Figure 3). We extracted means, error (standard deviation, standard error) and sample sizes from text, tables, figures and supplementary data files for both males and females on all personality traits described by the study authors. We used the R package *metaDigitise* (v1.0.0, Pick *et al.* 2019) to extract summary statistics from figures where needed.

*Effect size and sampling error*

To understand how males and females differ in their mean personality traits and their variability we used two contrast-based effect size measures. To understand changes in mean personality between the sexes we calculated Hedges’ *g* (sample size adjusted standardised mean difference; Hedges & Olkin 1985) and its associated sampling error as follows:

, *Eq. 1*

*Eq. 2*

Where is the mean, is the pooled standard deviation and N and SD are the sample size and standard deviation for males (M) and females (F), respectively. We used Hedges *g* instead of log response ratios as our chosen effect size as much of the data was not on a ratio scale and as such a response ratio could not be calculated. Effect size direction for a given behavioural trait depended on the way in which it was measured within a study. For example, boldness can often be measured as latency to flee, where a shorter time indicates a shyer individual. Yet the time to resume behaviour following a simulated predator approach, another common boldness measure, means that bolder individuals are those that resume normal behaviour more quickly. As such, we carefully examined all behavioural traits and ensured that the directionality (i.e., larger mean) was always consistent, flipping the direction of effect size estimates to ensure effect size interpretations were consistent.

We had strong mean-variance relationships in our data. As such, to understand differences in variance in personality traits across the sexes we used the log coefficient of variation (lnCVR) (Nakagawa *et al.* 2015), calculating the effect size and samping variance as follows:

,

*Eq. 3*

*Eq. 4*

where CV*M* and CV*F* are *SDM /M* and *SDF /F* , respectively. The lnCVR allows us to quantify the difference in variance between the sexes independent of changes in the mean (Nakagawa *et al.* 2015).

*Data transformations*

Our dataset contained some means, and their variances, that needed to be corrected before we could calculate SMD and lnCVR effect sizes. First, latency data (e.g. time to resume behaviour) that had not been corrected for normality (right-skewed) was log-transformed using the following calculations for means and standard deviations, respectively:

,

*Eq. 5*

*Eq. 6*

Second, proportion data constrain distributions of variances at high and low values, so we converted these data to a logit scale to meet assumptions of normality before calculating effect sizes.

*Moderator variables*

We collected a series of moderator variables that we *a priori* expected to moderate the effect size magnitude and direction. Personality traits extracted for males and females were categorised into one of five types suggested by Réale *et al.* (2007); ‘boldness’, ‘aggression’, ‘activity’, ‘sociality’ and ‘exploration’. We also recorded the taxonomic group (‘invertebrates’, ‘fish’, ‘amphibians’, ‘reptiles’, ‘birds’, ‘mammals’), the age of the sample (‘juveniles’ or ‘adults’), whether the study population was from the lab or the wild, and whether the study was conducted in the lab or field given this is expected to affect behaviour (Tarka *et al.* 2018). We also generated three moderator variables that attempt to capture the strength of sexual selection. First, we quantified the degree of sexual size dimorphism (SSD) between males and females; SSD is often condition-dependent and thus reflects male quality (Rowe and Houle 1996), and is strongly correlated with the strength of sexual selection (e.g. birds: Székely *et al.* 2004; Dale *et al.* 2007; fish: Walker & McCormick 2009; mammals: Promislow 1992). We created an SSD index by taking the ratio of male and female mean body size, mass or another sexually dimorphic trait (e.g. wing length), and used the following calculation (Lovich & Gibbons 1992):

*Eq. 7*

Using a ratio for SSD instead of raw means for male and females allowed us to: 1) have a single continuous moderator for sexual selection that tells us when the sexes are equal in size (0), when males are larger than females (positive values), and when females are larger than males (negative values), and 2) to compare SSD across a range of body size measures and scales. Our second measure attempting to capture the strength of sexual selection was the type of mating system (‘multiple mating’ and ‘monogamous’). Lastly, we also categorised the parental care strategies exhibited by each species within and across studies (‘biparental’, ‘maternal’, ‘paternal’, ‘none’). However, we decided to drop parental care as a moderator term because published data was sparse and unreliable for most species in our dataset. Where body size measures for males and females, mating system and parental care were not reported in the research article itself, we obtained these data by searching Web of Science, Scopus and Google Scholar. We used the search terms: “species name” AND male AND female AND body size OR length (for body size measures), “species name” AND parental care OR mating system for parental care and mating system.

*Meta-analyses*

We modelled SMD (standardised mean difference) and lnCVR (log-coefficient of variance) using multi-level meta-analytic (MLMA) and multi-level meta-regression (MLMR) models in R using the package *metafor* (version 2.4.0 Viechtbauer 2010; R Core Team 2019). In all our models we subset the data into broad taxonomic groups (i.e., birds, mammals, fish, invertebrates and reptilia (amphibians and reptiles combined due to low sample sizes)) as we were: 1) specifically interested in the extent to which patterns varied across these taxonomic groups and 2) because constructing phylogenies were easier within these sub-groups. We first estimated the overall evidence for a sex-specific effect in mean and variance across all personality traits for each specific taxonomic group using MLMA models. We accounted for effect size measurement error in our model and included study and phylogenetic random effects. We also estimated an observational-level (residual) variance as *metafor* does not estimate this by default. Phylogenetic correlation matrices were derived for each group either from existing phylogenetic tree databases (e.g. BirdTree.org for birds; Jetz *et al.* 2012, 2014) or using TimeTree.org (Hinchliff *et al.* 2015). For the bird phylogeny, we used the Ericson tree backbone (Ericson *et al.* 2006) to generate 1,000 trees and took a sample of 100 trees. We used *ape* and *phytools* in R to take an average tree from these 100 which we used for analyses. TimeTree phylogenies were derived by exporting a list of taxa from TimeTree.org. In some instances, a species that was present in our dataset had no phylogenetic representation so we found and substituted the closest sister species. In all cases, we resolved synonymous taxa across our dataset and trees and pruned our trees where needed. From our MLMA models we derived heterogeneity estimates (*I*2; Higgins & Thompson 2002; Nakagawa & Santos 2012). More specifically we derived between-study heterogeneity estimates (*I*2study), phylogenetic heterogeneity or phylogenetic heritability (*I*2phylo; Hadfield & Nakagawa 2010; Nakagawa & Santos 2012) along with total heterogeneity (*I*2Total), which is simply the proportion of total variance in effect size estimates excluding the total sampling variance.

We also fitted a series of models that included key moderator variables we predicted would explain variation in effect size. In all models, we again controlled for measurement error, study and phylogeny. Our first MLMR model estimated overall-meta-analytic mean estimates for each personality trait type separately (i.e. ‘activity’, ‘aggression’, ‘boldness’, ‘sociability’, ‘exploration’) within each taxonomic group. We expected that any mean and variance differences between the sexes would depend on the type of personality trait measured, especially where trait types are correlated with different sex roles (e.g. parental care; Schuett *et al.* 2010), life-histories (e.g. sex-biased dispersal and exploration or activity trait types; Cote *et al.* 2010), or are under sexual selection (e.g. male aggression might reflect the level of male-male competition or female mate choice; Munson *et al.* 2020). We also tested whether the degree of sexual selection, as measured by 1) the degree of sexual size-dimorphism (SSD) or 2) mating system, moderated effect sizes. We predicted that species with strong SSD or that have multiple mating systems would show stronger differences in mean and variance. However, we also expected such effects to depend on the personality trait being measured. As such, we fitted a model that included personality trait type, SSD and their interaction and a second model that included personality and mating system (we removed the personality trait-level interaction from mating system models due to lack of data). Levels with fewer than 10 species with SSD data were dropped from meta-regression models, and there were not enough data on reptiles to include them in these interaction models (Table 3). We present meta-analytic mean estimates and 95% confidence intervals from our models throughout.

Studies included in our meta-analysis varied greatly in their design and there was a number of additional sources of non-independence within studies (Noble *et al.* 2017). First, multiple personality traits were quantified on the same sets of individuals (e.g. ‘boldness’ and ‘aggression’). Additionally, some studies measured the same individuals multiple times for the same trait (i.e. repeatability), or for the same trait type using a different test (e.g. several measures of boldness). We conducted a series of sensitivity analyses to ascertain the impact that these sources of non-independence had on our final results. We created correlation matrices among effect sizes that shared the same sets of individuals in the sample used to derive effect sizes. Given that we did not know the exact correlation among traits we created three different dependency matrices (**D**; i.e., correlation matrices) that assumed r = 0.3, r = 0.5 and r = 0.8. We refit our models replacing the identify matrix (**I**) that was assumed when estimating the residual error variance with our **D** matrices.

*Publication bias*

To look for evidence of publication bias in our dataset, we first checked for funnel plot asymmetry across all effect sizes for both SMD and lnCVR. Instead of using an Egger’s regression, we decided to use effect size precision. When studies with small sample sizes (e.g. low precision) are published significantly more often that studies with larger sample sizes, those low-precision studies can drive funnel asymmetry. Therefore, we calculated and included precision as a moderator term in our personality trait MLMR models to look for significance of the slope (Van Aert *et al.* 2019). Where precision was significant, it is likely that publication bias is present. Precision was calculated using the following calculation:

*Eq. 8*

Where is the effect size Hedge’s *g* and its corresponding sampling variance.

**References**

Van Aert, R.C.M., Wicherts, J.M. & Van Assen, M.A.L.M. (2019). Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. *PLoS One*.

Amos-Landgraf, J.M., Cottle, A., Plenge, R.M., Friez, M., Schwartz, C.E., Longshore, J., *et al.* (2006). X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. *Am. J. Hum. Genet.*, 79, 493–499.

Archer, J. & Mehdikhani, M. (2003). Variability among Males in Sexually Selected Attributes. *Rev. Gen. Psychol.*, 7, 219–236.

Arden, R. & Plomin, R. (2006). Sex differences in variance of intelligence across childhood. *Pers. Individ. Dif.*, 41, 39–48.

Arnold, A.P. (2004). Sex chromosomes and brain gender. *Nat. Rev. Neurosci.*, 5, 701–708.

Blanckenhorn, W.U. (2005). Behavioral causes and consequences of sexual size dimorphism. *Ethology*, 111, 977–1016.

Borkenau, P., McCrae, R.R. & Terracciano, A. (2013). Do men vary more than women in personality? A study in 51 cultures. *J. Res. Pers.*, 47, 135–144.

Buss, D.M. (1995). Psychological Sex Differences: Origins Through Sexual Selection. *Am. Psychol.*, 50, 164–168.

Campbell, A. (1999). Staying alive: Evolution, culture, and women’s intrasexual aggression. *Behav. Brain Sci.*, 22, 203–252.

Charlesworth, B. (1996). The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol.*, 6, 149–162.

Costa, P.T. & McCrae, R.R. (1992). Four ways five factors are basic. *Pers. Individ. Dif.*, 13, 653–665.

Cote, J., Clobert, J., Brodin, T., Fogarty, S. & Sih, A. (2010). Personality-dependent dispersal: Characterization, ontogeny and consequences for spatially structured populations. *Philos. Trans. R. Soc. B Biol. Sci.*, 365, 4065–4076.

Dale, J., Dunn, P.O., Figuerola, J., Lislevand, T., Székely, T. & Whittingham, L.A. (2007). Sexual selection explains Rensch’s rule of allometry for sexual size dimorphism. *Proc. R. Soc. B Biol. Sci.*, 274, 2971–2979.

Dall, S.R.X., Houston, A.I. & McNamara, J.M. (2004). The behavioural ecology of personality: Consistent individual differences from an adaptive perspective. *Ecol. Lett.*, 7, 734–739.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. Murray, London.

Deary, I.J., Thorpe, G., Wilson, V., Starr, J.M. & Whalley, L.J. (2003). Population sex differences in IQ at age 11: The Scottish mental survey 1932. *Intelligence*, 31, 533–542.

DeCasien, A.R., Sherwood, C.C., Schapiro, S.J. & Higham, J.P. (2020). Greater variability in chimpanzee (Pan troglodytes) brain structure among males. *Proc. R. Soc. B Biol. Sci.*, 287, 20192858.

Dingemanse, N.J. & Wolf, M. (2010). Recent models for adaptive personality differences: A review. *Philos. Trans. R. Soc. B Biol. Sci.*, 365, 3947–3958.

Dochtermann, N.A., Schwab, T. & Sih, A. (2014). The contribution of additive genetic variation to personality variation: Heritability of personality. *Proc. R. Soc. B Biol. Sci.*, 282.

Ericson, P.G.P., Anderson, C.L., Britton, T., Elzanowski, A., Johansson, U.S., Källersjö, M., *et al.* (2006). Diversification of Neoaves: Integration of molecular sequence data and fossils. *Biol. Lett.*, 2, 543–547.

Feingold, A. (1992). Sex Differences in Variability in Intellectual Abilities: A New Look at an Old Controversy. *Rev. Educ. Res.*, 62, 61–84.

Graves, J.A.M. & Shetty, S. (2001). Sex from W to Z: Evolution of vertebrate sex chromosomes and sex determining genes. *J. Exp. Zool.*, 290, 449–462.

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

Halpern, D.F. & LaMay, M.L. (2000). The Smarter Sex: A Critical Review of Sex Differences in Intelligence. *Educ. Psychol. Rev.*, 12, 229.

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

Higgins, J.P.T. & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Stat. Med.*, 21, 1539–1558.

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

James, J.W. (1973). 353. Note: Covariances Between Relatives due to Sex-Linked Genes. *Biometrics*, 29, 584.

Janicke, T., Häderer, I.K., Lajeunesse, M.J. & Anthes, N. (2016). Evolutionary Biology: Darwinian sex roles confirmed across the animal kingdom. *Sci. Adv.*, 2, 1–11.

Jennions, M.D. & Petrie, M. (1997). Variation in mate choice and mating preferences: A review of causes and consequences. *Biol. Rev.*, 72, 283–327.

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.

Jetz, W., Thomas, G.H., Joy, J.B., Redding, D.W., Hartmann, K. & Mooers, A.O. (2014). Global Distribution and Conservation of Evolutionary Distinctness in Birds. *Curr. Biol.*, 24, 919–930.

Johnson, W., Deary, I.J. & Carothers, A. (2008). Sex Differences in Variability in General Intelligence: A New Look at the Old Question. *Perspect. Psychol. Sci.*, 3, 518–531.

Johnson, W., Deary, I.J. & Carothers, A. (2009). A Role for the X Chromosome in Sex Differences in Variability in General Intelligence? *Perspect. Psychol. Sci.*, 4, 598–611.

Jones, C.M., Braithwaite, V.A. & Healy, S.D. (2003). The evolution of sex differences in spatial ability. *Behav. Neurosci.*, 117, 403–411.

Ju, C., Duan, Y. & You, X. (2015). Retesting the greater male variability hypothesis in mainland China: A cross-regional study. *Pers. Individ. Dif.*, 72, 85–89.

Karwowski, M., Jankowska, D.M., Gajda, A., Marczak, M., Groyecka, A. & Sorokowski, P. (2016). Greater Male Variability in Creativity Outside the WEIRD World. *Creat. Res. J.*, 28, 467–470.

Lehre, A.C., Lehre, K.P., Laake, P. & Danbolt, N.C. (2009). Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. *Dev. Psychobiol.*, 51, 198–206.

Liebgold, E.B., Brodie, E.D. & Cabe, P.R. (2011). Female philopatry and male-biased dispersal in a direct-developing salamander, Plethodon cinereus. *Mol. Ecol.*, 20, 249–257.

van der Linden, D., Dunkel, C.S. & Madison, G. (2017). Sex differences in brain size and general intelligence (g). *Intelligence*, 63, 78–88.

Lovich, J.E. & Gibbons, J.W. (1992). A review of techniques for quantifying sexual size dimorphism. *Growth Dev Aging*, 56, 269–281.

Lyon, M.F. (1961). Gene Action in the X-chromosom (Mus musculus L.). *Nature*, 190, 372–373.

Moiron, M., Laskowski, K.L. & Niemelä, P.T. (2020). Individual differences in behaviour explain variation in survival: a meta-analysis. *Ecol. Lett.*, 23, 399–408.

Munson, A.A., Jones, C., Schraft, H. & Sih, A. (2020). You’re Just My Type : Mate Choice and Behavioral Types. *Trends Ecol. Evol.*, 1–11.

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., *et al.* (2015). Meta-analysis of variation: Ecological and evolutionary applications and beyond. *Methods Ecol. Evol.*, 6, 143–152.

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

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *Am. Psychol.*, 61, 622–631.

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

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

Pomiankowski, A. & Moller, A.P. (1995). A resolution of the lek paradox. *Proc. R. Soc. B Biol. Sci.*, 260, 21–29.

Promislow, D.E.L. (1992). Costs of sexual selection in natural populations of mammals. *Proc R Soc L. B*, 247, 203–210.

R Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria.

Réale, D., Reader, S.M., Sol, D., McDougall, P.T. & Dingemanse, N.J. (2007). Integrating animal temperament within ecology and evolution. *Biol. Rev.*, 82, 291–318.

Reinhold, K. & Engqvist, L. (2013). The variability is in the sex chromosomes. *Evol.*, 67, 3662–3668.

Rowe and Houle. (1996). The lek paradox and the capture of genetic variance. *Proc. Biol. Sci.*, 263, 1415–1421.

Schuett, W. & Dall, S.R.X. (2009). Sex differences, social context and personality in zebra finches, Taeniopygia guttata. *Anim. Behav.*, 77, 1041–1050.

Schuett, W., Tregenza, T. & Dall, S.R.X. (2010). Sexual selection and animal personality. *Biol. Rev.*, 85, 217–246.

Sih, A., Bell, A.M. & Johnson, J.C. (2004). Behavioral syndromes: an ecological and evolutionary overview. *TREE*, 19, 372–378.

Smith, B.R. & Blumstein, D.T. (2008). Fitness consequences of personality: A meta-analysis. *Behav. Ecol.*, 19, 448–455.

Székely, T., Freckleton, R.P. & Reynolds, J.D. (2004). Sexual selection explains Rensch’s rule of size dimorphism in shorebirds. *Proc. Natl. Acad. Sci. U. S. A.*, 101, 12224–12227.

Tarka, M., Guenther, A., Niemelä, P.T., Nakagawa, S. & Noble, D.W.A. (2018). Sex differences in life history, behavior, and physiology along a slow-fast continuum: a meta-analysis. *Behav. Ecol. Sociobiol.*, 72, 132.

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

Walker, S.P.W. & McCormick, M.I. (2009). Sexual selection explains sex-specific growth plasticity and positive allometry for sexual size dimorphism in a reef fish. *Proc. R. Soc. B Biol. Sci.*, 276, 3335–3343.

Wolf, M., Van Doorn, G.S., Leimar, O. & Weissing, F.J. (2007). Life-history trade-offs favour the evolution of animal personalities. *Nature*, 447, 581–584.

Wyman, M.J. & Rowe, L. (2014). Male Bias in Distributions of Additive Genetic, Residual, and Phenotypic Variances of Shared Traits. *Am. Nat.*, 184, 326–337.

Table 1. Intercept models for each taxonomic group looking for differences in mean (SMD) and variability (lnCVR) between males and females. Positive mean effect sizes reflect a female-bias in mean and variability for personality traits while negative mean effect sizes show male-bias.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parameters** | **Estimated mean** | **95% CI** | **Prediction interval** | ***t score*** | ***p*-value** | ***n* effect sizes** |
| *SMD (Hedge’s g)* |  |  |  |  |  |  |
| Mammals | 0.08 | -0.28, 0.45 | -1.10, 1.27 | 0.46 | 0.65 | 684 |
| Birds | -0.13 | -0.36, 0.10 | -1.85, 1.59 | -1.12 | 0.26 | 493 |
| Reptilia | 0.07 | -0.08, 0.22 | -0.52, 0.67 | 0.94 | 0.35 | 95 |
| Fish | -0.04 | -0.35, 0.28 | -1.00, 0.93 | -0.24 | 0.28 | 493 |
| Invertebrates | 0.24 | -0.03, 0.51 | -1.76, 2.24 | 1.73 | 0.08 | 483 |
| *lnCVR* |  |  |  |  |  |  |
| Mammals | 0.07 | -0.20, 0.35 | -0.64, 0.79 | 0.54 | 0.59 | 684 |
| Birds | -0.15 | -0.66, 0.36 | -1.95, 1.65 | -0.59 | 0.56 | 493 |
| Reptilia | 0.05 | -0.04, 0.14 | -0.06, 0.15 | 1.13 | 0.26 | 95 |
| Fish | -0.04 | -0.09, 0.01 | -0.64, 0.56 | -1.44 | 0.15 | 493 |
| Invertebrates | -0.02 | -0.09, 0.06 | -0.72, 0.69 | -0.44 | 0.66 | 483 |

Table 2. MLMR models with personality trait type for each taxonomic group. Values highlighted in bold indicate significant differences between males and females. Positive mean effect sizes reflect a male-bias in mean and variability for personality traits while negative mean effect sizes show female-bias. These models are graphically represented in Figures 1 & 2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parameters** | **SMD**  **mean** | **SMD 95%CI** | **SMD**  ***p-value*** | **lnCVR**  **mean** | **lnCVR**  **95%CI** | **lnCVR**  ***p-value*** |
| ***Mammals*** | | | | | | |
| Activity | -0.16 | -0.56, 0.23 | 0.41 | 0.11 | -0.19, 0.41 | 0.48 |
| Aggression | 0.11 | -0.27, 0.48 | 0.58 | 0.11 | -0.20, 0.42 | 0.48 |
| Boldness | 0.15 | -0.20, 0.50 | 0.39 | 0.06 | -0.21, 0.34 | 0.65 |
| Exploration | 0.06 | -0.31, 0.42 | 0.77 | 0.05 | -0.24, 0.34 | 0.73 |
| Sociality | 0.09 | -0.29, 0.47 | 0.64 | 0.06 | -0.24, 0.37 | 0.68 |
| ***Birds*** | | | | | | |
| Activity | -0.13 | -0.42, 0.16 | 0.37 | 0.04 | -0.25, 0.33 | 0.79 |
| Aggression | -0.11 | -0.40, 0.18 | 0.47 | -0.06 | -0.39, 0.27 | 0.72 |
| Boldness | -0.19 | -0.44, 0.07 | 0.15 | -0.01 | -0.23, 0.21 | 0.93 |
| **Exploration** | 0.08 | -0.19, 0.35 | 0.56 | **-0.26** | **-0.52, 0.00** | **0.05** |
| **Sociality** | **-0.68** | **-1.16, 0.21** | **0.005** | 0.13 | -1.59, 1.86 | 0.61 |
| ***Reptilia*** | | | | | | |
| Activity | -0.05 | -0.45, 0.36 | 0.82 | -0.11 | -0.51, 0.29 | 0.60 |
| Aggression | -0.07 | -0.32, 0.19 | 0.60 | 0.33 | -0.05, 0.72 | 0.09 |
| Boldness | 0.08 | -0.15, 0.31 | 0.68 | 0.10 | -0.13, 0.33 | 0.41 |
| **Exploration** | **0.25** | **0.05, 0.45** | **0.02** | -0.10 | -0.35, 0.14 | 0.40 |
| Sociality | -0.05 | -0.60, 0.50 | 0.86 | -0.12 | -0.76, 0.52 | 0.70 |
| ***Fish*** | | | | | | |
| Activity | -0.16 | -0.65, 0.33 | 0.53 | -0.03 | -0.16, 0.09 | 0.63 |
| **Aggression** | -0.05 | -0.52, 0.42 | 0.53 | **-0.13** | **-0.25, -0.01** | **0.04** |
| Boldness | -0.16 | -0.63, 0.32 | 0.52 | -0.02 | -0.63, 0.59 | 0.59 |
| Exploration | -0.05 | -0.54, 0.44 | 0.84 | -0.03 | -0.16, 0.09 | 0.62 |
| Sociality | -0.40 | -0.91, 0.11 | 0.12 | 0.07 | -0.11, 0.24 | 0.47 |
| ***Invertebrates*** | | | | | | |
| Activity | 0.08 | -0.23, 0.39 | 0.62 | -0.06 | -0.17, 0.05 | 0.26 |
| **Aggression** | **0.45** | **0.00, 0.90** | **0.05** | 0.02 | -0.17, 0.21 | 0.83 |
| **Boldness** | **0.32** | **0.01, 0.62** | **0.04** | -0.02 | -0.12, 0.08 | 0.65 |
| Exploration | 0.06 | -0.31, 0.43 | 0.74 | 0.05 | -0.10, 0.20 | 0.51 |
| Sociality | 0.30 | -0.26, 0.86 | 0.29 | 0.23 | -0.08, 0.54 | 0.15 |

Table 3. MLMR models with personality trait type and SSD moderator terms, as well as their interactions, for each taxonomic group. Note that some levels were dropped from these meta-regression models due to low sample sizes (fewer than 10 different species with SSD data within each personality trait type). Values highlighted in bold indicate significant differences between males and females.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameters** | **SMD**  **mean** | **SMD 95%CI** | **SMD**  ***p-value*** | **lnCVR**  **mean** | **lnCVR**  **95%CI** | **lnCVR**  ***p-value*** | **n studies** | **n**  **species** | **n effect sizes** |
| ***Mammals*** | | | | | | |  |  |  |
| **Activity** | **0.27** | **-0.01, 0.54** | **0.05** | 0.07 | -0.22, 0.37 | 0.63 | 14 | 12 | 84 |
| Aggression | 0.15 | -0.06, 0.37 | 0.17 | 0.10 | -0.18, 0.39 | 0.48 | 15 | 13 | 85 |
| Boldness | 0.13 | -0.07, 0.33 | 0.20 | 0.10 | -0.18, 0.39 | 0.48 | 26 | 26 | 163 |
| Exploration | 0.03 | -0.15, 0.22 | 0.72 | 0.04 | -0.24, 0.32 | 0.78 | 19 | 16 | 223 |
| **SSD** | **-2.02** | **-2.94, -1.10** | **<0.0001** | 0.30 | -0.34, 0.94 | 0.36 |  |  |  |
| **Aggression x SSD** | **3.45** | **2.05, 4.85** | **<0.0001** | -0.15 | -1.14, 0.83 | 0.76 |  |  |  |
| **Boldness x SSD** | **1.96** | **0.96, 2.96** | **0.0001** | -0.30 | -1.02, 0.41 | 0.41 |  |  |  |
| **Exploration x SSD** | **1.92** | **0.91, 2.92** | **0.0002** | -0.07 | -0.80, 0.65 | 0.84 |  |  |  |
| ***Birds*** | | | | | | |  |  |  |
| Intercept (Boldness) | -0.25 | -0.85, 0.34 | 0.40 | 0.03 | -0.04, 0.11 | 0.38 | 21 | 78 | 235 |
| SSD | -0.24 | -1.45, 0.97 | 0.70 | 0.10 | -0.16, 0.37 | 0.44 |  |  |  |
| ***Fish*** |  |  |  |  |  |  |  |  |  |
| Aggression | -0.09 | -0.66, 0.47 | 0.75 | -0.11 | -0.31, 0.09 | 0.27 | 16 | 13 | 93 |
| Boldness | -0.13 | -0.70, 0.44 | 0.66 | -0.03 | -0.22, 0.16 | 0.74 | 23 | 12 | 172 |
| SSD | 0.16 | -0.85, 1.17 | 0.76 | -0.28 | -1.03, 0.48 | 0.47 |  |  |  |
| Boldness x SSD | -0.38 | -1.45, 0.69 | 0.48 | 0.39 | -0.44, 1.21 | 0.36 |  |  |  |
| ***Invertebrates*** |  |  |  |  |  |  |  |  |  |
| Activity | 0.03 | -0.37, 0.42 | 0.90 | -0.08 | -0.22, 0.06 | 0.28 | 22 | 19 | 204 |
| **Boldness** | **0.37** | **-0.02, 0.76** | **0.06** | -0.02 | -0.15, 0.11 | 0.79 | 27 | 26 | 178 |
| SSD | -0.87 | -2.15, 0.42 | 0.19 | 0.05 | -0.56, 0.66 | 0.87 |  |  |  |
| **Boldness x SSD** | **1.29** | **0.22, 2.36** | **0.02** | 0.07 | -0.62, 0.76 | 0.84 |  |  |  |

Table 4.Primary and secondary keyword searches used for our literature search on 11 December 2018. Keyword searches were altered slightly to account for differences in operator terms employed by each database and searches were further refined by relevant biology categories in ISI or SCOPUS.

|  |  |  |
| --- | --- | --- |
|  | ***ISI Web of Science*** | ***SCOPUS*** |
| *Primary keyword search terms* | (personalit\* OR “behavioural syndrome\*” OR “behavioral syndrome\*” OR temperament) AND (sex\*) NOT (man OR men OR woman OR women OR human) | personalit\* OR “behavioural syndrome” OR “behavioral syndrome” OR temperament AND sex AND NOT man AND NOT woman AND NOT human |
| *Secondary keyword search terms* | animal\* AND behav\* AND (bold\* OR shy\* OR neoph\* OR aggress\* OR explor\* OR hid\*) AND “sex differences” NOT man NOT woman NOT human | animal\* AND (bold\* OR shy\* OR neoph\* OR aggress\* OR explor\* OR hid\*) AND “sex differences” AND NOT man AND NOT woman AND NOT human |
| *Refined by categories* | zoology, ecology, biology, multidisciplinary sciences, evolutionary biology | agriculture & biology |



**Figure 1**. Phylogenetic relationships and orchard plots for a) mammals and b) birds. Node colours reflect SSD ratios for each species (no data is grey) and heatmap data shows *n* effect sizes for each personality trait type for each species. Orchard plots show all effect sizes and mean effect size for SMD and lnCVR MLMR models with personality trait as a moderator term. Thin bars are 95% CIs and thick bars are prediction intervals. Circle size reflects effect size precision where bigger circles have greater precision.



**Figure 2.** Phylogenetic relationships and orchard plots for c) reptilia, d) fish and e) invertebrates. Node colours reflect SSD ratios for each species (no data is grey) and heatmap data shows *n* effect sizes for each personality trait type for each species. Orchard plots show all effect sizes and mean effect size for SMD and lnCVR MLMR models with personality trait as a moderator term. Thin bars are 95% CIs and thick bars are prediction intervals. Circle size reflects effect size precision where bigger circles have greater precision.



**Figure 3.** *PRISMA* diagram showing the process of finding, screening and including/excluding studies for this meta-analysis. Searches were further refined to agriculture & biology (Scopus), zoology, ecology, biology, multidisciplinary sciences, evolutionary biology (WoS) categories.