doi:10.1111/evo.13554

How to quantify (the response to) sexual selection on traits

Jonathan M. Henshaw,[1,2,3] Michael D. Jennions,[1] and Loeske E. B. Kruuk[1]

[1] *Division of Ecology and Evolution, Research School of Biology, The Australian National University, Acton ACT 2601, Canberra, Australia*

[2] *Institute of Zoology, University of Graz, Universitätsplatz 2, Graz 8010, Austria*

[3] *E-mail: u4307221@anu.edu.au*

Received April 1, 2018
Accepted July 7, 2018

Natural selection operates via fitness components like mating success, fecundity, and longevity, which can be understood as intermediaries in the causal process linking traits to fitness. In particular, sexual selection occurs when traits influence mating or fertilization success, which, in turn, influences fitness. We show how to quantify both these steps in a single path analysis, leading to better estimates of the strength of sexual selection. Our model controls for confounding variables, such as body size or condition, when estimating the relationship between mating and reproductive success. Correspondingly, we define the Bateman gradient and the Jones index using partial rather than simple regressions, which better captures how they are commonly interpreted. The model can be applied both to purely phenotypic data and to quantitative genetic parameters estimated using information on relatedness. The phenotypic approach breaks down selection differentials into a sexually selected and a "remainder" component. The quantitative genetic approach decomposes the estimated evolutionary response to selection analogously. We apply our method to analyze sexual selection in male dusky pipefish, *Syngnathus floridae*, and in two simulated datasets. We highlight conceptual and statistical limitations of previous path-based approaches, which can lead to substantial misestimation of sexual selection.

KEY WORDS: **Animal model, opportunity for sexual selection, quantitative genetics, Robertson's secondary theorem of selection, selection gradient, structural equation modeling.**

We often care about the exact route by which traits affect fitness, and not just the strength of these effects. For instance, large males may be fitter, but is this because they live longer, mate more often, fare better in sperm competition, or are multiple pathways involved? To answer this type of question, we must consider selection as a multistep causal process (Kingsolver and Schemske 1991; Scheiner et al. 2000; Frank 2013; Morrissey 2014). Individual components of this process can be estimated and then integrated into an overall measure of "selection via" any causal intermediary of interest.

Here, we apply this philosophy to the quantification of sexual selection. Sexual selection arises via competition for mates or fertilization opportunities (Shuker 2010). It can be conceptualized as a two-step causal process (Fig. 1): phenotypic traits affect mating or fertilization success (Step 1); and mating or fertilization success affects reproductive success, defined as the total production of offspring over a given period (Step 2). Recent reviews have championed a unified view of this process (Jones 2009; Henshaw et al. 2016; Anthes et al. 2017). However, most empirical work still considers each step in isolation, usually via particular metrics such as the *opportunity for sexual selection*, an upper bound on Step 1, or the *Bateman gradient*, an estimate of Step 2 (see Table 1 for a glossary of terms). This has contributed to widespread controversy about how to define and quantify sexual selection (Klug et al. 2010; Krakauer et al. 2011; Fitze and Le Galliard 2011; Jennions et al. 2012; Mobley 2014).

Here, we formalize the causal structure of sexual selection using path analysis, an extension of regression modeling that accommodates complex causal patterns (Wright 1934; Loehlin 2004; Shipley 2016). Our approach quantifies the relationships between measured traits, mating success, and reproductive success in a single framework. It also allows us to correct for the

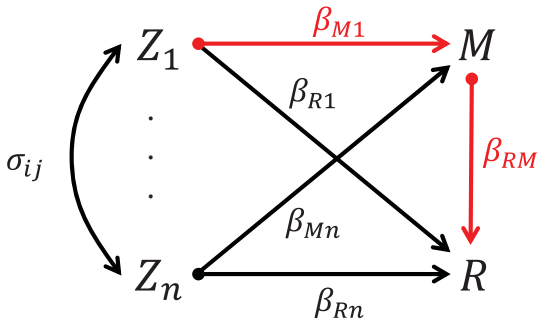

Figure 1. Path diagram for the causal relationships between traits (or environmental effects) $Z_1, \ldots, Z_n$, mating success M, and reproductive success R. Single-headed arrows indicate potential causal effects that are included in the model. Double-headed arrows indicate covariances between pairs of traits Z_i and Z_j, which are not analyzed causally. Path coefficients from A to B are written β_{BA} and covariances are written σ_{AB}. The pathway of direct sexual selection on Z_1 is shown in red, with strength estimated as $\sigma_1^2\beta_{M1}\beta_{RM}$, where σ_1^2 is the variance in Z_1.

potentially confounding effects of environmental variation. The strength of sexual selection acting on a trait can then be estimated by combining both steps in the causal pathway (Jones 2009). We focus on selection via mating success (often called "pre-mating" or "pre-copulatory" sexual selection). For brevity, we use "sexual selection" as a shorthand for "premating sexual selection" unless otherwise stated. However, our approach applies equally well to selection via fertilization success, or indeed–thinking beyond sexual selection–via any other fitness component, with only a change in terminology (e.g., replace "mating success" with "fertilization success").

SEXUAL SELECTION METRICS: WHAT'S WRONG WITH THE STATUS QUO?

Most popular indices of sexual selection capture only one step in the causal pathway linking a trait to reproductive success (Jones 2009; Henshaw et al. 2016; Anthes et al. 2017; summarized in Table 1). *Mating differentials* and *mating gradients* tell us how traits covary with mating success (Step 1 above) but not how, or even whether, this translates into differences in reproductive success or, ultimately, fitness (Lorch 2005). The same limitation holds for the *opportunity for sexual selection* I_s (i.e., the coefficient of variation in mating success), which provides an upper bound on the mating differential for a standardized trait (Wade and Arnold 1980; Jones 2009). In contrast, the *Bateman gradient* β_{ss} (Arnold and Duvall 1994) quantifies the relationship between mating and reproductive success (Step 2), but says nothing about the determinants of mating success. The Bateman gradient also fails to account for confounding variables (e.g., body size or condition) that may directly influence both mating and reproductive success (Gerlach et al. 2012; Collet et al. 2012, 2014; Janicke et al. 2016; Anthes et al. 2017).

These various metrics are not alternative ways to measure a single entity ("sexual selection"). Rather, they are estimates or upper bounds of distinct components of the sexual selection process. The overall relationship between traits and reproductive success can be quantified using *selection differentials* (defined here as the covariance between a trait and relative reproductive success) or *selection gradients* (partial regression coefficients of relative reproductive success on trait values: Lande and Arnold 1983; Henshaw and Zemel 2017). However, these metrics include selection from all sources, both sexual and nonsexual, and consequently do not illuminate the contribution of sexual selection to overall selection on traits. None of the above metrics is an appropriate stand-alone measure of the strength of sexual selection.

Jones (2009) recognized the need for metrics that integrate both steps in the sexual selection pathway, following on from previous path-based models of sexual selection (Arnold 1994; Arnold and Duvall 1994; Conner 1996). He defined the strength of premating sexual selection on a trait as the product $m\beta_{ss}$ of the mating differential m and the Bateman gradient β_{ss} (Table 1). The *Jones index* $s'_{\max} = \beta_{ss}\sqrt{I_s}$ provides an upper bound on $m\beta_{ss}$ for any variance-standardized trait. In a recent simulation study, the Jones index outperformed all other non-trait-based indices in predicting the strength of sexual selection on a trait (Henshaw et al. 2016). Nonetheless, like the Bateman gradient, Jones' metrics are vulnerable to confounding factors when estimating the relationship between mating and reproductive success.

Anthes et al. (2017) illustrated the pitfalls of current metrics with a simple example concerning sexual selection on females. Imagine a species where female fecundity increases with size, and where males have consequently evolved to prefer mating with larger females. Females need sperm from at least one male to fertilize their eggs, but matings are easily obtained and additional matings do not influence female reproductive success. In this scenario, there may be positive covariances between (i) female size and reproductive success; (ii) female size and mating success, due to males' mating preferences; and (iii) female mating and reproductive success, as both are influenced by female size. Mating and selection differentials (m and s) on female size will consequently be positive, as will the Bateman gradient β_{ss} and Jones index $s'_{\max}$. In fact, all currently used metrics will indicate positive sexual selection, even though there is no actual sexual selection on female size. Larger females have both more mates and more offspring, but the beneficial effect of size is not because higher mating success increases reproductive success. Similarly, we can imagine scenarios when favorable environmental conditions elevate both mating and reproductive success, but mating success does not directly influence reproductive success. These examples show how careful we need to be in defining and quantifying sexual selection to avoid drawing nonsensical conclusions.

Table 1. Glossary of terms. Note that we treat reproductive success as a proxy for fitness.

Phenotypic variance-covariance matrix ($\boldsymbol{P}$)	Matrix of phenotypic variances and covariances for a set of n traits ($\boldsymbol{P}_Z$) or n traits and mating success ($\boldsymbol{P}_{ZM}$)
Genetic variance-covariance matrix ($\boldsymbol{G}$)	Matrix of additive genetic variances and covariances for a set of n traits ($\boldsymbol{G}_Z$) or n traits and mating success ($\boldsymbol{G}_{ZM}$)
Selection differential (s)	Covariance between a trait (or set of traits) and reproductive success; also equals the difference in mean traits values between all individuals and the parents of offspring, with the latter weighted by the number of offspring per parent (Lande and Arnold 1983)
Selection gradient	Simple or partial regression coefficient of reproductive success on a trait (Lande and Arnold 1983)
Mating differential (m)	Covariance between a trait (or set of traits) and mating success (Jones 2009)
Mating gradient	Simple or partial regression coefficient of mating success on a trait (Jones 2009)
Path coefficient (β_{BA})	Estimate of the direct causal effect of one variable A on another variable B
Extended selection gradient (η)	Total effect of a trait on reproductive success via any causal pathway (Morrissey 2014)
Opportunity for sexual selection (I_s)	Variance in relative mating success or, equivalently, the coefficient of variation in absolute mating success (Wade and Arnold 1980; Jones 2009)
Simple Bateman gradient (β_{ss})	Simple regression coefficient of reproductive success on mating success (Arnold and Duvall 1994)
Partial Bateman gradient (β_{RM})	Partial regression coefficient of reproductive success on mating success, controlling for other measured traits or environments
Simple Jones index ($s'_{\max} = \beta_{ss}\sqrt{I_s}$)	Estimated maximum strength of premating sexual selection on a trait (Jones 2009)
Partial Jones index ($x'_{\max} = \beta_{RM}\sqrt{I_s}$)	Estimated maximum strength of premating sexual selection on a trait, controlling for other measured traits or environmental variables that might confound the relationship between mating and reproductive success (eq. (11))
Sexual selection (x)	Estimated component of the selection differential on a trait that is due to sexual selection (eq. (7))
Remaining selection (r)	Estimated component of the selection differential on a trait that is not due to sexual selection ($s = x + r$; eq. (8))
Predicted evolutionary response to selection (s_g)	Additive genetic covariance between a trait and reproductive success; also equals the estimated difference in mean breeding values between all individuals and the parents of offspring, with the latter weighted by the number of offspring per parent (Robertson 1966, 1968)
Predicted response to sexual selection (x_g)	Estimated component of the evolutionary response to selection that is due to sexual selection (eq. (14))
Predicted response to remaining selection (r_g)	Estimated component of the evolutionary response to selection that is not due to sexual selection (eq. (14))

OUR APPROACH

Our path-analytic model avoids these problems by estimating both steps in the sexual selection pathway simultaneously, while controlling for likely confounding variables. In particular, it controls for all measured traits (e.g., body size) and environments (e.g., local resource availability) when estimating the Bateman gradient (Collet et al. 2012, 2014). We are not the first to apply a path-based method to quantifying selection (e.g., Kingsolver and Schemske 1991; Scheiner et al. 2000, 2002; Morrissey 2014), nor to the particular case of sexual selection (Arnold 1994; Arnold and

Duvall 1994; Conner 1996). However, our approach avoids key conceptual and statistical problems associated with these previous approaches (see Comparisons to existing path-based models of selection).

Our method can be applied both to purely phenotypic data and to quantitative genetic parameters estimated using information on the relatedness between individuals (Lynch and Walsh 1998; Kruuk 2004; Wilson et al. 2010). As input it requires only the covariances (phenotypic or additive genetic) between all measured variables–at minimum, this means one needs to quantify mating success, reproductive success, and at least one trait. The phenotypic model decomposes the selection differential into components due to premating sexual selection (defined here as selection via mating success) and all remaining selection (which encompasses both nonsexual selection and any sexual selection that does not act via mating success, including postmating sexual selection and selection for mate quality or differential allocation: Jones 2009; Fitzpatrick 2015; Haaland et al. 2017). The quantitative genetic model decomposes the predicted evolutionary response to selection analogously.

Methods: Phenotypic Path Model

We model the relationships among traits $Z_1, \ldots, Z_n$, mating success M, and reproductive success R, according to the path diagram in Figure 1. We assume that traits can directly influence both mating and reproductive success, and that mating success can affect reproductive success. We aim to estimate the strength of these causal relationships. Environmental effects (e.g., local food availability) can be included as "traits" with no change to the model structure (see Statistical considerations). We initially present the model in its phenotypic form, but later show how it can be adapted to a quantitative genetic framework to decompose the predicted evolutionary response to selection.

COVARIANCES, PATH COEFFICIENTS, AND STANDARDIZATION OF TRAITS

Like all path analyses, our approach involves two types of statistic that are important to distinguish: covariances and path coefficients (Loehlin 2004; Kline 2016). Covariances are statistical associations between variables that can potentially arise from multiple alternative causal patterns (i.e., one variable influences the other and/or both are influenced by a third variable). Covariances between trait values and mating success are called *mating differentials*, and those between traits and reproductive success are *selection differentials* (Table 1). The selection differential is of particular interest because it equals the difference in mean trait values between all individuals and the parents of offspring, with the latter weighted by the number of offspring per parent (Lande 1979; Jones 2009; Henshaw and Zemel 2017).

Path coefficients are estimates of the strength of the causal influence of one variable on another, assuming that the chosen path model is correct (see Statistical considerations). *Mating gradients* and *selection gradients* can be thought of as path coefficients arising from the simple path model that underlies multiple regressions: a dependent variable (mating or reproductive success) is influenced by several independent variables (traits or environmental variables).

We write $m_i = \text{cov}(M, Z_i)$ and $s_i = \text{cov}(R, Z_i)$ for the mating and selection differentials on the ith trait, respectively. The selection differential on mating success is $s_M = \text{cov}(R, M)$. Column vectors of mating and selection differentials are written $\boldsymbol{m} = (m_1, \ldots, m_n)^T$ and $\boldsymbol{s} = (s_1, \ldots, s_n, s_M)^T$. Note that the vector of selection differentials includes s_M (i.e., mating success is treated as the $(n+1)$th trait).

The phenotypic variance-covariance matrix of the n traits is denoted $\boldsymbol{P}_Z$. It is useful to also consider an enlarged phenotypic variance-covariance matrix $\boldsymbol{P}_{ZM}$ that includes mating success as a trait. The relationship between $\boldsymbol{P}_Z$ and $\boldsymbol{P}_{ZM}$ is:

$$\boldsymbol{P}_{ZM} = \begin{bmatrix} \boldsymbol{P}_Z & \boldsymbol{m} \\ \boldsymbol{m}^T & \sigma^2_M \end{bmatrix} \quad (1)$$

We write β_{Mi} for the path coefficients from the ith trait to mating success, β_{Ri} for those from the ith trait to reproductive success, and β_{RM} for that from mating success to reproductive success (Fig. 1). Column vectors of path coefficients leading to mating and reproductive success are written respectively as $\boldsymbol{\beta}_M = (\beta_{M1}, \ldots, \beta_{Mn})^T$ and $\boldsymbol{\beta}_R = (\beta_{R1}, \ldots, \beta_{Rn}, \beta_{RM})^T$.

Our analysis does not depend on how traits are standardized, except where noted. For many applications, it is helpful to variance-standardize trait values (i.e., first deduct the mean from all trait values, and then divide by the standard deviation). Selection differentials are then measured in units of phenotypic standard deviations (Lande and Arnold 1983). Mean-standardization of traits (i.e., dividing by the mean) may be more appropriate for some purposes (Houle 1992; Hereford et al. 2004; Houle et al. 2011). *Relative* (i.e., mean-standardized) *reproductive success* should always be used to ensure compatibility with quantitative genetic theory, and we use *relative mating success* for similar reasons (Jones 2009).

SOLVING FOR THE PATH COEFFICIENTS

Following Wright's rules (Wright 1934; Loehlin 2004), the path coefficients in Figure 1 should obey the following equations (where we write $\sigma_{ij} = \text{cov}(Z_i, Z_j)$):

$$m_i = \sum_{j=1}^{n} \sigma_{ij} \beta_{Mj} \quad (2)$$

$$s_i = \sum_{j=1}^{n} \sigma_{ij} \left(\beta_{RM} \beta_{Mj} + \beta_{Rj} \right) \quad (3)$$

$$s_M = \sigma_M^2 \beta_{RM} + \sum_{i=1}^{n} \beta_{Ri} \sum_{j=1}^{n} \sigma_{ij} \beta_{Mj} \quad (4)$$

This is a system of $2n + 1$ equations with the same number of path coefficients as unknowns (remember that n is the number of traits). We should consequently expect an exact solution (in path-analytic terms, the system is "just-determined": Loehlin 2004). Equation (2) can be written in matrix form as:

$$\boldsymbol{m} = \boldsymbol{P}_Z \boldsymbol{\beta}_M \quad (5)$$

Similarly, by substituting (2) into (4) and using the definition of $\boldsymbol{P}_{ZM}$ from equation (1), we can summarize equations (3) and (4) as a single matrix equation:

$$\boldsymbol{s} = \boldsymbol{P}_{ZM} \boldsymbol{\beta}_R \quad (6)$$

By inverting these two matrix equations, we obtain simple expressions for all path coefficients, namely $\boldsymbol{\beta}_M = \boldsymbol{P}_Z^{-1} \boldsymbol{m}$ and $\boldsymbol{\beta}_R = \boldsymbol{P}_{ZM}^{-1} \boldsymbol{s}$. These results are intuitive. The path coefficients $\boldsymbol{\beta}_M$ are equal to the partial regression coefficients of mating success on the n traits under ordinary least squares regression (i.e., the mating gradients). Similarly, $\boldsymbol{\beta}_R$ equals the partial regression coefficients of reproductive success on the n traits and mating success (i.e., the selection gradients).

A SIMPLE MEASURE OF SEXUAL SELECTION

We can now partition total selection on the ith trait into sexual selection and remaining selection: $s_i = x_i + r_i$. Sexual selection (used here as a shorthand for "premating sexual selection") is the relationship between a trait and reproductive success via mating success. It consists of all pathways from the trait Z_i to reproductive success R that pass through mating success M (Fig. 1). It is given by:

$$x_i = \beta_{RM} \left(\sum_{j=1}^{n} \sigma_{ij} \beta_{Mj} \right) = \beta_{RM} m_i \quad (7)$$

Remaining selection consists of all pathways from Z_i to R that do not pass through M, and is given by:

$$r_i = \sum_{j=1}^{n} \sigma_{ij} \beta_{Rj} \quad (8)$$

More succinctly, the vector $\boldsymbol{s}_Z = (s_1, \ldots, s_n)^T$ of selection differentials on the n traits, excluding mating success, can be expressed as:

$$\boldsymbol{s}_Z = \boldsymbol{x} + \boldsymbol{r} = \boldsymbol{P}_Z \left(\beta_{RM} \boldsymbol{\beta}_M + \boldsymbol{\beta}_{RZ} \right), \quad (9)$$

where $\boldsymbol{\beta}_{RZ} = (\beta_{R1}, \ldots, \beta_{Rn})^T$.

Morrissey (2014) refers to the total effect of a trait on fitness as the *extended selection gradient* η (see Comparisons to existing path-based models of selection). For our model (Fig. 1), the vector of extended selection gradients on all n traits is given by $\boldsymbol{\eta} = \beta_{RM} \boldsymbol{\beta}_M + \boldsymbol{\beta}_{RZ}$. Equivalently, $\boldsymbol{\eta}$ is the vector of selection gradients obtained from the multiple regression of reproductive success on the n traits (not including mating success). "Direct" selection on the ith trait is given by $\sigma_i^2 \eta_i$, consistent with the Lande-Arnold framework (Lande and Arnold 1983; Scheiner et al. 2000).

Sexual selection $\boldsymbol{x}$ and remaining selection $\boldsymbol{r}$ take on different meanings depending on how fitness components are defined and measured (Evans and Garcia-Gonzalez 2016; Anthes et al. 2017). For instance, studies of premating sexual selection usually define an individual's mating success operationally as its number of matings or mates. If mating success is taken as the intermediate variable in our model, then "sexual selection" $\boldsymbol{x}$ is strictly confined to selection to increase this number. "Remaining selection" $\boldsymbol{r}$ covers everything else, including both nonsexual selection, and any type of sexual selection that does not act via mating success, such as sperm competition, cryptic choice, and, more subtly, selection for mate quality or differential allocation (Jones 2009; Fitzpatrick 2015; Haaland et al. 2017). If other variables such as paternity share or fertilization success are substituted for mating success, then the interpretation changes accordingly. It is important not to take the labels "sexual" and "remaining" too literally, as study-specific definitions of fitness components determine exactly which processes these pathways include.

"PARTIAL" DEFINITIONS OF THE BATEMAN GRADIENT AND THE JONES INDEX

The path coefficient β_{RM} estimates the causal effect of mating success on reproductive success. It is analogous to the Bateman gradient β_{ss}, which is the slope of the simple linear regression of R on M (Arnold and Duvall 1994; Jones 2009). Unlike the Bateman gradient, however, β_{RM} controls for spurious relationships between mating and reproductive success that arise from any common causes (i.e., traits or environments) that are included in the analyses (Collet et al. 2012, 2014). We consequently refer to β_{RM} as a *partial* Bateman gradient.

If all traits affecting mating success are uncorrelated with those affecting reproductive success (i.e., if $\sigma_{ij} = 0$ whenever both β_{Mi} and β_{Rj} are nonzero) then equation (4) gives us:

$$\beta_{RM} = \frac{s_M}{\sigma_M^2} = \beta_{ss} \quad (10)$$

In other words, if mating and reproductive success are influenced by two independent sets of traits, then the partial and simple Bateman gradients are equal. Total sexual selection on the ith trait is then $x_i = m_i \beta_{ss}$, which is the same

definition as used by Jones (2009) (see Introduction). For variance-standardized traits, this quantity is bounded in magnitude by the Jones index $s'_{\max} = \beta_{ss}\sqrt{I_s}$, where I_s is the opportunity for sexual selection (Wade and Arnold 1980; Jones 2009; Henshaw et al. 2016; see Table 1). The Jones index can also be thought of as a Bateman gradient calculated using variance-standardized, rather than mean-standardized, mating success (McDonald and Pizzari 2018). If mating and reproductive success are determined by independent sets of traits, then "remaining" selection can be expressed as $r_i = \text{cov}(Z_i, \varepsilon)$, where ε is the residual from the simple regression of reproductive success on mating success (see Jones (2009) for derivation).

When either a particular trait, or a pair of correlated traits, influences both mating and reproductive success directly, the simple and partial Bateman gradients may differ. In this case, the simple Bateman gradient reflects the overall association between mating and reproductive success, whereas the partial Bateman gradient controls for measured confounding factors that influence both mating and reproductive success. Sexual selection on the ith trait can then be bounded as follows:

$$|x_i| \leq |\beta_{RM}| \sqrt{\sigma_i^2 I_s} \tag{11}$$

For variance-standardized traits, the strength of premating sexual selection is thus bounded in magnitude by the *partial* Jones index $x'_{\max} = \beta_{RM}\sqrt{I_s}$. This estimated bound is informative even for traits that are not included in the analysis. It is unbiased only if the analysis includes all factors that confound the relationship between mating and reproductive success. The simple Jones index $s'_{\max}$ is generally a less accurate bound because it does not control for any confounding factors.

Worked Examples

To illustrate the use of our method we provide two worked examples, where we analyse sexual selection on body size in (1) a simulated dataset based on the example of Anthes et al. (2017) given in the Introduction, and (2) an empirical dataset on male dusky pipefish, *Syngnathus floridae*. Datasets and R code for both examples are provided at https://doi.org/10.5061/dryad.1fp7830.

FAT, SEXY, AND FECUND

First, we modeled the example of misidentified sexual selection on females from Anthes et al. (2017) (see Introduction). We also compared our model's estimates of sexual selection to the path-based models of Arnold (1994), Conner (1996), and Jones (2009) (for full details of these models see Comparisons to existing path-based models of selection). Two of these models (Arnold 1994; Conner 1996) incorporate a derived variable, "fecundity per mate," that is not included in our model (see Figs. 2 and 3).

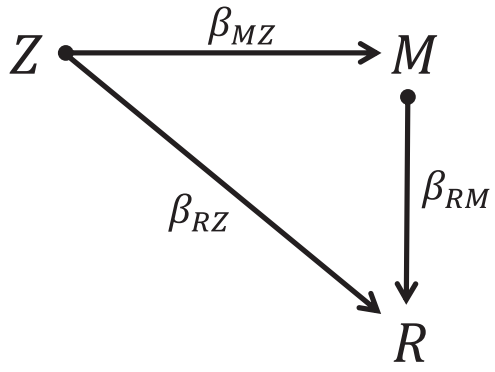

Figure 2. Path diagram for the two worked examples, "Fat, sexy, and fecund" and "Size selection in male pipefish," shown with body size or length *Z*, mating success *M*, and reproductive success *R*. Path coefficients from *A* to *B* are written β_{BA}.

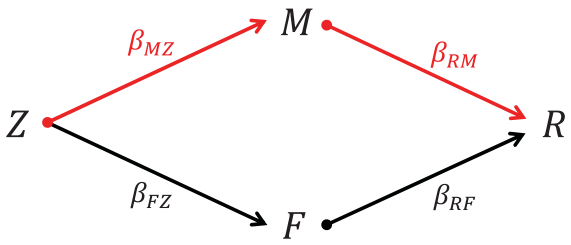

Figure 3. Path diagram for the causal relationships between a single trait *Z*, mating success *M*, fecundity per mate *F*, and reproductive success *R*, as used in the models of Arnold (1994) and Conner (1996). The coefficients β_{MZ} and β_{FZ} are simple regression coefficients of *M* on *Z* and *F* on *Z*, respectively. The coefficients β_{RM} and β_{RF} are either partial regression coefficients in the multiple regression of *R* on both *M* and *F* (Arnold 1994) or simple regression coefficients in separate regressions of *R* on *M* and *R* on *F*, respectively (Conner 1996). The pathway of direct sexual selection on *Z* is shown in red, with strength estimated as $\sigma_Z^2\beta_{MZ}\beta_{RM}$.

Fecundity per mate F is defined as an individual's reproductive success divided by its mating success. In this worked example, mating success is always at least one, so there is no issue of division by zero; in general, however, the definition of fecundity per mate is problematic (see Comparisons to existing path-based models of selection).

We simulated female body size Z, mating success M and reproductive success R in a sample of 1000 females (details in Supporting Information). We assumed that all females find at least one mate, and that mating success increases roughly linearly with body size. Reproductive success also increases linearly with size, but is not directly influenced by mating success.

Although there was strong selection on body size ($s_Z = 0.30$), our model correctly attributed this to the direct effect of body size on reproductive success ($r_Z = 0.29$), rather than to premating sexual selection ($x_Z = 0.00$: Table 2). In contrast, the alternative models all indicated substantial sexual selection (i.e., selection via mating success), even though none is acting. They

Table 2. Results of the worked example "Fat, sexy, and fecund," showing how various modeling approaches estimate sexual versus remaining selection on body size *Z*, based on its relationships with mating success *M*, fecundity per mate *F*, and reproductive success *R*.

Parameter	Model	Formula	Raw value	Relativized value
Mating gradient on Z	All	β_{MZ}	1.73	0.33
Fecundity-per-mate gradient on Z	A, C	β_{FZ}	−0.30	−0.12
Selection gradient on Z (controlling for M)	Ours	$\beta_{RZ} = \beta_{RZ\|M}$	3.00	0.29
Simple Bateman gradient	C, J	$\beta_{ss} = \beta_{RM\|\emptyset}$	0.80	0.40
Partial Bateman gradient	Ours	$\beta_{RM} = \beta_{RM\|Z}$	0.00	0.00
Selection gradient on M (controlling for F)	A	$\beta_{RM\|F}$	1.39	0.70
Selection gradient on F (no covariates)	C	$\beta_{RF\|\emptyset}$	0.36	0.09
Selection gradient on F (controlling for M)	A	$\beta_{RF\|M}$	1.30	0.33
Sexual selection	A	$\sigma_Z^2\beta_{RM\|F}\beta_{MZ}$	2.40	0.24
	C, J	$\sigma_Z^2\beta_{ss}\beta_{MZ}$	1.38	0.14
	Ours	$x_Z = \sigma_Z^2\beta_{RM}\beta_{MZ}$	0.00	0.00
Remaining selection	A	$\sigma_Z^2\beta_{RF\|M}\beta_{FZ}$	−0.39	−0.04
	C	$\sigma_Z^2\beta_{RF\|\emptyset}\beta_{FZ}$	−0.11	−0.01
	J	$\mathrm{cov}(Z, \varepsilon)$	1.62	0.16
	Ours	$r_Z = \sigma_Z^2\beta_{RZ}$	3.00	0.29
Total selection (selection differential)	All	$s_Z = \sigma_{ZR} = x_Z + r_Z$	3.00	0.30

We present parameter estimates from analyses of raw data ("Raw value") and also from analyses where *F*, *M*, and *R* were mean-standardized ("Relativized value"). In both cases, body size is standardized so that $\sigma_Z^2 = 1$. We write $\beta_{AB|CD}$ for the partial regression coefficient of *A* on *B* with additional covariates *C* and *D* (i.e., based on the regression $A \sim B + C + D$). The notation $\beta_{AB|\emptyset}$ explicitly indicates that there are no covariates (i.e., the simple regression coefficient of *A* on *B*). The residual in the simple regression of *R* on *M* is denoted ε (Jones 2009). Models are abbreviated as: A = Arnold (1994); C = Conner (1996); J = Jones (2009); and Ours = our model (see Comparisons to existing path-based models of selection). All nonzero parameters values are statistically significant ($P < 0.05$).

also drastically misestimated the strength of remaining selection on body size (Table 2).

The failure of the alternative models has two main causes. First, they overestimate the causal effect of mating success on reproductive success by failing to control for body size. A naive interpretation of the simple Bateman gradient would suggest that each additional mating partner results in 0.8 additional offspring on average (unstandardized value of the simple Bateman gradient $\beta_{ss} = 0.80$: Table 2). In contrast, our model correctly estimates that there is no causal effect of mating success on reproductive success (partial Bateman gradient $\beta_{RM} = 0.00$). Second, if reproductive success is independent of mating success M, then fecundity per mate F is proportional to $1/M$ for any given female. The models of Arnold (1994) and Conner (1996) include fecundity per mate as a variable, but fail to account for this dependence (see Comparisons to existing path-based models of selection). Note that these models would still misestimate the strength of sexual selection even if selection via both M and F were classified as "sexual selection" (Table 2).

SIZE SELECTION IN MALE PIPEFISH

For our second example, we analyzed sexual selection in dusky pipefish (*Syngnathus floridae*), a sex-role reversed species, using both our model and that of Jones (2009). Male dusky pipefish brood embryos in a specialised pouch, which may contain offspring from multiple females simultaneously. Mate choice is largely driven by females; even so, our focus here is on sexual selection in males.

Mobley and colleagues collected pregnant male dusky pipefish across five sites on the Eastern coast of North America (Mobley and Jones 2007, 2009, 2013; Mobley et al. 2014). They measured body length (tip of snout to base of caudal fin) and dissected brood pouches to extract developing embryos. Parentage analysis was performed on a subset of 42 embryos from each male (details in Mobley and Jones 2007). A male's reproductive success was defined as the total number of embryos in his brood pouch. Mating success was estimated as the number of unique mothers identified in the subsample of 42 embryos. The resulting datasets contain information on male body length, mating success, and reproductive success across six samples (corresponding to five sites, one of which was sampled in two different years). Here, we only consider data from pregnant males, which means that the estimated Bateman gradients (whether simple or partial) may not be reflective of the whole male population.

We compared our model's estimates of sexual and remaining selection on body length to those of Jones (2009), based

Table 3. Results of the worked example "Size selection in male pipefish," showing estimates of sexual and remaining selection on male body length Z using our model and that of Jones (2009).

Site		VA	NC	TB	SJ	TX1	TX3
Sample size		30	24	30	22	13	26
Mating gradient on male size β_{MZ}		0.20*	0.14	0.22*	0.19*	0.12	0.02
Selection gradient on male size β_{RZ}		0.25*	0.25*	0.12	0.13	0.35*	0.01
Simple Bateman gradient β_{ss}		0.47*	0.36	0.58*	0.33	1.02*	0.49*
Partial Bateman gradient β_{RM}		−0.11	0.13	0.36	0.18	0.70*	0.47*
Sexual selection	Jones (2009)	0.10	0.05	0.12	0.06	0.12	0.01
	Our model (x)	−0.02	0.02	0.08	0.03	0.09	0.01
Remaining selection	Jones (2009)	0.13	0.21*	0.07	0.10	0.31*	0.01
	Our model (r)	0.25*	0.25*	0.12	0.13	0.35*	0.01
Total selection (s)		0.23*	0.27*	0.20*	0.17	0.44*	0.02

Within each population, body length is variance-standardized, whereas mating and reproductive success are mean-standardized. Data originate from Mobjack Bay, Virginia (VA); Morehead City, North Carolina (NC); Tampa Bay, Florida (TB); St. Joseph Bay, Florida (SJ); and Aransas Pass, Texas (TX1 and TX3, representing two different years). For details of data collection see (Mobley and Jones 2007; 2009; 2013). Note that some values differ from those in the original publications because we only used data for pregnant males.

*Statistically significant ($P < 0.05$).

on the path diagram in Figure 2. We first calculated variance-standardized body length and relative (i.e., mean-standardized) mating and reproductive success separately within each sample (corresponding to an assumption of soft selection among sites: De Lisle and Svensson 2017). Reproductive success increased with male size in all six samples. Selection differentials on body length ranged from 0.02 to 0.44 and were significant in four of the six samples (Table 3).

Two non-mutually-exclusive mechanisms might explain the greater success of larger males. First, larger males might be more attractive to females, whose eggs are a limiting resource (sexual selection). Second, larger males may have a greater capacity to brood eggs (remaining selection). Our model suggests that both mechanisms might operate in male dusky pipefish, but that sexual selection is less important (Table 3). Our model produced consistently smaller estimates of sexual selection and larger estimates of remaining selection than Jones' model (Wilcoxon signed-rank test, $P = 0.036$). This is because Jones' model did not control for the confounding effect of body length when estimating the relationship between mating and reproductive success. Crucially, larger males had both more mates and more offspring, but the effect of mate number on reproductive success was smaller than a simple regression (i.e., the simple Bateman gradient) would suggest.

Methods: Quantitative Genetic Path Model

Until now we have only discussed sexual selection on phenotypes. Our phenotypic model describes how realised trait values are related to reproductive and mating success; for example which phenotypes have higher reproductive success, and to what extent can this be explained by differences in mating success? As input, the model requires only the matrix of covariances among measured traits, environmental effects, mating success, and reproductive success. Equations (5) though (9) then allow the selection differential on each trait to be split into components for sexual and remaining selection.

However, our method can be extended to estimate the various components of the sexual selection process at the genetic level. Given detailed knowledge of relatedness among individuals (e.g., a pedigree, controlled breeding design, or high-density marker data), one can estimate the additive genetic covariances among measured variables (Lynch and Walsh 1998; Kruuk 2004; Wilson et al. 2010). By Robertson's secondary theorem of selection, the predicted evolutionary response to selection (i.e., the estimated difference in mean breeding values between all individuals and parents of offspring) is equal to the additive genetic covariance between a trait and fitness (Robertson 1966, 1968; Morrissey et al. 2010; Walsh and Lynch 2014). A quantitative genetic framing of the path model (Fig. 1) allows the estimated response to selection to be broken down along the proposed causal pathways (similarly, Stinchcombe et al. (2014) and Kruuk et al. (2014) used multivariate analyses to estimate genetic selection gradients). This allows one to estimate the evolutionary responses to sexual and remaining selection.

For the quantitative genetic model, the analogs of equations (5) and (6) are:

$$\boldsymbol{m}_g = \boldsymbol{G}_Z \boldsymbol{\beta}_{Mg} \tag{12}$$

$$\boldsymbol{s}_g = \boldsymbol{G}_{ZM} \boldsymbol{\beta}_{Rg} \tag{13}$$

The *genetic mating differential* $\boldsymbol{m}_g$ is a vector of estimated additive genetic covariances between the n traits and mating success. The *genetic selection differential* $\boldsymbol{s}_g$ consists of the additive genetic covariances between the $n+1$ traits (including mating success) and reproductive success (Stinchcombe et al. 2014). It estimates the evolutionary response to selection for these traits. The matrices $\boldsymbol{G}_Z$ and $\boldsymbol{G}_{ZM}$ are additive genetic variance-covariance matrices, excluding and including mating success, respectively. The genetic path coefficients $\boldsymbol{\beta}_{Mg}$ and $\boldsymbol{\beta}_{Rg}$ are analogous to their phenotypic counterparts, except that they describe causal relationships among breeding values, rather than realized trait values.

The estimated response to selection can be broken down into sexual and remaining components using the analog of equation (9):

$$\boldsymbol{s}_{Zg} = \boldsymbol{x}_g \; + \; \boldsymbol{r}_g = \boldsymbol{G}_Z \left(\beta_{RMg}\boldsymbol{\beta}_{Mg} + \boldsymbol{\beta}_{RZg}\right) \qquad (14)$$

The quantitative genetic form of our model describes how breeding values for traits and reproductive success covary, and to what extent any association can be explained by genetic covariance with mating success. A worked example is provided in the Supporting Information.

Statistical Considerations

CAUSALITY, HIDDEN VARIABLES, AND CHOICE OF PATH MODELS

The path coefficients in our model are interpreted as estimates of the strength of causal effects. This interpretation is only fully valid if the path model correctly represents the underlying causality, meaning that (i) if *A* influences *B* then there is an arrow from *A* to *B* in the path diagram, and (ii) there are no "hidden variables" that independently influence two or more path variables, but are absent from the model (Loehlin 2004). For instance, it is assumed that mating success influences reproductive success, but not the other way around. This assumption will be violated, for example, in fish species where females prefer to mate with males tending larger nests, leading to a "backward" causal effect of male reproductive success on male mating success (Wisenden 1999). Mating and reproductive success can also directly affect phenotypic traits, such as when body condition deteriorates due to breeding (Milenkaya et al. 2013). Some of these backward causal effects can, however, be ameliorated by careful choice of variables (e.g., measure body condition at the start of the breeding season).

The assumption that there are no hidden variables is more problematic, as is well known from phenotypic studies of selection (Mitchell-Olds and Shaw 1987; Rausher 1992; Morrissey et al. 2010; Walker 2014; Reed et al. 2016). Analyses will be improved by including obvious confounders such as body size or condition (Scheiner et al. 2002). Importantly, our quantitative genetic approach is resistant to confounding by environmental variables, as these will usually show no additive genetic variance. Although some relevant variables will always be missed, our approach provides a better estimate of sexual selection than simple regression-based methods, which also assume no hidden variables. Ultimately, however, manipulative experiments are essential to resolve causal patterns (see Concluding remarks).

In some cases, traits are only expressed in a subset of individuals (e.g., breeding phenotypes are not expressed in individuals that die as juveniles). Moorad and Wade (2013) show how to estimate selection in the presence of such "nonexistent" traits, and their approach can be incorporated into our phenotypic model without issue (note that mating and reproductive success may be zero, but never "nonexistent"). For our quantitative genetic model, this is not necessary, because breeding values for a trait can be estimated even in individuals that do not express the trait (Hadfield 2008). This means that additive genetic variances and covariances can be estimated across all individuals, even those that do not express all traits.

For well-studied species, researchers may have strong prior hypotheses about the causal relationships among traits and fitness components. The best *a priori* path model may then differ from that in Figure 1 (e.g., by excluding some paths or by explicitly modeling the internal causal structure among traits). Tailored path models can lend considerable insight into the causes of selection (e.g., Conner et al. 1996; Sheldon and Ellegren 1999; Latta and McCain 2009; Dai and Galloway 2013).

The simple analytic equations that we present do not apply to these alternative models, which can be fitted using dedicated structural equation modeling (SEM) software (e.g., commercial packages like Mplus and Stata, or the R package *lavaan*: Rosseel 2012). However, the broad conceptual points above still apply, namely (1) that premating sexual selection should be quantified using the causal pathways that pass via mating success to reproductive success, and (2) that it is important to control for factors that potentially confound the relationship between mating and reproductive success.

MORE FLEXIBLE MODELS: GLMs, GLMMs, AND PIECEWISE SEM

Standard path analyses share many assumptions with linear regression. First, causal effects, represented by single-headed arrows in path diagrams (e.g., Fig. 1), are assumed to be linear and additive. In some cases, nonlinearity can be dealt with by transforming variables. It is preferable not to transform reproductive success, however, as this obscures connections between the path model and broader evolutionary theory (Stanton and Thiede 2005: see Comparisons to existing path-based models of selection).

Second, residuals are assumed to be normally distributed and homoscedastic (see Estimating uncertainty and model fit).

As mating and reproductive success are generally count variables (e.g., number of mates or offspring), these assumptions might be severely violated in practice. As a result, point estimation of parameters will be inefficient, requiring larger sample sizes to achieve the same error, and estimates of uncertainty may be biased (Finney and DiStefano 2006). Most dedicated SEM software includes "robust" estimation procedures that help to correct for nonnormality and heteroscedasticity (Finney and DiStefano 2006; Rosseel 2012). Another solution is to estimate each step of the sexual selection pathway using a more flexible model, such as a generalized linear-mixed model (GLMM), to accommodate departures from linearity and nonnormality. For instance, mating and reproductive success can be modeled using Poisson or negative binomial distributions (Broquet et al. 2015; Martin et al. 2015; Turnell and Shaw 2015; Worthington and Kelly 2016). Random effects (e.g., year or location) can also be incorporated using a GLMM (O'Hara 2009).

Traditional SEM has been slow to accommodate these more flexible regression approaches because all parameters are estimated globally (although Mplus and Stata have "generalized SEM" functions). In contrast, "piecewise" SEM estimates each piece of the model locally (Lefcheck 2016; Shipley 2016). The R package *piecewiseSEM* imports most mainstream regression techniques into an SEM context, using familiar R syntax (Lefcheck 2016).

One disadvantage of more general modeling approaches is that their parameters have no straightforward link to central constructs of evolutionary theory like the selection differential or the predicted response to selection. In theory, such a link could be forged by transforming statistics (e.g., selection differentials, path coefficients) from the link scale to the data scale via integration (de Villemereuil et al. 2016), although such techniques have not yet been developed for SEM. Study-specific decisions on how to construct these models may also make it harder to compare sexual selection among species or between the sexes, as has been done with standard metrics of sexual selection (Janicke et al. 2016; Janicke and Morrow 2018; Janicke et al. 2018). In contrast, our method can be applied to any system, and requires no more than the three variables needed to measure sexual selection on traits (i.e., trait values, mating or fertilization success, and reproductive success). Our approach is consequently well-suited for comparative analyses.

ESTIMATING UNCERTAINTY AND MODEL FIT

For the phenotypic model, point estimates of path coefficients can be obtained by inverting equations (5) and (6), or by fitting the associated regression models. Confidence intervals and *P*-values for path coefficients can, in principle, be taken directly from the fitted regressions. In practice, it is easier to use dedicated SEM software (e.g., the R package *lavaan*: Rosseel 2012) that can also provide uncertainty estimates for compound variables such as the total strength of sexual selection (see R code at https://doi.org/10.5061/dryad.1fp7830). For models using piecewise SEM, uncertainties in compound variables can be estimated using the Delta method (Lynch and Walsh 1998), but we do not explore this approach in detail here. The explanatory power of these models is best assessed by calculating the two R^2 values for the regressions of mating and reproductive success respectively on their explanatory variables (Nakagawa and Schielzeth 2013).

For the quantitative genetic model, point estimates of path coefficients can be obtained by inverting equations (12) and (13) (see worked example in Supporting Information). Estimating uncertainty is trickier, because it is important not to neglect the uncertainty in additive genetic covariances arising from the mixed model. This is currently best done by using a Bayesian framework to estimate posterior distributions for the genetic variance-covariance matrix (e.g., using the R package *MCMCglmm*: Hadfield 2010) and thence for all other derived parameters (see e.g., Stinchcombe et al. 2014; Kruuk et al. 2014). This approach can also be used to estimate parameter uncertainty in purely phenotypic models.

Sample size requirements for path analysis depend on many factors, including the complexity of the model, precision of measurement, and the reasonableness of assumptions such as normality and linearity. As a rule-of-thumb, Jackson (2003) and Kline (2016) recommend sample sizes of at least 10, or ideally 20, times the number of model parameters, but certainly not less than 100. For our phenotypic approach, there are $2n + 1$ model parameters, where n is the number of traits. Our model is very simple by SEM standards, so in practice somewhat smaller sample sizes may be adequate. The component equations (5) and (6) can also be studied separately using sample sizes sufficient for multiple regression. The quantitative genetic approach will require substantially larger sample sizes for precise estimates of the additive genetic covariances among variables, the exact value of which will depend on the relatedness structure, but would seem foolhardy to attempt without sample sizes in the several hundreds (Lynch 1999).

In the SEM literature, goodness-of-fit is sometimes assessed by how closely models reproduce the observed covariance matrix among measured variables. Our model necessarily reproduces this matrix perfectly (as it is "just determined": Loehlin 2004), and so most covariance-based measures of fit are uninformative. Researchers who wish to change the path structure in Figure 1 may find these metrics useful, however (see e.g., Kline (2016) for appropriate statistics). Model selection methods can also be used to compare the path model in Figure 1 to alternatives (e.g., models with fewer arrows), although we suspect that evolutionary biologists will rarely have the reasons, or the data, to do this satisfactorily.

Comparisons to Existing Path-Based Models of Selection

COMPARISON OF PHENOTYPIC MODELS

We are not the first to propose a path-based model of sexual selection, but our model avoids many previous conceptual and statistical limitations. Here, we briefly outline the models of Arnold (1994), Conner (1996), and Jones (2009) and highlight their differences from our model. The models of Arnold (1994) and Conner (1996) rely on expressing fitness as an exact product of fitness components. In the current context, this means that reproductive success R (in place of fitness) is expressed as mating success M times fecundity per mate F (Fig. 3). The models then contain two tiers of regression. First, each component of reproductive success (i.e., M or F) is regressed on the measured traits, yielding gradients of the form β_{MZ} or β_{FZ}, where Z is a trait. Second, reproductive success is regressed on its components, using either a combined multiple regression over all components (Arnold 1994) or separate simple regressions for each component (see Conner 1996 for the advantages of separate regressions). This yields gradients of the form β_{RM} and β_{RF}. Direct selection on Z via either M or F is then calculated as $\sigma^2_Z\beta_{MZ}\beta_{RM}$ or $\sigma^2_F\beta_{FZ}\beta_{RF}$, respectively. Jones' (2009) definition of sexual selection $m\beta_{ss}$ is equivalent to applying Conner's model to a single trait. However, Jones calculates remaining selection by deducting his estimate of sexual selection from the selection differential.

All three models differ from ours in one crucial respect: they calculate the relationship between mating and reproductive success either without controlling for other variables (Conner 1996; Jones 2009) or controlling only for fecundity per mate (Arnold 1994). In contrast, our model estimates this relationship using a partial Bateman gradient that controls for other measured traits, and can also accommodate potentially relevant environmental effects. In addition, the models of Arnold and Conner assume that mating success does not influence fecundity per mate. This assumption will be violated whenever total fecundity does not increase in proportion to mating success (i.e., for females of almost all species and males of most species). For example, if total fecundity is fixed, then fecundity per mate will be proportional to $1/M$, but this dependence is not accounted for by Arnold (1994) or Conner (1996).

Three additional technical issues affect previous path-based approaches, although these issues do not apply to all models. First, the derived variable "fecundity per mate" is not defined for individuals that never mate. Moorad and Wade (2013) resolve the problem of "nonexistent" values for regression-based models of selection; their approach could likely be extended to the path-analytic context. This issue does not arise under our approach, however, because fecundity per mate is not included as a variable.

Second, Arnold (1994) and Conner (1996) do not estimate gradients in a traditional path-analytic framework, but rather as stacked regressions. Deviations from Wright's path analysis rules (Wright 1934; Loehlin 2004) are consequently not minimised in any controlled way (unlike our model, those of Arnold and Conner are underdetermined: Loehlin 2004). Consequently, these models do not provide an additive (or even an "additive as possible") breakdown of the selection differential into sexual and remaining components (see Table 2, where sexual and remaining selection on body size do not add up to the selection differential). Jones' model does not suffer from this limitation, because remaining selection is defined explicitly to satisfy additivity.

Third, reproductive success is a multiplicative, rather than additive, function of its components (i.e., $R = MF$). This relationship is exact and known a priori, so estimating it via multiple regression, as Arnold does, merely creates an error-prone model without adding explanatory power (Conner 1996). Arnold suggests taking logarithms of reproductive success and its components to meet the additivity assumption of linear regression, but, even then, the regression remains uninformative.

These drawbacks of previous approaches can lead to substantial misestimation of sexual selection, even in simple, biologically plausible cases (see Worked examples and Table 2).

COMPARISON OF QUANTITATIVE GENETIC MODELS

Our quantitative genetic model extends the approach of Stinchcombe et al. (2014), who provided genetic analogs of selection gradients. It uses a similar methodology to Scheiner et al. (2002), who applied path analysis to the means of full and half-sib families. The advantage of our model is that it accommodates any method for estimating additive genetic covariances. Morrissey (2014) provided a conceptually similar approach with a distinct methodology. He uses a path model to estimate selection at the phenotypic level, and then infers the genetic consequences of selection via an analog of the breeder's equation. A great strength of Morrissey's approach is that it incorporates a very literal representation of the causal relationships among traits and their genetic underpinnings (i.e., genes only interact causally via their effects on phenotypes). However, it also suffers from a core problem with the classical breeder's equation: unmeasured environmental variables can confound the analysis, even if they lack genetic variance (Rausher 1992; Morrissey et al. 2010; Morrissey 2014).

Morrissey (2014) approximates the evolutionary response to selection as

$\Delta\bar{\boldsymbol{Z}} = \boldsymbol{\Phi}\boldsymbol{G}_{\epsilon}\boldsymbol{\eta}$. Here, $\boldsymbol{\Phi}$ is a matrix representing the total causal effects of each trait on all other traits (excluding fitness), estimated using a phenotypic path model. The extended selection gradients $\boldsymbol{\eta}$ are estimates of the total causal effects of each trait on fitness. The matrix $\boldsymbol{G}_{\epsilon}$ represents additive genetic variances and covariances beyond those attributable to causal relationships

in the path model. For instance, in our model, mating success is influenced by other phenotypic traits, and so some of the additive genetic variance in mating success will result from additive genetic variance in these traits. The matrix $\boldsymbol{G}_{\epsilon}$ contains *additional* additive genetic variance in mating success that does not result from the underlying traits.

For our model, we have:

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{I_n} & \mathbf{0} \\ \boldsymbol{\beta}_M^T & 1 \end{bmatrix}, \tag{15}$$

where $\boldsymbol{I_n}$ is an $n \times n$ identity matrix and $\mathbf{0}$ is a column of zeroes, and

$$\boldsymbol{\eta} = \boldsymbol{\Phi}^T \boldsymbol{\beta_R} = \begin{bmatrix} \beta_{RM}\beta_{M1} + \beta_{R1} \\ \vdots \\ \beta_{RM}\beta_{Mn} + \beta_{Rn} \\ \beta_{RM} \end{bmatrix} \tag{16}$$

Because $\boldsymbol{\eta}$ is based on phenotypic rather than additive genetic relationships, it is sensitive to confounding by environmental variables, including those with zero additive genetic variance. This confounding then affects the estimated evolutionary response to selection $\Delta\bar{\boldsymbol{Z}}$.

In contrast, our quantitative genetic approach is only sensitive to confounding factors that covary genetically with measured variables (e.g., other unmeasured traits, extended phenotypes, or parental effects: Stinchcombe et al. 2014). For instance, suppose that the mean temperature during development has direct positive effects on both snout-vent length (SVL) and reproductive success in a reptile. If temperature is excluded from the analysis, then Morrissey's approach will overestimate the causal phenotypic relationship between SVL and reproductive success, and this overestimate is carried through to the genetic analysis. Our model instead relies on the additive genetic covariance between SVL and reproductive success, which is not expected to be biased by temperature. We provide a worked example of our quantitative genetic approach in the Supporting Information.

Concluding Remarks

Path-analytic approaches have yielded fundamental insights into how selection operates (Kingsolver and Schemske 1991; Scheiner et al. 2000; Morrissey 2014), which are especially relevant to understanding and quantifying sexual selection (Arnold 1994; Arnold and Duvall 1994; Jones 2009; Anthes et al. 2017). We have built upon this tradition by providing a consistent framework to quantify sexual selection that avoids many of the pitfalls of previous approaches. Our approach estimates causal relationships based on observations of unmanipulated natural or laboratory populations. These causal inferences will necessarily be uncertain, because it is impossible to measure and control for all possible confounding factors. We agree with Anthes et al. (2017) that manipulative studies should play a larger role in sexual selection research. Nonetheless, many important variables are difficult to manipulate (e.g., attractiveness, fighting ability, mating success). We consequently believe that observational studies will continue to play a central role in sexual selection research, and it is important to analyze them in a conceptually sound statistical framework. We see two important goals for future theory: (1) integrate pre- and postmating sexual selection into a single framework that breaks down the relative contributions of these pathways to selection; and (2) incorporate estimates of mate quality into the definition of mating success to bring the operational definition of premating sexual selection in line with our conceptual understanding.

AUTHOR CONTRIBUTIONS

J.M.H. conceived the ideas and wrote the manuscript; J.M.H. and L.E.B.K. performed the analyses for the worked examples; J.M.H., M.D.J., and L.E.B.K. revised the manuscript and approved its final form.

ACKNOWLEDGMENTS

We would like to thank Adam Jones and Kenyon Mobley for sharing their data and their insights, as well as Michael Morrissey, Luc Bussière, Nina McLean, Karoline Fritzsche, and the Evolutionary Ecology journal club at the Australian National University, for their invaluable feedback on the manuscript. Thank you also to Associate Editor Jacob Moorad and two anonymous reviewers, whose comments greatly improved the manuscript. Funding was provided by the Australian Research Council, the Gwendolyn Woodroofe PhD Scholarship in the Biological Sciences, and an Ernst Mach Grant from the Österreichischer Austauschdienst.

DATA ARCHIVING

The doi for our data is https://doi.org/10.5061/dryad.1fp7830.

LITERATURE CITED

Anthes, N., I. K. Häderer, N. K. Michiels, and T. Janicke. 2017. Measuring and interpreting sexual selection metrics: evaluation and guidelines. Meth. Ecol. Evol. 8:918–931.

Arnold, S. J. 1994. Bateman's principles and the measurement of sexual selection in plants and animals. Am. Nat. 144:S126–S149.

Arnold, S. J., and D. Duvall. 1994. Animal mating systems: a synthesis based on selection theory. Am. Nat. 143:317–348.

Broquet, T., A. Barranger, E. Billard, A. Bestin, R. Berger, G. Honnaert, and F. Viard. 2015. The size advantage model of sex allocation in the protandrous sex-changer *Crepidula fornicata*: role of the mating system, sperm storage, and male mobility. Am. Nat. 186:404–420.

Collet, J. M., D. S. Richardson, K. Worley, and T. Pizzari. 2012. Sexual selection and the differential effect of polyandry. Proc. Natl. Acad. Sci. USA 109:8641–8645.

Collet, J. M., R. F. Dean, K. Worley, D. S. Richardson, and T. Pizzari. 2014. The measure and significance of Bateman's principles. Proc. R Soc. B 281:20132973.

Conner, J. K. 1996. Understanding natural selection: an approach integrating selection gradients, multiplicative fitness components, and path analysis. Ethol. Ecol. Evol. 8:387–397.

Conner, J. K., S. Rush, and P. Jennetten. 1996. Measurements of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). I. Selection through lifetime female fitness. Evolution 50:1127–1136.

Dai, C., and L. F. Galloway. 2013. Sexual selection in a hermaphroditic plant through female reproductive success. J. Evol. Biol. 26:2622–2632.

De Lisle, S. P., and E. I. Svensson. 2017. On the standardization of fitness and traits in comparative studies of phenotypic selection. Evolution 71:2313–2326.

de Villemereuil, P., H. Schielzeth, S. Nakagawa, and M. B. Morrissey. 2016. General methods for evolutionary quantitative genetic inference from generalized mixed models. Genetics 204:1281–1294.

Evans, J. P., and F. Garcia-Gonzalez. 2016. The total opportunity for sexual selection and the integration of pre- and post-mating episodes of sexual selection in a complex world. J. Evol. Biol. 29: 2338–2361.

Finney, S. J., and C. DiStefano. 2006. Non-normal and categorical data in structural equation modeling. Pp. 269–314 *in* G. R. Hancock and R. O. Mueller, eds. Structural equation modeling: A second course. Information Age Publishing, Greenwich, Connecticut.

Fitze, P. S., and J.-F. Le Galliard. 2011. Inconsistency between different measures of sexual selection. Am. Nat. 178:256–268.

Fitzpatrick, C. L. 2015. Expanding sexual selection gradients; a synthetic refinement of sexual selection theory. Ethology 121:207–217.

Frank, S. A. 2013. Natural selection. VI. Partitioning the information in fitness and characters by path analysis. J. Evol. Biol. 26:457–471.

Gerlach, N. M., J. W. McGlothlin, P. G. Parker, and E. D. Ketterson. 2012. Reinterpreting Bateman gradients: multiple mating and selection in both sexes of a songbird species. Behav. Ecol. 23:1078–1088.

Haaland, T. R., J. Wright, B. Kuijper, and I. I. Ratikainen. 2017. Differential allocation revisited: when should mate quality affect parental investment? Am. Nat. 190:534–546.

Hadfield, J. 2008. Estimating evolutionary parameters when viability selection is operating. Proc. R Soc. B 275:723–734.

———. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J. Stat. Softw. 33:1–22.

Henshaw, J. M., A. T. Kahn, and K. Fritzsche. 2016. A rigorous comparison of sexual selection indexes via simulations of diverse mating systems. Proc. Natl. Acad. Sci. USA 113:E300–E308.

Henshaw, J. M., and Y. Zemel. 2017. A unified measure of linear and nonlinear selection on quantitative traits. Meth. Ecol. Evol. 8:604–614.

Hereford, J., T. F. Hansen, and D. Houle. 2004. Comparing strengths of directional selection: how strong is strong? Evolution 58:2133–2143.

Houle, D. 1992. Comparing evolvability and variability. Genetics 130: 195–204.

Houle, D., C. Pélabon, G. P. Wagner, and T. F. Hansen. 2011. Measurement and meaning in biology. Q Rev. Biol. 86:3–34.

Janicke, T., and E. H. Morrow. 2018. Operational sex ratio predicts the opportunity and direction of sexual selection across animals. Ecol. Lett. 21:384–391.

Janicke, T., M. G. Ritchie, E. H. Morrow, and L. Marie-Orleach. 2018. Sexual selection predicts species richness across the animal kingdom. Proc. R Soc. B 285:20180173.

Janicke, T., I. K. Häderer, M. J. Lajeunesse, and N. Anthes. 2016. Darwinian sex roles confirmed across the animal kingdom. Sci. Adv. 2:e1500983.

Jennions, M. D., H. Kokko, and H. Klug. 2012. The opportunity to be misled in studies of sexual selection. J. Evol. Biol. 25:591–598.

Jones, A. G. 2009. On the opportunity for sexual selection, the Bateman gradient and the maximum intensity of sexual selection. Evolution 63: 1673–1684.

Kingsolver, J. G., and D. W. Schemske. 1991. Path analyses of selection. TREE 6:276–280.

Kline, R. B. 2016. Principles and practice of structural equation modeling. 4th ed. Guilford Press, New York.

Klug, H., J. Heuschele, M. D. Jennions, and H. Kokko. 2010. The mismeasurement of sexual selection. J. Evol. Biol. 23:447–462.

Krakauer, A. H., M. S. Webster, E. H. Duval, A. G. Jones, and S. M. Shuster. 2011. The opportunity for sexual selection: not mismeasured, just misunderstood. J. Evol. Biol. 24:2064–2071.

Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the "animal model." Phil. Trans. R Soc. B 359:873–890.

Kruuk, L. E. B., T. Clutton-Brock, and J. M. Pemberton. 2014. Case study: quantitative genetics and sexual selection of weaponry in a wild ungulate. Pp. 160–176 *in* A. Charmantier, D. Garant, and L. E. B. Kruuk, eds. Quantitative genetics in the wild. Oxford Univ. Press, Oxford.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:402–416.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Latta, R. G., and C. McCain. 2009. Path analysis of natural selection via survival and fecundity across contrasting environments in *Avena barbata*. J. Evol. Biol. 22:2458–2469.

Lefcheck, J. S. 2016. piecewiseSEM: piecewise structural equation modelling in R for ecology, evolution, and systematics. Meth. Ecol. Evol. 7: 573–579.

Loehlin, J. C. 2004. Latent variable models: An introduction to factor, path, and structural equation analysis. 4th ed. Lawrence Erlbaum Associates, Mahwah, NJ.

Lorch, P. D. 2005. Using upper limits of "Bateman gradients" to estimate the opportunity for sexual selection. Integr. Comp. Biol. 45:924–930.

Lynch, M. 1999. Estimating genetic correlations in natural populations. Genet. Res. 74:255–264.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Martin, A. M., M. Festa-Bianchet, D. W. Coltman, and F. Pelletier. 2015. Comparing measures of breeding inequality and opportunity for selection with sexual selection on a quantitative character in bighorn rams. J. Evol. Biol. 28:223–230.

McDonald, G. C., and T. Pizzari. 2018. Structure of sexual networks determines the operation of sexual selection. Proc. Natl. Acad. Sci. USA 151:E53–E61.

Milenkaya, O., N. Weinstein, S. Legge, and J. R. Walters. 2013. Variation in body condition indices of crimson finches by sex, breeding stage, age, time of day, and year. Conserv. Physiol. 1:cot020.

Mitchell-Olds, T., and R. G. Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. Evolution 41:1149–1161.

Mobley, K. B. 2014. Mating systems and the measurement of sexual selection. Pp. 99–144 *in* K. Yasukawa, ed. Animal behavior: How and why animals do the things they do. Praeger, Santa Barbara, CA.

Mobley, K. B., and A. G. Jones. 2009. Environmental, demographic, and genetic mating system variation among five geographically distinct dusky pipefish (*Syngnathus floridae*) populations. Mol. Ecol. 18:1476–1490.

Mobley, K. B., and A. G. Jones. 2007. Geographical variation in the mating system of the dusky pipefish (*Syngnathus floridae*). Mol. Ecol. 16: 2596–2606.

Mobley, K. B., and A. G. Jones. 2013. Overcoming statistical bias to estimate genetic mating systems in open populations: a comparison of Bateman's principles between the sexes in a sex-role-reversed pipefish. Evolution 67:646–660.

Mobley, K. B., M. Abou Chakra, and A. G. Jones. 2014. No evidence for size-assortative mating in the wild despite mutual mate choice in sex-role-reversed pipefishes. Ecol. Evol. 4:67–78.

Moorad, J. A., and M. J. Wade. 2013. Selection gradients, the opportunity for selection, and the coefficient of determination. Am. Nat. 181:291–300.

Morrissey, M. B. 2014. Selection and evolution of causally covarying traits. Evolution 68:1748–1761.

Morrissey, M. B., L. E. B. Kruuk, and A. J. Wilson. 2010. The danger of applying the breeder's equation in observational studies of natural populations. J. Evol. Biol. 23:2277–2288.

Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Meth. Ecol. Evol. 4:133–142.

O'Hara, R. B. 2009. How to make models add up—a primer on GLMMs. Ann. Zool. Fennici 46:124–137.

Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. Evolution 46:616–626.

Reed, T. E., P. Gienapp, and M. E. Visser. 2016. Testing for biases in selection on avian reproductive traits and partitioning direct and indirect selection using quantitative genetic models. Evolution 70:2211–2225.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. Anim. Prod. 8:95–108.

Robertson, A. 1968. The spectrum of genetic variation. Pp. 5–16 *in* R. C. Lewontin, ed. Population biology and evolution. Syracuse Univ. Press, Syracruse, NY.

Rosseel, Y. 2012. lavaan: an R package for structural equation modeling. J. Stat. Softw. 48:1–36.

Scheiner, S. M., K. Donohue, L. A. Dorn, S. J. Mazzer, and L. M. Wolfe. 2002. Reducing environmental bias when measuring natural selection. Evolution 56:2156–2167.

Scheiner, S. M., R. J. Mitchell, and H. S. Callahan. 2000. Using path analysis to measure natural selection. J. Evol. Biol. 13:423–433.

Sheldon, B. C., and H. Ellegren. 1999. Sexual selection resulting from extra-pair paternity in collared flycatchers. Anim. Behav. 57:285–298.

Shipley, B. 2016. Cause and correlation in biology: A user's guide to path analysis, structural equations and causal inference with R. 2nd ed. Cambridge Univ. Press, Cambridge.

Shuker, D. M. 2010. Sexual selection: endless forms or tangled bank? Anim. Behav. 79:e11–e17.

Stanton, M. L., and D. A. Thiede. 2005. Statistical convenience vs biological insight: consequences of data transformation for the analysis of fitness variation in heterogeneous environments. New Phytol. 166: 319–337.

Stinchcombe, J. R., A. K. Simonsen, and M. W. Blows. 2014. Estimating uncertainty in multivariate responses to selection. Evolution 68: 1188–1196.

Turnell, B. R., and K. L. Shaw. 2015. High opportunity for postcopulatory sexual selection under field conditions. Evolution 69:2094–2104.

Wade, M. J., and S. J. Arnold. 1980. The intensity of sexual selection in relation to male sexual behaviour, female choice, and sperm precedence. Anim. Behav. 28:446–461.

Walker, J. A. 2014. The effect of unmeasured confounders on the ability to estimate a true performance or selection gradient (and other partial regression coefficients). Evolution 68:2128–2136.

Walsh, B., and M. Lynch. 2014. Theorems of natural selection: results of Price, Fisher, and Robertson. *In* Evolution and selection of quantitative traits. Advanced copy: http://nitro.biosci.arizona.edu/zbook/NewVolume_2/newvol2.html.

Wilson, A. J., D. Réale, M. N. Clements, M. B. Morrissey, E. Postma, C. A. Walling, L. E. B. Kruuk, and D. H. Nussey. 2010. An ecologist's guide to the animal model. J. Anim. Ecol. 79:13–26.

Wisenden, B. D. 1999. Alloparental care in fishes. Rev. Fish. Biol. Fisher 9:45–70.

Worthington, A. M., and C. D. Kelly. 2016. Direct costs and benefits of multiple mating: are high female mating rates due to ejaculate replenishment? Behav. Process 124:115–122.

Wright, S. 1934. The method of path coefficients. Ann. Math. Stat. 5:161–215.

Associate Editor: J. Moorad
Handling Editor: P. Tiffin

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Path diagram for the worked example 'Cheating vs caring', shown with care effort *C*, ornament expression *O*, mating success *M*, and reproductive success *R*.

Figure S2. Estimated phenotypic and additive genetic variance-covariance matrices for the worked example 'Cheating vs caring'.

Table S1. Results of 'Cheating vs caring', showing phenotypic path coeffcients and estimates of sexual, remaining and total selection on care effort *C* and ornament expression *O*.

Table S2. Results of 'Cheating vs caring', showing genetic path coeffcients and the estimated evolutionary response to sexual, remaining and total selection on care effort *C* and ornament expression *O*.