COMMENTARY

The opportunity for sexual selection: not mismeasured, just misunderstood

A. H. KRAKAUER*, M. S. WEBSTER†, E. H. DUVAL‡, A. G. JONES§ & S. M. SHUSTER¶

*Department of Evolution and Ecology, University of California, Davis, CA, USA
†Laboratory of Ornithology and Department of Neurobiology and Behavior, Cornell University, Ithaca, NY, USA
‡Department of Biological Science, Florida State University, Tallahassee, FL, USA
§Department of Biology, Texas A&M University, College Station, TX, USA
¶Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA

Introduction

Biologists have long been interested in the prominent role that sexual selection plays in organismal evolution (Darwin, 1871; Fisher, 1929; Bateman, 1948). This key evolutionary process is thought to be responsible for many of the observed morphological and behavioural differences between the sexes (Andersson, 1994), including extreme ornaments that appear detrimental to survival (Promislow et al., 1992) and those that contribute to macroevolutionary processes such as speciation (Qvarnstrom et al., 2006; van Doorn et al., 2009). Demonstrating the action of sexual selection can be difficult, and theoreticians have developed a number of metrics to measure the actual or potential strength of sexual selection under varying ecological and social conditions.

Keywords:
Bateman gradient; mating system; operational sex ratio; selection gradient.

Abstract

Evolutionary biologists have developed several indices, such as selection gradients (β) and the opportunity for sexual selection (Is), to quantify the actual and/or potential strength of sexual selection acting in natural or experimental populations. In a recent paper, Klug et al. (J. Evol. Biol. 23, 2010, 447) contend that selection gradients are the only legitimate metric for quantifying sexual selection. They argue that Is and similar mating-system-based metrics provide unpredictable results, which may be uncorrelated with selection acting on a trait, and should therefore be abandoned. We find this view short-sighted and argue that the choice of metric should be governed by the research question at hand. We describe insights that measures such as the opportunity for selection can provide and also argue that Klug et al. have overstated the problems with this approach while glossing over similar issues with the interpretation of selection gradients. While no metric perfectly characterizes sexual selection in all circumstances, thoughtful application of existing measures has been and continues to be informative in evolutionary studies.

In a recent review of several of these methods, Klug et al. (2010a) strongly advocated for a trait-based approach using sexual selection gradients or differentials and were critical of fitness variance-based approaches such as the opportunity for sexual selection (Is = Imates Crow, 1958; Wade, 1979; Shuster & Wade, 2003) and the Bateman gradient (Arnold & Duvall, 1994), as well as population-based approaches such as the operational sex ratio (OSR; Emlen & Oring, 1977). Specifically, Klug et al. argued that (i) Is, the Bateman gradient and the ORS do not reflect actual selection on traits, (ii) these metrics may fail to capture meaningful variation in reproductive success and (iii) researchers inappropriately employ these metrics to quantify the strength of sexual selection.

We agree with several of the points raised by Klug et al. (2010a), but we disagree strongly with their conclusion that mating system measures such as Is, the Bateman gradient and the ORS have no place in the study of sexual selection. In this paper, we argue that many of the flaws in the use of these metrics are overstated. We focus our discussion on the many uses of Is, because it is...
developed from a population-genetic selection framework, but also discuss other variance- and population-based metrics of sexual selection. We also suggest that Klug et al. take too restrictive a view of the goals of sexual selection studies, leading them to unnecessarily reject metrics based on population mating system parameters. Finally, we discuss some of the limitations in the use of selection gradients and suggest that over-reliance on this approach could create many of the same problems Klug et al. seek to avoid.

The utility of the opportunity for selection

We agree with Klug and colleagues that sexual selection gradients (β) and the opportunity for sexual selection (ls) are neither equivalent nor interchangeable. Sexual selection gradients (derived from the more general concept of selection gradients) measure how a given trait covaries with relative fitness, measured in terms of mating success or offspring produced (Lande & Arnold, 1983; Arnold & Duvall, 1994) and therefore reflect the direct (and/or indirect) strength of sexual selection acting on that trait. In contrast, the opportunity for sexual selection (derived from the more general concept of opportunity for selection, c.f., Crow, 1958) measures the variance in mating success divided by the squared mean fitness, or (and/or indirect) strength of sexual selection acting on that trait. In contrast, the opportunity for sexual selection may or may not serve to greatly increase variation in mating and/or reproductive success (Webster et al., 1995; Whittingham & Dunn, 2005; Freeman-Gallant et al., 2005). Measures of ls effectively integrate the contributions of mate choice, social interactions, mate monopolization and other factors affecting mating patterns in a single value, thereby providing a concise description of the distribution of fertilizations (Vanpe et al., 2008; Krakauer, 2008; Mobley & Jones, 2007). We are uncertain why Klug et al. (2010a) find this usage of ls ‘disconcerting’.

Second, when properly employed, ls is not only sufficient for measuring the potential for selection to occur, it is necessary for identifying whether sexual selection may operate at all (Hersch & Phillips, 2004). Estimates of ls are therefore fundamental to examining ecological and social factors that have the potential to influence sexual selection. Because it is standardized by mean fitness, ls is a unit-less fitness metric that can be compared across classes of individuals (e.g. males vs. females within a population) or across populations (e.g. those that show strong sexual dimorphism vs. those that do not). This measure describes the distribution of mating success within a population, and it is this variance in mating success that drives sexual selection (Arnold & Duvall, 1994; Arnold & Wade, 1984a; Shuster & Wade, 2003; Jones, 2009). Thus, ls can be used to separate populations or conditions where sexual selection can have a strong evolutionary effect from those where it cannot. This application of ls is useful for broad comparative studies seeking to understand general patterns of sexual dimorphism and ornamentation (Clutton-Brock et al., 1980; Bro-Jorgensen, 2007), and also for
theoretical and empirical studies seeking to understand how ecological and social factors shape the potential for sexual selection (Mobley & Jones, 2007; DuVal & Kempeanaa, 2008; Webster et al., 2007).

Klug et al. (2010a) specifically criticize the use of $I_s$ when the traits relevant to mate choice or intrasexual competition are unspecified. We disagree and argue instead that the opportunity for selection’s freedom from assumptions about specific targets of selection provides several advantages if the researcher is not focused exclusively on the evolution of a particular trait. Because $I_s$ is calculated from measures of total mating success or, when available, measures of the mean and variance in offspring numbers from parentage analyses (Wade & Shuster, 2004; Shuster, 2009), it integrates the potential of all the organism’s traits to contribute to selection. This feature is particularly useful when targets of selection are poorly known or are difficult to measure, or when patterns of actual selection may vary across time, space or social contexts (Moller & Jennions, 2002; Coleman et al., 2004; DuVal & Kempeanaa, 2008).

Moreover, when $I_s$ is calculated separately for males and females, these measures allow comparison of the potential strength of selection in each sex; the magnitude and sign of the sex difference in the opportunity for selection, $\Delta I = I_{\text{males}}-I_{\text{females}}$, identifies the direction and degree to which sexual dimorphism is expected to evolve (Shuster & Wade, 2003; Shuster, 2009). Bateman (1948) suggested that for many species the variance in mating success will be higher for males than for females, thus facilitating stronger sexual selection on male traits and behaviours used in inter-sexual attraction or intra-sexual competition. However, in many systems of particular interest to biologists, including brood parasites (Woolfenden et al., 2002), cooperative breeders (Hauber & Lacey, 2005) and sex-role-reversed species (Mobley & Jones, 2007), it may not be clear a priori whether males or females should exhibit higher reproductive potential. For example, behavioural evidence and the occurrence of male pregnancy suggested that the broad-nosed pipefish, Syngnathus typhle, was sex-role-reversed with respect to the intensity of sexual selection (Berglund et al., 1986). An examination of patterns of parentage showed that females displayed higher values of $I_s$ and other mating system metrics compared to males, further supporting the occurrence of sex-role-reversal in S. typhle and validating Bateman’s assertion that the sex experiencing stronger sexual selection should display higher variances in mating and reproductive success (Jones et al., 2000, 2005).

An extension of measures of the opportunity for selection is that $I_s$ can be partitioned into component parts to further examine which aspects of mate acquisition or monopolization may be most important in determining the overall variance in reproductive success (Arnold & Wade, 1984a,b; Webster et al., 1995). This partitioning can involve additive terms, such as the relative contributions of within pair vs. extra pair paternity to male reproductive success (Webster et al., 1995, 2007; Whittingham & Dunn, 2005), or the relative opportunity for sexual selection from male-male competition and female mate choice (DuVal & Kempeanaa, 2008). It can also include multiplicatively nested terms, such as the contributions of mate number, average mate fecundity and average paternity share (Webster et al., 1995; Freeman-Gallant et al., 2005; Webster et al., 2007; Krakauer, 2008). These terms, along with the associated covariance terms, indicate where most of the reproductive variance exists and whether trade offs may occur between fitness components, and therefore may indicate where selection has the greatest potential to act. Partitioning $I_s$ in this way provides greater resolution of the source, the possible intensity and the mechanisms by which sexual selection operates (e.g. Webster et al., 2007), providing the foundation for trait-based studies aimed at measuring sexual selection on phenotypes that affect mating success through increased social status, mate attraction or mate quality.

The utility of other metrics: OSR and Bateman gradients

We have focused on $I_s$, but similar arguments can be made for other population- and mating-system-based metrics that Klug et al. (2010a) seek to discard. For example, the OSR measures the number of sexually active males relative to the number of sexually active females in a population, and therefore is a useful measure of the potential level of competition among individuals of one sex for mating access to individuals of the other sex (Emlen & Oring, 1977). We agree with Klug et al. that the OSR is an inaccurate measure of selection intensity because it does not measure selection at all. Moreover, estimates of OSR must be used with some caution because instantaneous estimates of the intensity of competition tend to over-estimate the overall intensity of sexual selection throughout the breeding season (Shuster & Wade, 2003; Shuster, 2009). However, OSR does provide useful information on instantaneous levels of intra-sexual competition, conditions likely to be perceived and responded to by individuals within field and laboratory populations alike (Berglund et al., 1986; Mobley & Jones, 2007). When the OSR is strongly biased towards one sex, individuals of that sex are likely to compete intensely for mates (Emlen & Oring, 1977). Selection favouring behavioural or morphological traits that lead to success in this competition is expected to occur under these circumstances, even if individual mating success does not remain constant over the entire breeding season (Shuster & Wade, 2003).

Similarly, the Bateman gradient (Bateman, 1948) measures the covariance between reproductive success (number of offspring) and mating success (number of mates), and so describes the extent to which total
reproductive success is tied to the ability to acquire mates (Arnold & Duvall, 1994; Jones, 2009). When this gradient is steep, then fitness is strongly tied to mating ability, and we expect strong selection on traits that help an individual acquire mates. In contrast, when the gradient is near zero, then total fitness is relatively independent of mating success and precopulatory sexual selection is unlikely to be acting strongly on any trait(s), measured or unmeasured. Moreover, comparison of the magnitude and sign of Bateman gradients estimated for each sex provide a quantitative measure of the degree to which sexual conflict may exist within and between species (Wade & Shuster, 2005). In each of these cases, metrics for the potential strength of sexual selection provide insights regardless of whether specific traits used in competition for mates have been identified.

Reliability of the opportunity for selection

Klug et al. (2010a) raise two key objections to \( l_s \) (as well as other measures of sexual selection). Specifically, they argue that (i) \( l_s \) is frequently misused as a measure of the strength of selection, and (ii) even when it is not misused, it is unreliable and provides a poor measure of selection. In this section, we address both of these claims.

The first claim rests on an assumption that most researchers calculating \( l_s \) are ultimately (and exclusively) interested in natural or sexual selection on specific phenotypic traits. Klug et al. correctly point out that there need not be a correlation between the opportunity for selection and the selection gradient for a given trait, and we agree with this important point. However, many if not most papers examining \( l_s \) are interested in the total potential for sexual selection rather than the evolution of specific traits. Klug et al. support their case with several quotes (their box 2) suggesting frequent misuse of \( l_s \), but closer examination of the cited papers indicates that misuse of \( l_s \) is far less rampant than Klug et al. suggest. For example, statements from Shuster & Wade (2003) represent two quotations taken from a more than 500-page book describing many ways that calculations of \( I \) (and \( l_s \)) can be useful for understanding organismal ecology and evolution. Shuster and Wade state: ‘This is another reason why \( l_{mates} \) sets an upper limit on response to sexual selection: not all mating is differential with respect to male characters’ (p. 31, emphasis in original). Similarly, the DuVal & Kempenaers (2008) quotation, cited as erroneous by Klug et al., is factually correct in describing \( l_s \) as a measure of the opportunity for selection and in pointing out that this measure is distinct from the trait-based evolutionary response to selection.

The second claim made by Klug et al. – that \( l_s \) and related measures do a poor job of quantifying or characterizing sexual selection – is argued using a series of thought experiments and simulations. In these examples, Klug et al. assert that the strength of selection acting on a specific trait depends only on how fitness covaries with the trait (\( \beta \)), and does not depend on the variance in relative fitness, estimated by mating success (\( l_s \)), or on the numbers of competing males and females (OSR). For studies seeking to understand the evolution of specific traits, this is an excellent point. However, the simulations provided by Klug et al. do not negate the utility of \( l_s \) or OSR and are not sufficiently general to warrant removal of these measures from the evolutionary biologists’ toolbox.

We disagree that the apparent discrepancies Klug et al. point out between the expected and observed performance of \( l_s \) (e.g. in their Fig. 2) invalidate it as a useful measure of sexual selection. Contrary to the authors’ interpretation, we maintain that \( l_s \) does exactly what it should do in the hypothetical scenarios presented, in that \( l_s \) is high in scenarios where there is large potential for sexual selection to operate and low when there is little potential for sexual selection. Take, for example, the situation where the breeding sex ratio (OSR) becomes strongly biased towards males by exchanging one male for one female in a hypothetical small population (depicted in their Fig. 2d–f). Here, Klug et al. find fault with \( l_s \) because it is higher when the OSR is biased towards males, compared to when the OSR is unity (Fig. 2d), regardless of whether copulations are monopolized by a small number of males (Fig. 2e) or are ‘as egalitarian as possible’ (Fig. 2f) given the sex ratio constraint. To Klug et al., this result casts doubt on \( l_s \) and they conclude that ‘[u]sing \( l_s \) as a proxy for actual selection is therefore problematic’. But there is nothing problematic here at all: \( l_s \) is behaving exactly as it should, showing increasing potential for sexual selection (i.e. maximum strength of sexual selection) as one moves from a balanced OSR with weak bias in male mating success (\( l_s = 0.67 \)), to a strongly male-biased (2:1) OSR with monogamous mating by those few males able to obtain mates (\( l_s = 1.00 \)), to a strongly male-biased OSR combined with strong monopolization of mates by a small number of breeding males (i.e. polygyny, \( l_s = 3.00 \)). Indeed, competition for mates under a biased breeding sex ratio was one of the two major mechanisms that Darwin (1871) proposed for sexual selection in socially monogamous species (e.g. Price, 1984), as the strength of selection depends not on how ‘egalitarian’ the system is, but rather on variance in reproductive/mating success. \( l_s \) is a measure of the maximum covariance between trait values and mating success; it is a mathematical concept, not a sociological one, and there is no reason why \( l_s \) should coincide with qualitative assessments of whether a mating system is egalitarian or not.

Klug et al. also provide a second quantitative example using larger samples sizes (their Fig. 3), this time using simulations to calculate \( l_s \) and selection differentials (closely related to \( \beta \)) in populations of set size that differ in their OSR, and they present comparisons of these values using different rules to assign male mating success. Klug et al. observe that \( l_s \) and the selection differential are
not always strongly correlated with each other, and that $l_s$ tends to vary based on the social system when underlying selection is held constant. In particular, in cases where selection was independent of mating system, either when there was no selection (their Fig. 3a) or when selection was constant and directional (their Fig. 3b), $l_s$ increased considerably as the OSR became more male biased. Klug et al. take this as evidence that the relationship between $l_s$ and the strength of sexual selection is too inconsistent for $l_s$ to be used as a reliable proxy of sexual selection.

However, once again, and despite their claims to the contrary, here Klug et al. do an excellent job of showing how $l_s$ does indeed correlate with the selection differential under biologically realistic conditions. First, in two of their five simulations (Fig. 3c,d), the selection differential and $l_s$ correlate remarkably well; the intercept for $l_s$ is consistently higher than it is for the selection differential, but this is as expected because the opportunity for selection will always be larger in magnitude than the selection differential (Crow 1958). Second, many of the simulations depicted are based on assumptions that are biologically unrealistic and/or uninformative. For example, in the first simulation (Fig. 3a), ‘mating is random and no variation exists in the trait of interest’ — if mating is indeed random, then selection is not occurring and the maximum strength of selection ($l_s$) will clearly be far greater than the selection differential ($= 0$) under any population scenario; this merely illustrates a well-known property of $l_s$ and other measures for the opportunity for selection. Similarly, several simulations use unrealistically biased breeding sex ratios, which tend to generate unusually high values of $l_s$, whereas most natural systems exhibit far lower values of $l_s$ (Jones et al., 2001; Table 2, $l_s = 0.25–2.03$; Wade & Shuster, 2004; Table 1, $l_s = 1.36–8.99$; Tatarenkov et al., 2008, Table 4, $l_s = 0.16–4.52$).

Third, a simple explanation exists for all three of the uncorrelated examples: in each of these cases, Klug et al. have intentionally made the phenotypic difference between mating and nonmating males independent of the sex ratio, i.e. the relationship between OSR and the selection differential is flat (Fig. 3a,b,e,f). As is well known, for a covariance to be positive, both variables must have nonzero variances. However, in each of these simulations, the selection differential is a constant value, making the exercise of asking whether or not any variable covaries with the selection differential pointless. We suggest that this approach is similar to asking whether or not smoking causes lung cancer, but restricting attention to those people who do not have lung cancer. In that set of people, a researcher would find some people who do not smoke, some who smoke moderately, and some who smoke a great deal, but none of them would have lung cancer, so the researcher could conclude that smoking must not cause lung cancer. We find it telling that in the two cases in which Klug et al. did allow variation in the selection differential, $l_s$ tracked the selection differential perfectly.

We believe Klug et al. mischaracterize the use of $l_s$ in most studies of sexual selection, possibly because they rely too heavily on colloquial definitions of $l_s$ (i.e. the degree of ‘reproductive monopolization’ in a population); in actuality, $l_s$ is defined in terms of standardized variance (Arnold & Wade, 1984b,a; Shuster & Wade, 2003). Thus, although values for $l_s$ may deviate from expectations based on colloquial definitions, those values are actually consistent with the more precise and mathematical definitions as the opportunity (or potential) for selection to operate. As emphasized above, $l_s$ is not a measure for actual selection on any trait(s), and most of the opportunity for sexual selection will remain unrealized if mating patterns are random with respect to male traits. We further suggest that if entirely random patterns of fertilization between males and females are a common feature of most biological systems, then overestimation of selection by opportunity for selection measures will be the least of our frustrations!

Selection gradients are not a panacea

Klug et al. advocate strongly for the use of direct, trait-based measures of sexual selection (e.g. sexual selection gradients, arguing that these measures more accurately capture the strength of sexual selection acting on a trait. We agree that the sexual selection gradient is a powerful statistical measure of the covariance between a trait and relative fitness, and that this is an excellent approach if one is interested in the extent to which sexual selection favours a particular trait. However, Klug et al. allude only briefly to some of the problems associated with selection gradients. First, this approach requires an assumption that a researcher has correctly identified and appropriately measured the key trait that is under selection. If, for example, the researcher has failed to measure the trait that is the target of sexual selection (e.g. those used by females to select a mate), then the conclusion would be that sexual selection is acting weakly on the traits that were measured, but selection on any unmeasured traits would go undetected despite being potentially strong. In contrast, $l_s$ would indicate the potential strength of selection in the population; whether or not a high potential for sexual selection is typically matched by actual selection on phenotypic traits is an important empirical (but not theoretical) question that remains to be answered.

Second, the ability of researchers to detect selection gradients depends first on whether the variance in relative fitness within a population (i.e. the opportunity for selection) is sufficiently large for selection to operate on any trait, and secondarily on whether the traits that the experimenter identified are important. Hersch & Phillips (2004) have concluded ‘that power to detect total selection (the selection differential) depends on sample
size and the strength of selection relative to the opportunity of selection’. These authors recommend that ‘studies should be designed so that selection is analyzed on at least several hundred individuals, the total opportunity of selection be considered along with the pattern of selection on individual traits, and nonsignificant results be actively reported combined with an estimate of power’.

Third, and importantly, selection gradients are at their core a correlational analysis, and are therefore subject to the same difficulties that are inherent with any such approach, which include issues surrounding the study sample and degree of natural variation (Grafen, 1987, 1988). In particular, with a correlational selection gradient approach, it can be difficult to separate cause from effect: showing that a particular ornamental trait is correlated with mating success does not demonstrate that the trait causes high mating success. This difficulty can arise if the trait of interest is not a target of sexual selection (i.e. does not affect mating success), but is phenotypically or genotypically correlated with an unmeasured trait that is. Partial regression analysis can help circumvent this problem by separating direct from indirect selection on the trait (Lande & Arnold, 1983), but this does not help if the actual targets of selection have not been measured (as above). Moreover, it is possible that high mating success causes enhanced trait expression, as might be the case if expression of an ornament is mediated by hormones and increased attention from females leads to higher hormone levels (Safran et al., 2008), or the level of male courtship behaviour depends upon female proximity and interest (Patricelli & Krakauer, 2010). Thus, sexual selection gradients can suggest that a particular trait is a target of sexual selection; but if one truly wants to know whether the trait affects mating success, then experimental manipulation is needed (Lande & Arnold, 1983; Grafen, 1988; Wade & Kalisz, 1990).

Lessons learned

In summary, the review by Klug et al. attempts to clarify the distinctions among some metrics used to measure sexual selection in natural populations, and we agree with many of their points regarding limitations of population-based metrics that measure the opportunity for sexual selection. First, $I_s$ describes the variance in relative mating success within a population, thus quantifying the maximum potential fitness differential available to drive sexual selection on one or more traits: researchers must be careful not to interpret $I_s$ as a measure of the actual strength of selection acting on specific traits. This point warrants repeating: in fact, this is a warning that advocates of this measure have long made, beginning with Crow (1958). Second, although the dimensionless property of $I_s$ allows comparisons between species or populations, such comparisons must be conducted carefully, as $I_s$ is sensitive to differences in sample size as well as to researcher assumptions about the number of breeding adults in the population (Freeman-Gallant et al., 2005). Direct comparisons across populations must be interpreted with caution to the extent that these factors differ substantially among studies.

Although we agree with some of the caveats raised by Klug et al. (2010a), we disagree strongly with their conclusion that measures such as $I_s$, OSR, and the Bateman gradient are useless or misleading, or that they are regularly misused as substitutes for the strength of selection. Indeed, numerous studies have employed these measures to elucidate the operation of sexual selection, including studies by authors of the Klug et al. paper (e.g. Klug et al., 2010b; Anthes et al., 2010). Measures such as $I_s$ can provide an integrative snapshot of the maximum fitness differentials available to drive selection on one or more traits, and can be used to better understand the effects of demographic, social and ecological factors on the maximum possible strength of sexual selection. $I_s$ and related measures also provide a handy metric to compare across populations or mating systems. Further, $I_s$ values are appropriate for some comparative studies, particularly those comparing the opportunity for selection between males and females or between distinct classes of males, because computational artefacts will be limited when all samples come from the same population. We also see great utility in the partitioning $I_s$ into its relevant component parts, which can shed light on the effects of social and ecological factors on the potential for various selective mechanisms to operate (e.g. Webster et al., 2007). Finally, studies of sexual selection can be complicated by numerous factors, including but not limited to variation in female preferences, environments, densities and metapopulation dynamics (Cornwallis & Uller, 2010); the opportunity for selection is one valuable tool for making sense of these intricate real-world systems. Contrary to the claims made by Klug et al., we believe biologists are, in general, correctly interpreting studies involving the opportunity for selection. We hope this debate will further clarify the proper use and interpretation of measures of $I_s$.

Acknowledgments

We thank Hope Klug and Michael Wade for discussion, and the National Science Foundation for support. AHK was supported by the National Science Foundation (IOS-0925038) during the writing of this manuscript.

References


Received 30 August 2010; revised 28 April 2011; accepted 4 May 2011