

Measuring Sexual Selection: Why Bother?

A. Grafen

*Department of Zoology
Animal Behaviour Research Group
Oxford OX1 3PS, England*

Abstract. Most of what we know about sexual selection comes from Darwin (1871) who never measured it. There is a difference between measuring things to do with sexual selection and measuring sexual selection itself. Many interesting problems in sexual selection are historical questions, requiring the application of historical methods for their study. The interesting intraspecific empirical question about sexual selection is what mechanisms of advantage favor extravagant characters. Fitness variance component analyses may have some weak connection with this. Quantification has no intrinsic virtues. A good field worker is nobody's poodle.

INTRODUCTION

Lord Kelvin, the second most distinguished Scottish physicist of the 19th century, was fond of asserting that until we can measure something we have no idea what we are talking about (Thomson 1889, 73). It was also Lord Kelvin who cast grave doubts on Darwin's theory of natural selection by proving that the earth was less than a hundred million years old (thus proving that even when we can measure something we may still not know what we are talking about) (Burchfield 1975). Darwin's thorough and logically sound but unquantitative methods have triumphed in the end over Lord Kelvin's calculations. The morals I draw for students of sexual selection are that quantification has no intrinsic virtues, is associated with the vice of hubris, and can set back the cause of science by giving a false air of great certainty of conclusion.

My starting point in this essay is therefore that there *may* be virtues in measuring sexual selection, but they need to be discovered, and the time and effort invested in measurement need to be justified. I begin with a consideration of the phrase "measuring sexual selection," which can be understood in two ways. The first is that it refers loosely to "measuring things that may be relevant to sexual selection," such as the length of a peacock's tail or linkage disequilibrium, and the second is that the quantity measured is "sexual selection itself."

There are dangers in claiming to measure sexual selection itself. What would we be tacitly asserting by taking some measurable quantity and naming it sexual selection? First that there exists one single quantity that captures the essence of sexual selection and second that we know what it is. There is certainly scope for argument on the second point (Wade and Arnold 1980; Sutherland 1985b), but even the existence of such a measure seems doubtful. Only very strong theoretical reasons could make us confident that our measures would be sensible in every species, with breeding systems of possibly as yet unobserved kinds.

If we decided that exactly one particular kind of numerical comparison between species, relevant to sexual selection, was always sensible and interesting, then calling that numerical measure “sexual selection” might not cause too much confusion. However, if there were no such measure, or more than one such measure, it would be best not to. It would be completely the wrong way round to choose a measure of sexual selection on other grounds and then assert that its comparison between species must be interesting. A premature christening preempts the question of the value of comparisons, among others. There is, in any case, no advantage to be gained from calling any quantity sexual selection, rather than in describing it more objectively and then arguing for a proposed connection between it and sexual selection either in a particular case or in general.

I shall, therefore, interpret my assigned title as “Measuring things to do with sexual selection: why bother?” Darwin (1871) discovered almost everything important now known about sexual selection and did so without measurement. The discovery of sexual selection and the outline of its main features are in no need of quantification: the arguments stand in our day as they did in Darwin’s. We should remember that Darwin was always very careful in fitting his methods to his problems (Ghiselin 1969), that he conducted experiments in some of his enquiries, but that he eschewed decimal places and field studies in his work on sexual selection. This is why the question of my title requires an answer.

I shall argue that the desirability of precise mathematical modelling does not imply the desirability of measuring parameters of those models in the field. Then the classes of reasons given as justification for measuring sexual selection will be considered. I shall argue that the study of sexual selection is an historical study and assess the role of measurement on that basis. I will then present the case that the interesting thing about sexual selection from an intraspecific point of view is the mechanisms of advantage, the selective forces at work on a character. Then I will discuss the relevance to these mechanisms of advantage of analyses of fitness variance components. To close, advice is offered to theoreticians and empiricists on the course of future research.

MODELLING VERSUS FIELD STUDIES

The desirability of precise models of sexual selection is beyond dispute.

Lande's (1981) partial mathematization of Fisher's (1958) runaway process confirmed that despite contrary claims (O'Donald 1982) it could work as Fisher asserted, and also showed that open-ended preferences were not necessary for the runaway process. The main technical problem which the modelling solved was the difficulty in keeping track of what happened to the genetic variances in the male trait and female preference, and how they in turn affected the selection of those two characters. Before Lande's model it was necessary to take Fisher's cryptic word for it.

Other current debates about sexual selection in which precise modelling is necessary are a) can an extravagant male character be selected *because* it is deleterious, as Zahavi (1975, 1977) claimed, and b) can host-parasite interactions maintain enough variability in a species without paternal care to ensure that females should expend time and energy in choosing a male on the basis of the fitness of the offspring the mating will produce, as Hamilton and Zuk (1982) have claimed? To attempt these problems by verbal arguments alone seems doomed to failure. Mathematics' major roles are to assist communication between different theoreticians and to help individual theoreticians marshal their thoughts.

It is tempting to believe that if a mathematical model is important in clarifying theoretical problems, then it must be interesting to measure the parameters of that model in natural populations. This modeller's hubris may be unjustified. The point of the kind of modelling of interest here is not to make precise predictions about particular species, indeed it is not to predict at all. Even if we follow Popper's view of how science works, prediction is the touchstone by which the *scientificness* (not the truth, interest, or importance) of theories is judged, but prediction is not usually the purpose of a theory. The purpose of most evolutionary models is not to foretell the future but to explain the present. Convenient assumptions are often false, and particular cases have peculiarities that would require a tailor-made model.

Even when precise predictions are not made about particular cases, the verbal conclusions from a mathematical model will often be important in interpreting observations in the field. Indeed, strong justification is required for making, in terms of equations rather than through an intermediate stage of words, the link between a model and a field study. All those explicit and implicit *caeteris paribus* assumptions that make a model worthwhile as an abstraction must be dealt with as objections to the equational link.

REASONS GIVEN FOR MEASURING THINGS TO DO WITH SEXUAL SELECTION

I have tried to understand the reasons given for measuring sexual selection by Ralls (1977), Payne (1979), Wade (1979), Arnold (1983), Lande and Arnold (1983), Clutton-Brock (1983), and Arnold and Wade (1984a,b). The reasons may be grouped into five categories, which I will discuss in turn:

- 1) to test predictions of sexual selection theory,
- 2) to measure things about selective surfaces,
- 3) to replace rhetorical claims,
- 4) to provide security of argument through precision,
- 5) to explain sex differences.

Testing predictions of sexual selection theory is all very well if there are such predictions. Danger lies in interpreting theoretical discussions too literally. Some authors, for example, have been under the illusion that sexual selection theory predicts a greater variance in male lifetime reproductive success (LRS) than in female LRS in species whose males (but not females) possess extravagant characters (e.g., Payne 1979). Let us examine one reasonable sounding line of argument that reaches this conclusion and spot the flaws. First, a) a female red deer's production of offspring is limited by her ability to gestate and suckle, and therefore has small scope for big gains in fitness, making selection for extravagant characters unlikely; b) as one female's production of offspring cannot be much varied by extravagant characters, different females must have similar offspring production; and c) variance in LRS is low in females. Second, a) a male red deer's production of offspring is limited by his ability to gain and defend a harem, and by the harem's size — this makes likely selection of extravagant characters such as those for male combat, and possibly attractiveness to females; b) as one male's offspring production can be varied by extravagant characters, different males must have dissimilar offspring production; and c) variance in LRS is high in males. In both cases step a) is correct, step b) is wrong, and step c) is “wronger.”

Step a) is Darwin's original argument. Step b) simply doesn't follow. One flaw in the argument is the supposition that because LRS cannot be influenced by extravagant characters, it cannot be influenced at all. Factors exist which contribute to variability in LRS but which cannot be influenced by extravagant characters, and this invalidates the jump from a) to b). They include longevity, randomness (Sutherland 1985a), and individual differences in quality arising from parental and other influences while a juvenile. This means that around the core of variability, which is concerned with sexual selection, there may be much variability from other causes which would obscure comparisons.

The second flaw shows that even the core is rotten. It is wrong to conclude from the fact that LRS can be influenced by extravagant characters, that it must be variable. This step depends on the existence of variability *now* in the extravagant character. Step a) relies on the possibility of variability in LRS, not on its existence in the population now. During selection for antlers, we expect that there was variability in antler length and consequent variability in LRS. If antlers are no longer growing in evolutionary time, then the amount of heritable variability in antler length depends presumably on a balance between selection and mutation. The magnitude of this variability in LRS will depend

on mutation rates, which are not obviously relevant to a measure of the scope for selection of extravagant characters.

I described step c) as “wronger” than step b), but I have just explained that step b) is completely wrong. This is because using the word variance gives the predictions a false air of scientificness and precision. Step a) is correct as a verbal argument and does not need to be made precise. Even if step b) were admitted, there is nothing to say that when it was made precise it would result in step c). Variance may not be the best measure of dissimilarity. The point of having precise predictions is lost if they are spuriously precise in the manner of step c). The invalidity of steps b) and c) shows the dangers of making perfectly good verbal arguments precise for the sake of it.

The measurement of adaptive landscapes is advanced by Lande and Arnold (1983) as a good thing. The extent of maladaptation, and the frequency of multiple peaks, are facts of general interest to geneticists, particularly to those of a Wrightian bent. The topic of this volume is sexual selection and in this context, while no one would throw away measurements of an adaptive landscape, they do not seem to me to be of burning interest. Others may disagree.

A second argument of Lande and Arnold is that measurement is “the best alternative to the fabrication of adaptive scenarios Our optimism that measurement can replace rhetorical claims for adaptation is founded on the growing success of field workers in their efforts to measure major components of fitness in natural populations.” In my view the construction of adaptive scenarios is half of what evolutionary biology is about, and their validation is the rest. What can we make of Lande and Arnold’s rhetorical claims for measurement?

Presumably the kind of replacement they have in mind is adding to an unsupported statement of the form “Character *X* is an adaptation for purpose *Y*” a set of persuasive measurements, which should include measurements of major components of fitness in natural populations. Supporting facts are good things but they are not always necessary. Furthermore, for claims about adaptations, measurements of major components of fitness in natural populations may be relevant only rarely, as hinted by the following example.

No measurement is needed to persuade us that eyes are for seeing. They may also be for the detection of the earth’s magnetic field (Leask 1977). How would we test such a claim? Neurophysiological experiments on eyes, and on the optic nerves and their destination areas in the brain, would be one way (Semm et al. 1984). Testing the ability of organisms to navigate in the presence and absence of light is another as the proposed photochemical mechanism requires light in the eye. Field studies on natural variability in eyes, and its contribution to LRS, would not be high on my list of suitable methods of investigation. There are, then, at least some kinds of claim about adaptation for which field measurements of components of fitness in natural populations are irrelevant. I have argued elsewhere that these fitness measurements may be more generally

rather irrelevant to the study of complex adaptations (Grafen 1988).

The fourth argument is that without measurement our arguments are insecure through imprecision. Taken as referring to theoretical arguments this is a plea for precise modelling, whose justice is discussed and affirmed above. Taken as referring to interpretation of data this rather depends on the kind of assertion desired. Certainly, no one would wish to learn only that a sex ratio “seemed to be female biased” when counts of individuals by sex were available. On the other hand, to try to prove by measurement that the male peacock’s plumage is more extravagant than the female’s is unnecessary. The truth of the statement can be admitted even though the meaning of “extravagant” is still imprecise, and in many ways is better left so. The idea that seems to me to underlie the previous two arguments in favor of measuring is that “you have to be numerically precise to be right” — and I tackle this general philosophy in the next section.

We come at last to the only good reason for studying sexual selection at all, and *a fortiori*, the only good reason for measuring it (if there is one). Darwin invented sexual selection to account for puzzling characters, often associated with sex differences, that seemed to hinder males (usually) in the struggle for existence. He used it to explain much of bird coloration, antlers, human hairlessness, and much sexual dimorphism in size. Measuring is desirable to the extent that it is helpful in pursuing Darwin’s aim in a logical way. Cases in which it is helpful are discussed at the end of the next section.

YOU DON’T HAVE TO BE NUMERICALLY PRECISE TO BE RIGHT

In considering the role of measurement in the study of sexual selection, it is helpful to consider which methods have been of most value so far. These are the arguments of Darwin (1871). They are of two main types: speculation as to possible mechanisms within a species, and the detection and interpretation of pattern usually between species. With this pair of techniques, Darwin invented sexual selection and discovered all its main features. I shall discuss why these techniques are so unnumerical. Then I shall outline what I consider to be a set of currently interesting problems in sexual selection and how measurement is likely to fit into the logic of their resolution.

Darwin’s techniques are unnumerical partly because measurements were not available and the statistical techniques which make measurement more valuable were not available. But the major reason is that Darwin was interested in showing the truth of historical statements, and that is because many of the questions of interest about sexual selection are historical. For example, is the generally brighter coloration of male birds explained by sexual selection? Is sexual dimorphism in ungulates the result of sexual selection? Is human hairlessness the result of sexual selection? These exemplify one central type of interesting statement about sexual selection, perhaps the most interesting type.

These questions are answered in much the same way as questions of the sort: “Does the success of a popular revolution depend on the existence of an educated and sympathetic middle class?” First, some plausible case must be made about the likely effect of the education and sympathy of the middle class on a popular revolution. Ideally, documentary evidence from particular examples would support this case. Second and most important, a pattern should be sought in the outcomes of all sufficiently studied popular revolutions. Can examples be found with an educated but unsympathetic middle class, with an uneducated but sympathetic middle class, and with neither? Do the ways in which those categories of revolution fail support the ideas about why education and sympathy are necessary?

Measurement is relevant to these questions, but obliquely and not centrally. The education of the middle classes could be measured by the number of years in school, by literacy levels, or by the number of books or magazines sold per head. Different measures might be necessary for different cases because of differences in the ways the original studies were carried out. The same measure might have to be interpreted in different ways because of special circumstances; for example, the number of books sold per head will depend on the cost and therefore available technology of production.

Darwin’s arguments about sexual selection are of this type in which measurement is relevant obliquely, not centrally. The role of measurement is the same today, as I argue now in a brief discussion of my personal list of interesting topics in sexual selection:

- 1) Female choice or male competition?
- 2) Number or quality of females?
- 3) Does male competition lead to organs designed as weapons or displays?
- 4) Is female choice for genes or for paternal care?
- 5) If for genes, what maintains the genetic variability among males?
- 6) Does sexual selection work on both sexes at once?
- 7) Does sexual selection do for a species more of what sex already does?

These questions may have different answers for different species. Interesting answers would likely not be yes or no, but an account of in what circumstances the answer was yes, and when no, and an explanation. My concern here, however, is how measurement is relevant to answering these questions.

Answering the questions will involve recognizing a pattern or design. Take the question of whether female choice or male competition is responsible for sexually selected male body size in *Drosophila*. It is imaginable that a female mates in the same time with a male of any size when confined with him alone, but generally mates with the larger when confined with two males. This would suggest that male competition and not female choice was important. Alternatively, a comparison between different species of *Drosophila* might reveal that the extent of the influence of sexual selection on male body size correlated well with the scope for male competition but not with the scope for female choice. I

stress these are hypothetical possibilities. These obvious possible ways of answering the question illustrate that measurement is obliquely relevant in that it may be necessary to measure to detect a pattern. Measurement is not central: there is no single quantity, not even one whose precise definition escapes our minds now but would become clear after much thought, whose measurement is the answer to the question.

Whether sexual selection acts in both sexes at once could be answered by comparative studies in the Darwinian tradition. A second reason for measurement is that the comparative method has changed since Darwin's day. A modern day exponent of this method has more measurements available, because more studies have been performed. This means that he can study finer patterns than Darwin. He is also more concerned about taxonomic artifacts, and the various methods designed to avoid them require quantification to greater or lesser degrees (Ridley 1983; Felsenstein 1985). He also lives in a community with a fetish for statistics. (By fetish I mean to suggest an irrationally strong attraction coupled with lack of understanding, as of magic.)

A third reason, which could apply to any of the questions, is to have a well-worked-out example. That old war-horse *Biston betularia* is didactically invaluable. Nothing of special interest was proved about natural selection, no new principles were discovered. To have more than a few examples of this kind would be wasteful, for it is not the job of biologists to do the same thing over and over again with different examples. But one or two well worked-out examples of how sexual selection operates through female choice, how it operates through male competition, how it can operate for good genes, how it can operate for paternal care, and so on, would be very welcome. They help in teaching, they help to clarify thought about the processes involved, and they may also turn up interesting quirks of natural history on the way. The best examples I know of in sexual selection are Endler's guppies (1980) and Andersson's widow birds (1982). Both could be improved upon as classic instances of sexual selection, but they will not easily be surpassed.

The major reason for measuring is then to assist in the recognition of patterns. Any kind of pattern will do, intraspecific, interspecific, or an anatomical pattern that we would call design. Any of these could persuade us of an important point about sexual selection, and all three techniques were used by Darwin. The logic of being persuaded by a pattern is that something intrinsically implausible demands an explanation. The nature of that explanation will ideally be closely dictated by the nature of the pattern itself.

Measuring in some other fields has a much stronger place than it does in the study of adaptations under natural and sexual selection. The body of theory in physics is very strong and is tightly related to questions of central importance in physics. We do not have such a body of theory in the study of sexual selection, and I guess never will have. This is not a matter for regret, as it arises naturally from the intrinsically historical nature of most of the interesting questions about sexual selection.

Reading discussions of exactly how sexual selection should be measured (e.g., Payne 1979; Wade 1979; Wade and Arnold 1980; Arnold 1983; Arnold and Wade 1984b; Sutherland 1985b) makes me imagine historians debating whether the significance of the Boston Tea Party should be assessed by the wet or the dry weight of the tea thrown overboard. It is not necessary to be numerically precise to be right.

WHAT A GOOD INTRASPECIFIC EMPIRICAL STUDY OF SEXUAL SELECTION SHOULD DO

The main points are that the mechanisms of advantage of the character should first be discovered, and that the question of which mechanisms count as natural and which as sexual selection should be answered afterwards, if at all.

In an imaginary crustacean species, females have a reproductive cycle and lay a batch of eggs every two weeks. They must mate in each cycle to lay fertilized eggs. Males vary in size fourfold, and the future size of young males can be determined by measuring their size at an early moult. Females are concealed in the mud at the bottom of a stream, where they spend their time feeding. Adult males search for females in the mud, helped by the detection of chemicals involuntarily released by females. Once found a female may or may not consent to mate. The consequences of male adult size are studied and it is found by making appropriate measurements that

- 1) young males trying to become large adults suffer higher juvenile mortality;
- 2) large adult males can search twice as fast because of their superior powers of locomotion;
- 3) large adult males are in addition twice as good at detecting females because their larger sensory apparatus can detect in lower concentrations the chemicals released by the females;
- 4) large adult males are twice as likely to be accepted as mates by females because females resist mating except near their moult and large males are better able to overcome this resistance;
- 5) females are only half as likely to remate during a cycle after mating with a large male than after mating with a small male (this is because the males insert a plug at the end of mating to prevent future matings; larger males' plugs are more efficient and there is full sperm precedence for the last male to mate);
- 6) large males live half as long as adults because the major predator ambushes prey, and moving twice as fast makes capture twice as likely per unit time.

Given these facts, I feel I would understand the comparative advantages of being large and small. Division into sexual and natural selection is not entirely obvious. Some traits are likely to be sexually selected, such as the sensory apparatus and the mating plugs. Whether male size is sexually selected is a scholastic question I am not sure how to answer, and I am not absolutely sure I want to. It might depend on whether the sensory apparatus and powers of

locomotion and efficiency of mating plug were simply allometric with body size, or whether they improved qualitatively.

The point of this section is that *the interest of an empirical study can be extracted without having to decide whether a mechanism counts as sexual or natural selection*. Darwin's statement (1871, quoted by Arnold 1983) that ". . . in many cases it is scarcely possible to distinguish the effects of natural and sexual selection" is a sign that imprecision of definition was intended: I believe it is also desirable.

THE RELEVANCE OF ANALYSES OF COMPONENTS OF VARIANCE IN LRS

There are a number of suggested analyses of simple variance in LRS into components (Arnold and Wade 1984b; Brown 1988). (I intend to discuss only direct variance analyses and not techniques in which components of fitness are regressed on characters such as body size. The relationship between characters and fitness is of obvious biological interest). What is the relationship between such an analysis and an outline of mechanisms like that above? The analysis can be performed without any idea of the mechanisms, and the mechanisms can be studied without bothering to do the analysis or even having the data to do one. I suggest that the analysis can be considered as an accounting exercise, in which the effects of the mechanisms are coded into numbers in a complex way, rather as when a firm's assets are written down in a ledger with a valuation.

I will pursue this analogy. In investigating a firm to work out why it has been performing well recently, some questions are best answered by going over the books. What were last year's profits? Have costs been coming down, or have prices been going up? Is the volume of sales changing? But no inspection of the accounts will reveal that the reason more ice cream has been sold is the hot weather, nor that the reasons the costs have come down is that squabbles between OPEC countries have reduced the price of petrol. Now, of course, these factors are reflected *in a way* in the accounts but seeing that reflection does not really help us understand the reason an ice cream manufacturer does well in hot weather when a major energy cartel breaks down. The point here is therefore that the fitness analysis may be useful for some purposes, but very unhelpful for others.

The accounting is unhelpful because the factors of real interest are scattered in different parts of the accounts, contributing to many different headings. The weather will affect demand, the costs of heating factories, and the absentee rate among the work force. The weather could not be reconstructed from the accounts, as each of these accounting headings is also affected by many other factors.

Consider how well a firm is described by its accounts. Seen in one way, every aspect of the firm is represented there and evaluated relative to all other

aspects. In another way, it is fair to say that many vital aspects of the firm are missing entirely, or if they are represented are present in a mangled and distorted way.

Understanding the reasons for a sex difference is an aspect of a species' biology not present in the fitness component analyses, although those reasons will doubtless have their influence on some of the numbers. As this is the only purpose which I consider to be a good reason for measuring things to do with sexual selection, does this mean I think that fitness variance component analyses are not worth doing?

I have given much thought to this question and can find only rather feeble arguments in favor of those analyses. The first is that they are a contribution to the natural history of the species. Of course, so would any other kind of measurement or analysis undertaken. The second is that the analysis might point to where mechanisms are likely to be found. If there is much variability between male toads in number of mates per breeding season, and rather little in number of breeding seasons, then it might be worth looking for causal connections between numbers of mates and size or beauty, while looking for causal connections with longevity are likely to be less worthwhile. These are rather weak clues, as we saw before that variance at a given stage has no necessary connection with sexual selection. So while this might be a way to use an analysis that exists, it is unlikely to be worth collecting all the necessary data just for these weak clues. Notice that having compiled all the information for the fitness variance component analysis, following up the clues involves collecting more facts. The data required for the analysis are complete in an accounting sense but by no means complete for answering questions about mechanisms of advantage.

To conclude this section, the analyses of fitness variance components suggested by various authors have little to do with the mechanisms of advantage, which are to my mind the main targets in a study of sexual selection. It is possible that I have missed the point of them. I would be persuaded of this if a close logical connection were demonstrated between a fitness variance component analysis and any important problem of sexual selection such as those I listed above. Meanwhile, my personal advice to anyone considering collecting data with the main purpose of performing such an analysis is the same as Punch's advice to those about to marry.

RECOMMENDATIONS FOR FUTURE RESEARCH

It is traditional in Dahlem conferences to give advice for future research, in which the roles of theoreticians and empiricists are pontificated upon. I see no reason not to indulge in this homely sport myself, to close this essay. I intend to redress a balance, not to be balanced myself.

To theoreticians, I would say: resist the temptation to make your models

more realistic. Your function now is to provide the material which will determine what field workers will find interesting in five years' time and beyond. Particular species are in themselves boring and deserve modelling only to make general points, or for didactic purposes. Particular genetic systems are also of interest only if used to make general points.

To field workers, I would say: whatever you do, don't go out and measure a parameter in somebody else's model. Few things could be more fruitless. Your role is not to be a theoretician's technician. Instead you are at least the theoretician's equal and would be better employed doing one of three things. First, study an as yet unmodelled problem. Show the world that this problem is interesting and important, and outline what you see as its main features. Let theoreticians then busy themselves, working out in more and more detail that you had indeed got it nearly all right in the first place. Meanwhile, you can have discovered another interesting problem. Make the theoreticians come to you. After all, your function now is to provide the material which will determine what theoreticians will find interesting in five years' time and beyond.

The second useful role of the field worker is to cut the Gordian knots of the theoreticians. For example, a debate is raging overhead about whether or not sufficient variability can be maintained in heritable fitness to make it worthwhile for a female to pay costs to choose a male solely for the inherited fitness of her offspring. Ultimately, this is a question that in theory will depend on the plausibility of various parameter values and details of the models. An empirical demonstration in reasonably natural conditions that fitness is often inherited from father to offspring in the absence of paternal care would radically change the nature of the debate. No longer would the plausibility of the conditions for maintaining this heritability be the hot topic, but rather what forces are actually responsible for it, and how strong are they? The empiricist must try to stop theoreticians wasting their time on empirical questions, if it is thought they could spend their time more profitably.

The third role of the empiricist is to undermine the theoretician by casting doubt on his accepted premises. What empirical demonstration would most embarrass this model? How can I make this theory seem truly pointless? And so on. Go for the jugular.

If after all three activities some field workers still have time to compile a fitness variance component analysis, then I take my hat off to their tirelessness.

REFERENCES

- Andersson, M. 1982. Female choice selects for extreme tail length in a widow bird. *Nature* 299: 818–820.
- Arnold, S.J. 1983. Sexual selection? the interface of theory and empiricism. In: *Mate Choice*, ed. P.P.G. Bateson, pp. 67–107. Cambridge: Cambridge Univ. Press.
- Arnold, S.J., and Wade, M.J. 1984a. On the measurement of natural and sexual selection: applications. *Evolution* 38: 720–734.

- Arnold, S.J. and Wade, M.J. 1984b. On the measurement of natural and sexual selection: theory. *Evolution* 38: 709–719.
- Brown, D. 1988. In: *Reproductive Success*, ed. T.H. Clutton-Brock. Chicago: Univ. of Chicago Press.
- Burchfield, J.D. 1975. *Lord Kelvin and the Age of the Earth*. New York: Science History Publications.
- Clutton-Brock, T.H. 1983. Selection in relation to sex. In: *Evolution from Molecules to Men*, ed. D.S. Bendall, pp. 457–481. Cambridge: Cambridge Univ. Press.
- Darwin, C. 1871. *The Descent of Man and Selection in Relation to Sex*. London: John Murray.
- Endler, J.A. 1980. Natural selection and color patterns in *Poecilia reticulata*. *Evolution* 34: 76–91.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Natural.* 125: 1–5.
- Fisher, R.A. 1958. *The Genetical Theory of Natural Selection*. New York: Dover.
- Ghiselin, M.T. 1969. *The Triumph of the Darwinian Method*. Berkeley: Univ. of California Press.
- Grafen, A. 1988. On the uses of data on lifetime reproductive success. In: *Reproductive Success*, ed. T.H. Clutton-Brock. Chicago: Univ. of Chicago Press.
- Hamilton, W.D., and Zuk, M. 1982. Heritable true fitness and bright birds: a role for parasites? *Science* 218: 384–387.
- Lande, R. 1981. Models of speciation by sexual selection of polygenic traits. *Proc. Natl. Acad. Sci.* 78: 3721–3725.
- Lande, R., and Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* 37: 1210–1226.
- Leask, M.J.M. 1977. A physicochemical mechanism for magnetic field detection by migratory birds and homing pigeons. *Nature* 267: 144–145.
- O'Donald, P. 1982. Sexual selection by female choice. In: *Mate Choice*, ed. P.P.G. Bateson, pp. 53–66. Cambridge: Cambridge Univ. Press.
- Payne, R.B. 1979. Sexual selection and intersexual differences in variance of breeding success. *Am. Natural.* 114: 447–452.
- Ralls, K. 1977. Sexual dimorphism in mammals: avian models and unanswered questions. *Am. Natural.* 111: 917–938.
- Ridley, M. 1983. *The Explanation of Organic Diversity*. Oxford: Clarendon Press.
- Semm, P.; Nohr, D.; Demaine, C.; and Wiltchko, W. 1984. Neural basis of the magnetic compass: interactions of visual, magnetic and vestibular inputs in the pigeon's brain. *J. Comp. Phys. A* 155: 283–288.
- Sutherland, W.J. 1985a. Chance can produce a sex difference in variance in mating success and account for Bateman's data. *Anim. Behav.* 34: 1349–1352.
- Sutherland, W.J. 1985b. Measures of sexual selection. *Oxford Surv. Evol. Biol.* 2: 90–101.
- Thomson, W. 1889. *Popular Lectures and Addresses*, vol 1. London: Macmillan.
- Wade, M.J. 1979. Sexual selection and variance in reproductive success. *Am. Natural.* 114: 742–746.
- Wade, M.J., and Arnold, S.J. 1980. The intensity of sexual selection in relation to male sexual behaviour, female choice and sperm precedence. *Anim. Behav.* 28: 446–461.
- Zahavi, A. 1975. Mate selection — a selection for a handicap. *J. Theor. Biol.* 53: 205–214.
- Zahavi, A. 1977. The cost of honesty. *J. Theor. Biol.* 67: 603–605.