



Optimality Models and the Test of Adaptationism

Steven Hecht Orzack; Elliott Sober

The American Naturalist, Vol. 143, No. 3. (Mar., 1994), pp. 361-380.

Stable URL:

<http://links.jstor.org/sici?sici=0003-0147%28199403%29143%3A3%3C361%3AOMATTO%3E2.0.CO%3B2-W>

The American Naturalist is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

OPTIMALITY MODELS AND THE TEST OF ADAPTATIONISM

STEVEN HECHT ORZACK* AND ELLIOTT SOBER†

*Department of Ecology and Evolution, 1101 East 57th Street, University of Chicago, Chicago, Illinois 60637; †Philosophy Department, University of Wisconsin, Madison, Wisconsin 53706

Submitted September 28, 1992; Revised April 2, 1993; Accepted April 19, 1993

Abstract.—The use of optimality models in the investigation of adaptation remains controversial. Critics charge that advocates of the optimality approach assume that the traits they analyze are optimal. Advocates of the approach deny this but admit to assuming that the traits have adaptive explanations. This controversy is part of the ongoing debate about adaptationism. We believe that this controversy remains unresolved in part because of ambiguity in the definition of adaptationism. In this article, we clarify the thesis of adaptationism, show how the structure of optimality models relates to that thesis, and describe how the thesis of adaptationism is testable. In addition, we describe the types of analyses that are essential to a test of an optimality model if the optimality of the trait is to be assessed and if assessments of the success of specific models are to contribute to a test of adaptationism. These analyses allow one to distinguish between the hypothesis that natural selection has had some influence or an important influence on a trait and the hypothesis that the trait is optimal. At present, to our knowledge, there are only two sets of studies in evolutionary biology in which this critical distinction has been made.

Although the controversy over adaptationism has gone on for many years, several important issues have not been resolved. First, adaptationism has not been clearly defined. Second, it has remained unclear as to how adaptationism is testable, how the testing of specific optimality models is involved in that test, and how the test of a specific model should be structured in order to assess the hypothesis of optimality. We address these issues in this article. Some of our discussion of optimality models will focus on evolutionarily stable strategy (ESS) models because the relevant theory illustrates an important distinction we will make concerning different hypotheses about the power of natural selection. Despite this focus, our arguments apply to any optimality model and, more generally, to any model used to support the claim that natural selection is the only important cause of a trait.

DEFINING OPTIMALITY

Instead of the population-genetic elaboration of genetic mechanisms and of modes of selection as a means of assessing the adaptiveness of traits, optimality analysis focuses on the determination of evolutionarily stable phenotypes. A phenotype of an individual is optimal (relative to a variety of alternatives) because it outperforms the other phenotypes and thereby results in a higher fitness. (The

fitness measure used will, of course, depend on the nature of the trait.) As a result, other phenotypes are eliminated from the population (or nearly so) or prevented from invading (see Maynard Smith 1982; Hines 1987; and Parker and Maynard Smith 1990 for further details). The rationale for the focus on phenotypes is the notion that natural selection will overcome any genetic or developmental constraints on an evolutionarily important trait. Perhaps the main motivation for the focus on stability is that many of the traits of interest to optimality modelers, such as feeding behaviors or sex ratios, are regarded as so clearly related to evolutionary success that it is reasonable to think that natural selection must lead to the appearance of an evolutionarily stable phenotype (cf. Eshel 1982).

DEFINING ADAPTATIONISM

Assessing the validity of the optimality approach is part of the larger project of evaluating the thesis of adaptationism. What does this thesis mean? In order to answer this question, it is useful to distinguish three propositions that concern some trait T of an *individual* in a given population:

- (U) Natural selection played some role in the evolution of T . (U stands for ubiquitous since we believe this proposition applies to most traits.)
- (I) Natural selection was an important cause of the evolution of T . (I stands for important.)
- (O) Natural selection is a sufficient explanation of the evolution of T , and T is locally optimal. (O stands for optimal.)

These propositions are presented in order of increasing strength. Proposition (O) entails (I) but not conversely, and (I) entails (U) but not conversely. The phrases *important cause*, *sufficient explanation*, and *locally optimal* require clarification, which we provide later.

To understand the contrast between (U) and (I), consider the dynamics of a trait controlled by two alleles A and a at a single diploid locus that is affected only by selection and genetic drift. Let A be completely dominant over a , where the fitnesses of AA and Aa are related to the fitness of aa as $1:1 - s$, where s is small but positive. Let N be the effective population size. Then the two alleles are said to be effectively neutral if $2Ns \ll 1$ (Crow and Kimura 1970). In this case, (U) is true but (I) is not. Natural selection occurs, but it is not an important cause of allele frequency change; that is, when predicting, say, the mean phenotype, natural selection can be ignored with no loss of accuracy. When the inequality is reversed, (U) and (I) are both true; natural selection not only occurs, but it is an important cause.

The distinction between (I) and (O) is illustrated by the case of a trait controlled by a single diallelic locus subject to natural selection in an infinite panmictic population. Suppose that the trait of the heterozygote has the highest fitness. Proposition (I) is true because natural selection cannot be ignored when predicting, say, the mean trait value. Yet since heterozygotes do not breed true, (O) is not true because this genetic constraint prevents fixation of the optimal trait.

There is a big difference between saying that natural selection was an important

cause of the evolution of a particular trait and saying that it was the only important cause. It is this contrast that underlies one difference between (I) and (O). In the model of selection and drift, for example, when $2Ns \approx 1$, it would be a mistake to predict the mean trait value by assuming that drift played no role. In this case, (I) is true but (O) is false. Natural selection is not a sufficient explanation of the trait's mean value, even though natural selection is an important factor at work. More specifically, we define the concept of sufficient explanation by reference to the predictions one obtains from a "censored" model (i.e., a model in which the only evolutionary force is natural selection). Optimality models are censored models of this kind. In effect, such a model says that there was no mutation pressure, that genetic drift did not significantly interfere with selective dynamics or with the attainment of the optimum by individuals, and so on. If the predictions of this model fit the observations in a statistical sense (see below), (O) can be regarded as true. Natural selection here provides a sufficient explanation because taking other factors into account could not significantly enhance the predictive accuracy of the optimality model. The factors we have in mind here include mutation, migration, genetic drift, and genetic and phylogenetic constraints. Obviously, an optimality model must still contain constraints in order for it to be well motivated. So, for example, in many sex ratio models the optimal sex ratio is selected from among competing phenotypes whose brood sizes are constrained to be equal. The retention of such constraints in the censored model reflects the fact that in all optimization analyses, some features of the biology are assumed to be "global" constraints (i.e., to be invariant across individuals), while others are assumed to be potentially subjected only to "local" constraints like genetic drift and pleiotropy and therefore capable of evolving to an optimal state. The important issues of how global and local constraints are meaningfully distinguished and of how the presence of constraints can be meaningfully reconciled with optimality are partially addressed later.

We stress that all claims of optimality imply that natural selection is a sufficient explanation of the population phenotype. The reverse is not true in that evolution driven *only* by natural selection can occur in situations in which there cannot be an optimal phenotype, by definition. One such instance is when natural selection acts on a temporally varying trait for which all individuals have the same mean and variance (see, e.g., Kimura 1954; Gillespie 1977). Another instance can be seen by considering a selective model of the evolution of, say, habitat preferences in which it is assumed that each individual has a fixed preference and all offspring of a given mother have her habitat preference. Suppose the model predicts an equilibrium with a mix of phenotypes. By assumption, natural selection is not impeded by any kind of local genetic constraint, genetic drift, migration, or mutation. However, it is not a sufficient explanation of individual phenotypes because some nonselective process explains the differences *among* individuals in the population. That natural selection here is a sufficient explanation of the *group* phenotype does not mean that it explains the trait of each individual. A claim of optimality cannot be defended in this instance. Consider, in contrast, a selective model in which it is assumed that mothers differ in the mix of fixed habitat preferences that each passes on to her set of offspring. The new model predicts the same mix of phenotypes at equilibrium as the old model did. However, this

mix can be manifested by *a* mother in that her set of offspring could display the correct proportions of habitat preferences. The trait would be locally optimal in a particular population if every mother assigned the fixed habitat preferences with the correct proportions (corresponding to the predicted population mix). We have defined optimality and adaptationism (discussed later) in terms of explaining individual phenotypes in a given population in order to reflect the traditional and, in our opinion, generally well motivated focus of almost all adaptationists.

Of course, additional considerations may contribute to a decision to accept a claim as to the sufficiency of natural selection even given statistical accuracy of the censored model's predictions. For example, Gillespie (1977) showed that a neutral model and a selective model of trait evolution have the same sampling distribution. Natural selection is a sufficient explanation of trait evolution in the selective model because constraints, mutation, migration, and genetic drift are excluded. Yet even exact congruence of the sampling distribution and an observed distribution would not support acceptance of the sufficiency of natural selection because the neutral model also fits the data. Similarly, an optimal explanation of a trait does not have precedence over nonoptimal or nonselective explanations. If competing explanations work equally well, additional data or analyses are needed.

We note that the claim that a trait is "adaptive" differs markedly from the claim that proposition (O) is true. It is one thing to believe that natural selection is involved in the evolution of, say, the vertebrate eye; it is something quite different to believe that the eye is optimal.

We now can turn to the question of what adaptationism is. An adaptationist view of a trait is given by (O). Adaptationists do not deny that factors other than natural selection played some role in evolution. However, they believe that these other influences may safely be ignored. This is a stronger claim than the more modest claims of propositions (U) and (I). Adaptationists are often prepared to concede that (O) and even (I) may be false with respect to molecular traits (see, e.g., Maynard Smith 1978). This makes it reasonable to formulate the thesis of adaptationism as

Natural selection is a sufficient explanation for most nonmolecular traits, and these traits are locally optimal.

This is a generalization of (O).

IS ADAPTATIONISM TESTABLE?

If adaptationism were the claim that all nonmolecular traits were locally optimal, a single counterexample would be enough to refute it. However, this strong form of adaptationism is something that few biologists would endorse. We believe that our formulation is closer to the real issue that currently exercises biologists.

A possible complaint about our formulation of adaptationism is that it is untestable. The thought behind this criticism might be that if one optimality model is refuted in the analysis of a trait, this outcome can be attributed to ignorance of the biology involved and another such explanation can be constructed; as a result, the assumption of optimality need not be questioned (see Lewontin 1978; May-

nard Smith 1978; Gould and Lewontin 1979; Mayr 1983; Sober 1988; Cain 1989; Mitchell and Valone 1990; and Parker and Maynard Smith 1990 for claims and counterclaims about this issue).

It is worth pondering the full implications of this familiar line of argument. If adaptationism is untestable, so is pluralism (the view that several forces have contributed significantly to evolutionary change). After all, nonoptimal or non-adaptive "just so" stories are as easy to construct as optimal ones. For those of us who hope that generalizations about how evolution usually works are not beyond the reach of science, this charge of untestability must be countered. We think adaptationism is testable by the accumulation of successes and failures of *specific* optimality models (see below) and that there is no general prior issue of testability to be resolved about such models.

The importance we attach to testing specific models reflects our belief that some of the most important issues about adaptationism have been obscured in the debate between those who believe that traits are usually not locally optimal and those who believe that they almost certainly must be. We agree and disagree with both sides of this debate. On the one hand, we agree that population-genetic data and theory reveal no general reason why optimality is to be expected, although this is a possibility for any given trait. On the other hand, we agree that natural selection is ubiquitous, but we do not believe that optimality should be assumed just because it could occur (see the discussion of optimality as a null hypothesis in Parker and Maynard Smith 1990). We can think of no prior reason why an optimality model cannot provide a correct explanation of a particular trait. One cannot refute the notion of optimality simply by affirming the general importance of understanding context, interaction, and the particularity of nature. Indeed, one could think of prior reasons why any particular evolutionary outcome should not occur. Consider, for example, why heterozygotes for the sickle-cell trait "should" not have an advantage in human populations in Africa. After all, heterozygote advantage is rare in a general sense. Yet, biologists generally agree that it is present in this instance.

Accordingly, we believe that the critical question to be asked about optimality models is not "Is the general approach correct?" but instead "How should specific models be tested?" Our answer to the latter question, if correct, is worth noting, since optimality models have rarely been examined in the way we will propose. In conjunction with our proposal, we will argue that the common practice of qualitatively testing the predictions of optimality models is a very weak scientific exercise at best and an approach that will not lead to a valid assessment of the thesis of adaptationism.

As noted earlier, our focus is on all models used to support claims that a trait is locally optimal. In the following section, we focus on one type of optimality model, the ESS model (see Maynard Smith 1982), because the important distinction between (I) and (O) is especially clear in this context.

EVOLUTIONARILY STABLE STRATEGIES AND STATES

Evolutionarily stable strategy models are an important subclass of optimality models that are often used to analyze traits whose "payoff" is frequency-

dependent. A phenotype is an ESS (relative to a variety of alternatives) if, when fixed (or nearly so), none of the other phenotypes can enter the population (see Maynard Smith 1982; Hines 1987, 1990 for further details).

An ESS can be of two types. A "pure" strategy is one in which a single phenotype is manifested by an individual. Such a strategy can be a conditional phenotype, in that, for example, an individual always manifests one trait when young and always another trait when old. A "mixed" strategy, in contrast, is one in which an individual can manifest any one of a set of phenotypes at a particular time, the choice being random. If A and B are phenotypes, then "Always produce A" is a pure strategy and "Produce A $x\%$ of the time and B $(100 - x)\%$ of the time" is a mixed strategy.

An evolutionarily stable state, on the other hand, describes a group "phenotype" without specifying the phenotypes of individuals. When a pure strategy is fixed, the terms *strategy* and *state* may be used interchangeably. However, when the equilibrium involves a mix of phenotypes, it is worth taking care in how the terms are applied. Suppose the evolutionarily stable state is $x\%$ A and $(100 - x)\%$ B. This state is compatible with many arrays of individual strategies. One possibility is that $x\%$ of the individuals always produce A and $(100 - x)\%$ of the individuals always produce B. Another possibility is that each individual produces A $x\%$ of the time and B $(100 - x)\%$ of the time. We will refer to the first as a polymorphism of pure strategies and to the latter as a monomorphic mixed strategy. Of these two possibilities, only the latter is an ESS.

Evolutionarily stable strategy models differ in their formal consistency with both arrangements (Thomas 1984). But given that a model is consistent with both, which configuration should we expect to find in a natural population to which the model is said to apply? The expectation is that an evolutionarily stable state should be realized as an ESS. The reason is that a selective benefit accrues to individuals possessing the ESS if the population is finite in size and thereby not at selective equilibrium (Hines 1980, 1982). This point has been made in the context of specific ESS models by Maynard Smith (1988) and Vickery (1988) (see also Verner 1965; Lloyd 1983; Poethke 1988).

Understanding the distinction between an ESS and an evolutionarily stable state is critical to understanding what the structure of a test of *any* optimality model should be if (O) is to be assessed. One cannot draw a conclusion about the optimality of individual phenotypes simply from the fact that the population exhibits a particular phenotype (e.g., an evolutionarily stable state). It is entirely possible that a population exhibits the evolutionarily stable state (natural selection thereby being a sufficient explanation of the population phenotype) and yet no individual manifests the optimal phenotype (i.e., the ESS). It follows that claims about the optimality of an observed phenotype advanced in the light of an optimality model must stem from comparisons of what individual organisms are doing (see below).

An evolutionarily stable state and an ESS differ significantly in their implications as to the power of natural selection. The presence of an evolutionarily stable state is evidence only that natural selection has played an important role in specifying the population phenotype and so leaves open the possibility that

within-population differences in phenotype are the result of a nonselective process. In contrast, the presence of an ESS is evidence that natural selection has produced a locally optimal phenotype. Accordingly, the distinction between state and strategy parallels the distinction between (I) and (O).

QUALITATIVE AND QUANTITATIVE TESTING

It is often maintained that optimality models provide "qualitative" predictions about data. Typically, in such a test a predicted trend is compared statistically with trends in the data. For example, given a prediction that a sex ratio should be female biased, one would assess theoretical success by determining whether, say, the average observed sex ratio was female biased. Although qualitative tests can be useful for some purposes (discussed later), there is a danger that investigators will adopt arbitrary and variable criteria for assessing the outcome of such tests. Indeed, in some instances, investigators have adopted contradictory criteria (see the examples depicted in fig. 1 of Orzack 1990). Perhaps the main defect in such assessments is that they usually are not based on a prior expectation as to what constitutes an unacceptable match between theory and data. The absence of prior criteria is especially problematic because it can lead to the too-easy acceptance of an optimality explanation for a particular trait. This imprecision will undermine the validity of the test of adaptationism (see below).

An additional problem with qualitative testing is that it often leads investigators to make visual assessments of the quantitative fit between predictions and data. Such assessments have tremendous potential for subjectivity. Witness the contradictory character of conclusions elicited by the sex ratio data of Werren (1980). Thornhill and Alcock (1983, p. 71) claim that the ESS "predictions are met with incredible precision." Similarly, after presenting a plot with some of the more deviant data missing, Trivers (1985, p. 285) states that the data "match closely the predicted values" and invites the reader to "notice the close fit of observed sex ratios to [the prediction]." In contrast, Leigh (1986, p. 207) claims, "There is a great deal of scatter about the quantitative prediction." None of these assessments can be regarded as meaningful in the absence of reasonable statistical criteria relating to the fit between predictions and data.

The presence of such criteria serves to distinguish between a qualitative test of a model (as defined earlier) and a quantitative test. In the latter, the investigator determines whether the model's numerical predictions are quantitatively accurate (i.e., whether they fit the numerical values in the data given standard assumptions about the nature and extent of sampling error). For example, in the case of the female-biased sex ratio just mentioned, one could determine whether the 95% confidence interval of the average of the data encompasses the prediction.

THE STRUCTURE OF A TEST FOR AN OPTIMALITY MODEL

Two questions can be asked about any causal model. One can ask whether the model describes some of the important processes underlying the data. Alternatively, one can ask whether the model describes all of the important processes

that need to be taken into account. In the context of our discussion of adaptationism, this distinction corresponds to the question of whether (I) or (O) correctly characterizes the relationship between data and the predictions of an optimality model. Consequently, talk about "testing a model" is ambiguous; it could mean that we are either testing (I) or testing (O). These considerations lead to a two-part test of an optimality model.

The first part is a comparison of observations and model predictions in which the quantitative accuracy of predictions is determined on the basis of standard statistical criteria. In addition to being the only objective way of determining whether a trait is locally optimal, a test of quantitative accuracy also allows one to determine whether models differ in the accuracy of their predictions.

An important clarification is needed with respect to the concept of quantitative accuracy. The level of analysis of an optimality model is the phenotype of an individual over its lifetime (or over some significant part thereof). Recall that the causal basis of such an optimality model is that a selective advantage accrues to the *individual* with a superior phenotype. This superiority implies that other phenotypes will be eliminated from the population (or nearly so) or prevented from entering it. To this extent, it is essential to assess quantitative accuracy with respect to a statistical description of *an* individual. This could be gained either by tracking individuals or by averaging the phenotypes of individuals derived from, say, an isofemale strain. The suitability of either approach obviously depends on the organism and trait in question.

The notion of optimality requires that one address the issue of individual variation. Indeed, the traits treated in many optimality models are of such palpable evolutionary significance that variation cannot be dismissed as mere bothersome noise. Perhaps within-population differences in, say, bristle number pose no great challenge to adaptationism; one can always say that such differences are selectively irrelevant. Proponents of the optimality approach, having spoken about the evolutionary importance of the traits they analyze (e.g., foraging behavior), cannot dismiss variation in this way. (This point has been acknowledged by some adaptationists; see, e.g., Cheverton et al. 1985.) The reason, of course, is that phenotypic variation could be due to evolutionary influences such as genetic constraints which prevent natural selection from creating the optimal phenotype. Without data on individual variation, the question of local optimality is left unaddressed.

There is no reason to think that optimality models cannot be quantitatively accurate. Lest the reader think we are "setting up" such models for attack in this regard, we note that many of their proponents share our view about the importance of quantitative accuracy as an indicator of theoretical success and about the potential that optimality models have for possessing this accuracy. There is no other explanation for numerous references in the literature to the quantitative agreement of data with the predictions, for example, of optimal foraging models (see, e.g., Stephens and Krebs 1986, table 9.1) or ESS models (see, e.g., Herre et al. 1987, p. 237). It is a separate matter that we happen to regard almost all such claims as unsubstantiated. (Our motivations for this statement are that quantitative testing has in fact often not occurred—instead a qualitative test

has been performed—and that when true quantitative testing has occurred, agreement between predictions and data was judged without a prior or clear definition as to what constitutes an unacceptable discrepancy.)

Of course, it is of concern that investigators might differ in the statistical tests they use and thereby come to contradictory conclusions about the fit between predictions and data. For example, consider the optimal sex ratios predicted for a range of foundress numbers by most models of local mate competition (see, e.g., Hamilton 1967). One could use a goodness-of-fit test in order to assess the fit between these sex ratios and observed sex ratios. This approach almost always must lead to the conclusion that such a model is quantitatively inaccurate if the sex ratios produced by isolated females are included in a test. In many species to which these models have been applied, such females almost always produce mixed-sex broods. The prediction is that such a female should produce an all-female brood. As a result, the expected number of males is zero and the χ^2 value for the goodness-of-fit test is infinite. On the other hand, one could assess quantitative accuracy of the model in a weaker sense by, for example, calculating the correlation between predicted and observed sex ratios and determining whether it matches or surpasses a given high value. Quantitative accuracy is a possible conclusion of such a test even when the predicted sex ratio for isolated females is included (see Orzack 1992 for further details and discussion). There is no general solution to this sort of problem. At present, at least, all we can ask is that investigators make explicit the tests they use so that subsequent investigators may apply the same test. Such standardization of protocols is essential to the construction of an unbiased test of adaptationism (see below).

The second part of the test of an optimality model is a statistical comparison of individuals (or isofemale strains, as noted earlier). In particular, one must determine whether between-individual heterogeneity in phenotype is consistent with the variation predicted by the model. If the model predicts one phenotype, one must determine whether repeated measurements of each individual imply statistical homogeneity of phenotype among individuals. This means that one must determine whether among-individual differences are repeatable (i.e., whether such differences are “real” or solely the result of measurement error). If the model predicts a set of phenotypes, one must determine whether an individual manifests (or can manifest) the predicted set of phenotypes. A necessary condition for the acceptance of (O) is that there be no significant differences among individuals in the fit of their phenotype(s) to the prediction(s). Our point is that whether the model predicts a single phenotype or many, claims of optimality must be based on multiple measurements of individuals (or isofemale strains) such that the within- and between-individual (or strain) components of phenotypic variation are characterized over the time period to which the optimality model is said to apply. In lieu of such measurements, even exact quantitative agreement between the prediction of an optimality model and a phenotypic average across individuals (or a group phenotype) does not support a claim of optimality (as in (O)). Such aggregation obscures the information needed to assess the relative importance of selective and nonselective forces that affect individuals. Acceptance of (O) requires that the relative importance of these forces be assessed.

		data on phenotypes of individuals		data only on population
		Homogeneity of fit to predictions	Heterogeneity of fit to predictions	phenotype or average phenotype of individuals
Relationship between predictions and data	quantitative disagreement, qualitative agreement	U and I	U or U and I	U or U and I
	quantitative and qualitative agreement	U, I, and O	U and I	U and I

FIG. 1.—The relationships among propositions (U), (I), and (O) and assessments of the fit of an optimality model (see text for further details).

With data limited to a phenotypic average or to a group phenotype, one can at best support a claim that (I) is true (see fig. 1). After all, an optimality model describes the phenotype of an *individual*, not the adaptation of a *group*.

HOW DOES ONE INTERPRET THE RESULTS OF A TEST OF AN OPTIMALITY MODEL?

Given a quantitative analysis of the fit of model predictions to the data and an analysis of the nature of individual variation, how do these results bear on accepting any of the three propositions concerning the power of natural selection?

If an optimality model proves to be quantitatively accurate with respect to its prediction of the population and individual phenotypes (e.g., the evolutionarily stable state and the ESS, respectively) and the individuals studied are homogeneous in the fit of their phenotypes to the prediction of the individual phenotype (e.g., the ESS), then it is appropriate to conclude that the trait is optimal; that is, (O) can be reasonably regarded as true (see fig. 1). If the model is quantitatively accurate with respect to its prediction of the population phenotype but not so with respect to individual phenotypes, (I) is reasonably regarded as true. Acceptance of (O) is not appropriate in this instance even though natural selection can be regarded as being a sufficient explanation of the group phenotype.

If an optimality model fails to be quantitatively accurate, the question remains as to whether it is qualitatively accurate. For example, suppose that the average sex ratios produced by every individual in some population are identically female biased although they depart significantly from the predicted female-biased value. The model correctly describes the direction in which selection is acting on the trait (because the average sex ratio is female biased). In this case, we would

conclude that the model is qualitatively accurate, though quantitatively inaccurate. Although (O) has been shown to be wrong by these findings, we believe that (I) would be a correct interpretation of the trait. In our opinion, (I) would be true even if the individuals were significantly heterogenous in the sex ratios they produce as long as the population sex ratio were significantly female biased (see fig. 1).

Another possible outcome is that the model proves to be both quantitatively and qualitatively inaccurate. The model fails not only to predict accurately the observed traits but also to capture the trends found in the data. This result would lead to the conclusion that both (I) and (O) are false. Of course, natural selection may have played some role in the evolution of the trait (i.e., (U) might still be true).

It follows from this picture of how an optimality model should be tested that whether qualitative testing is enough depends on the question one wishes to answer. If the question is whether natural selection has pushed the population in the direction of the optimum described in the model, then the issue of qualitative fit is a reasonable one to address. However, if one wishes to determine whether a trait is locally optimal, then qualitative fit is not enough (even if all individuals have an identical qualitative fit). Quantitative testing of quantitative predictions is a prerequisite to acceptance of (O). This underscores the important contrast between the idea that natural selection is important in shaping a trait and the idea that the trait is optimal. The latter thesis is more ambitious than the former and so demands more substantial evidence if it is to be accepted. (We note in this regard that all optimality models, being mathematically based, make quantitative predictions.)

The need for quantitative testing underscores the inadequacy in the context of testing optimality of the common attitude that a model captures the "essence" of a biological system if it successfully predicts a qualitative trend in the data. This attitude sometimes manifests itself in the metaphorical use of statistical language; the model is said to "account for most of the variation" or to be a "first-order" explanation, with the implication that unexplained variance is a "second-order" detail and, hence, biologically less important. We will not take issue with this attitude as a general rule of thumb in scientific inquiry. Our point, however, is that this attitude is incompatible with an adequate assessment of optimality. Qualitative testing cannot discriminate between (I) and (O).

It is implicit in the preceding discussion of how test results relate to acceptance of (U), (I), and (O) that the optimality of a given trait is a proposition to be proved. The idea that optimality models "do not test the proposition that animals are optimal" (Krebs et al. 1983, p. 204) glosses over the fact that claims of optimality require support in the context of explicit models if they are to be regarded as scientific claims in the standard sense. (The same, we emphasize, is true about claims of *non*optimality.) We suspect some biologists claim that models do not test optimality because they think that nature is complex and inscrutable, at least as judged relative to the tools we use to understand it. While many traits may ultimately prove optimal, we can see no justification at present for the general attitude that model "failure" always or almost always results from an investi-

gator's failure to properly appreciate the nature of the optimality. Assertions of optimality must be judged in light of explicit models. In this regard, it is essential to note that the test we propose "cuts both ways" with respect to confirming or rejecting the predictions of an optimality model. There is no reason to think that homogeneity of quantitative fit to predictions cannot be achieved, since it is certainly possible that natural selection and genetics should act in the way postulated by these models. As long as such homogeneity is based on adequate comparisons (see below), it supports the adaptationist assertion that natural selection has created an optimal phenotype. We stress that no general biological arguments exist that convincingly imply that natural selection cannot create an optimal phenotype. In particular, the claim that "populations are always variable" is ambiguous in the present context. One reason is that it is not obvious that genetic variation is present for most traits of interest to adaptationists (witness the apparent absence of such variability for sex ratio traits in many species) or that mutations have any functional effects on these traits. Therefore, it is not as though mutation must be invoked as an important force affecting any trait such that (O) cannot be true (as in the case of mutation-selection balance; see Haldane 1927). Perhaps more importantly, the evolutionary implication of variability in a character cannot be judged independently of a particular optimality model. Variability *by itself* does not imply rejection of (O). As noted previously, what needs to be assessed is whether the phenotypic variability of an individual is statistically consonant with the phenotypic variability predicted for an individual by the optimality model.

Local optimality is also not ruled out by a general claim that not all variants are possible because of genetic, phylogenetic, or biomechanical constraints. That such constraints surely act to prevent globally optimal phenotypes from evolving (e.g., the nonexistence of elephants that escape predators by flying) does not imply that such constraints must prevent locally optimal phenotypes from evolving. We think the distinction between global and local constraints is almost always clear in practice for two reasons. First, many optimality models are motivated locally (i.e., by a variable trait in a species or group of related species in which other traits are invariant). So, for example, when assessing the optimality of a sex ratio trait in a species of wasp, it is reasonable to regard, say, arrhenotoky or the physical constraints engendered by the need to fly as global constraints whose presence has little or no implication for whether a locally optimal sex ratio can evolve. Second, an optimality model must be tested against data on individuals or isofemale strains from the same population if (O) is to be properly assessed. This context necessarily restricts the variety of phenotypes considered in an optimality model. After all, when studying, say, the body size of individuals in a particular wasp species, the absence of warbler-sized wasps has no implication for whether the observed body size is locally optimal.

Three general points must be made about the test we propose. The first relates to the fact that phenotypic homogeneity has evolutionary significance. The problem is that the appearance of homogeneity can result from undersampling of individuals or strains. As a result, undersampling biases one toward accepting the optimality model. The solution is to structure investigations so that standard

criteria are used in studies. This means that investigators should study similar numbers of individuals or strains and have similar numbers of observations for each one so that the between- and within-individual components of variation are adequately characterized (at least in a relative sense). This type of standardization of protocols is needed if a test of adaptationism is to be possible (see below).

The second point is that the test is applicable to all types of optimality models including those that predict a pure phenotype. So, for example, an ESS model predicting a pure strategy is to be judged quantitatively inaccurate if a mixed strategy is observed. At best, qualitative agreement of this model with the data would support (I) as a description of the trait. Of course, another model might indicate that the trait is optimal (as in (O)).

The final point relates to a population's lineage. It is possible that a trait deemed optimal on the basis of the test described earlier evolved to that state in a different environment or in an ancestral species. We regard this sort of possibility to be just one of a number of the potentially important evolutionary "factors" to be investigated in the analysis of a specific trait (see below). To this extent, our test of adaptationism is complementary to (and does not compete with) the phylogenetic study of adaptation (see, e.g., Brooks and McLennan 1991), especially since we believe that the latter should rest on a determination of whether (U), (I), or (O) is the best description of a trait.

WHICH STUDIES ALLOW AN ASSESSMENT OF PROPOSITIONS (I) AND (O)?

Although the bearing of within- and between-individual variation and quantitative testing on the question of optimality may be obvious once stated, we can find only two sets of studies in evolutionary biology in which data and analyses are structured so as to allow one to distinguish between (I) and (O).

The first set of studies focuses on the reproductive behavior of the digger wasp, *Sphex ichneumoneus* (Brockmann and Dawkins 1979; Brockmann et al. 1979). A female can dig her own nest or enter the nest of another female. (A female typically makes 12 or so of these decisions in her 6-wk lifetime.) In either case, she provisions the nest with katydids and very likely will lay an egg in the nest unless ousted by an entering female. The authors developed an ESS model to predict the optimal mix of the digging and entering behaviors. For a population in New Hampshire, the model is quantitatively accurate in that the predicted distribution of behaviors closely matches the observed distribution (Brockmann and Dawkins 1979, p. 229), and one can show that they do not differ significantly ($\chi^2 = 0.608$, $df = 1$, $P > .05$). By comparing lifetime behavioral records of *individuals*, the authors also demonstrate that the evolutionarily stable state is realized as an ESS. In particular, individuals do not differ significantly in the mix of behaviors each produces (Brockmann and Dawkins 1979, p. 214). Hence, it is reasonable to conclude that individuals possess optimal nesting behavior in this population (as in (O)). In contrast, the authors conclude that their ESS model cannot be reconciled with the data from a population in Michigan (Brockmann et al. 1979, pp. 491–494).

The second set of studies focuses on the sex ratio behaviors of 12 isofemale

strains extracted from a population of the parasitic wasp, *Nasonia vitripennis* (Orzack 1990; Orzack and Parker 1990; Orzack et al. 1991). Three behaviors were examined: the sex ratio produced by a female when alone, the sex ratio produced by a female when she recognizes previous parasitization of a host, and the "conditional" sex ratios produced by a female when ovipositing in the presence of an array of different numbers of females. The repeatability of each behavior was assessed by measuring multiple females from each strain. The predictions of ESS models for these behaviors are not quantitatively accurate. For example, none of the strains has a Spearman rank correlation between observed and optimal conditional sex ratios with a 95% confidence interval that overlaps 0.95 (see table 8 of Orzack et al. 1991). (Perfect adherence would result in a correlation of 1.0. A value of 0.95 was used because the confidence interval of a less-than-perfect correlation cannot overlap 1.0.) For all three behaviors, the associated ESS model did predict the qualitative trend in the data for most (but not all) strains. There is significant between-strain heterogeneity of fit to predictions (see fig. 1 of Orzack and Parker 1990; table 1 and figs. 5 and 6 of Orzack 1990; table 8 and fig. 6 of Orzack et al. 1991). A reasonable interpretation of these results is that natural selection is an important force causing sex ratios to evolve in the direction predicted by the models; however, the stronger thesis of optimality is not supported by the data; that is, (I) but not (O) can reasonably be regarded as true.

WHY DO ALMOST ALL PRESENT TESTS OF OPTIMALITY MODELS FAIL
TO TEST OPTIMALITY?

In the previous section, we noted two sets of studies that allow assessment of (I) and (O). (It is, of course, a different matter as to whether the conclusions reached about optimality are correct.) These studies differ from others in the literature of evolutionary biology in that they include both quantitative analyses of the accuracy of model predictions and *comparisons* of the phenotypes of individuals or isofemale strains. As noted previously, quantitative accuracy of model predictions and demonstration of phenotypic homogeneity are essential to assessment of (O). We stress that the latter need not imply a lack of heterogeneity among individuals at one time. Recall, for example, that, at any one time in the New Hampshire population of *Sphex*, a given female may be digging a burrow while another female is entering a burrow. Nonetheless, the mix of these behaviors is statistically homogenous across the individuals studied.

We have looked at hundreds of other tests of optimality models. In particular, we examined studies relating to many kinds of behavioral, morphological, and physiological traits and our search included the journals the *American Naturalist*, *Animal Behaviour*, *Behavioral Ecology*, *Behavioral Ecology and Sociobiology*, *Ecology*, *Evolution*, and *Evolutionary Ecology*. In addition, we have looked at almost all of the tests cited in Maynard Smith (1982) and those cited in Krebs et al. (1983) and Stephens and Krebs (1986) for which these authors claim there is quantitative agreement with the model. Most of the studies we have looked at test only a qualitative prediction of the model against a population phenotype or a phenotypic average. As a result, each of these studies supports (I) at best. A

second group of studies includes *one* of the following analyses: a quantitative analysis of the accuracy of the predictions of the associated optimality model, or some comparison of the phenotypes of individuals and/or an assessment of phenotypic repeatability. Nonetheless, the absence of either one of these analyses implies that all of these studies lack the structures in their data or analyses that allow assessment of (I) and (O). We note that there seems to be no association between the lack of the appropriate data and analyses and the nature of the conclusions drawn in that the studies we have considered include those accepting optimality and those rejecting it.

We likely have overlooked or misjudged some studies that do allow one to distinguish between (I) and (O), and we hope readers will bring these studies to our attention. We make no absolute claim as to the number of studies in which the structure of the data and analyses allow one to choose between (I) and (O). Perhaps a few more such studies are present in the literature. What we are certain of is the great rarity of such studies relative to the total number of studies of optimality models.

WHAT DOES WITHIN-POPULATION HETEROGENEITY MEAN?

Within-population heterogeneity of fit to the prediction of an optimality model has several possible explanations. One is that the phenotype undergoing selection is some complex of traits, not the single trait treated in the model under test. In this case, the model is mistaken and needs to be replaced by some other optimality model. Another possibility is that a critical assumption underlying the optimality approach—that the genetics of the character does not “get in the way”—has been violated. The heterogeneity could be due to genetic phenomena such as epistasis, heterosis, or pleiotropy, which could prevent the evolution of the optimal phenotype. Alternatively, it may be that the genetic assumptions of the optimality model are correct but natural selection has acted to achieve only an evolutionarily stable state, for example. In the specific context of the sex ratio behaviors of *Nasonia vitripennis* discussed earlier, there are several reasons to believe that each individual should possess the ESS if genetically possible (Orzack et al. 1991). Hence, it is likely (but still unproven) that the observed heterogeneity of fit to ESS predictions is due to the fact that genetics “has gotten in the way.”

The interpretation of within-population heterogeneity relates to a critical issue concerning optimality. It is often claimed that disagreement between the predictions of an optimality model and the data stems from misunderstanding the constraints (see, e.g., Cheverton et al. 1985). The underlying notion is that the trait is optimal given the appropriate constraints. Such a conclusion may be reasonable if the revised model results in the “resolution” of between-individual heterogeneity—that is, if individuals are statistically homogenous with respect to the quantitative fit of their phenotypes to the predictions of the revised model. (Of course, there should be some reasonable independent basis for justifying the new choice of constraints.) However, the notion of optimality loses all meaning if between-individual heterogeneity of phenotypes is regarded as consistent with optimality

because one posits a different constraint for each individual. Optimality means that the trait is the best relative to traits possessed by other individuals. The implied context of this definition is a population defined in the modern microevolutionary sense (i.e., one in which individuals with different phenotypes interact with each another). To argue that every individual is optimal in a different way because of different constraints removes this essential causal basis of the definition of optimality.

THE TEST OF ADAPTATIONISM

We believe that the results obtained in the analyses of *Sphex ichneumoneus* and *Nasonia vitripennis* do not fully support the adaptationist claims about genetics and natural selection that underlie the predictions of optimality models. But the obvious weakness to any present claim about the validity of adaptationism relates to the ensemble nature of the test of this thesis. We simply need more studies in which the predictions of an optimality model are quantitatively compared with data on individual phenotypes and in which there is a statistical comparison of individuals or isofemale strains.

The process we envisage for testing adaptationism is the same that any science confronts when it tries to assess the correctness of an "ism." The validity of a research program—adaptationism in evolutionary biology, behaviorism in psychology, functionalism in cultural anthropology—cannot be assessed by examining the way a single specific model conforms to a single data set (Sober 1988, 1993; Mitchell and Valone 1990). Rather, the investigation of a trait like the reproductive behavior of *Sphex* is a project in which, as time goes on, some optimality models may be replaced by others and more limited data sets may be supplanted by more inclusive ones. Some reasonable optimality account of the trait might be found to be adequate. We can eventually assemble the results of investigations of many traits and assess the correctness of adaptationism. However, the issue cannot be addressed directly but only after separate and more focused investigations have been carried out. At present, as described earlier, of the two sets of appropriately structured studies, one results in the conclusion that the trait in question is optimal, while another leads to the conclusion that the traits are not optimal. (We ignore the results for the Michigan population of *Sphex* because the test of optimality for this population is not complete, there being no quantitative comparison of the predicted and observed distributions of reproductive behaviors. We also suspect the sex ratio traits in *Nasonia* are interrelated enough that they should be treated as one trait in the present context. See later discussion.) As the number of studies grows, we expect that the fraction of investigations in which optimality is demonstrated will come to differ from 0.5. Of course, there will be mistaken assessments of particular traits that may eventually be revised as better data and models become available. On the one hand, revised optimality models may often lead to the reassignment of some traits from the categories supporting (U) and (I) to the category supporting (O). On the other hand, questions can be raised about any optimality model (like any

evolutionary model) as to whether any given assumption about constraints or ancestral variation is correct and about whether parameters can be estimated precisely (R. Lewontin, personal communication). In many instances, such questions will probably lead to reassignment of traits from the category supporting (O) to the categories supporting (U) and (I). Nonetheless, there is no reason to think that misassessments will bias the proportion of studies favoring optimality at any one time. To this extent, our test of adaptationism does not depend on a final assessment of the optimality of any given trait. Nor is it dependent on the absence of investigator bias. We hope that readers see why it is appropriate to regard the optimality of a given trait not as a *premise* but as a possible *conclusion* resulting from a specific test. However, we suspect that some readers will be unconvinced. All we ask in particular instances is that investigators determine in the standard manner described earlier whether the presently available optimality model is a sufficient explanation of the data. Such assessments contribute to the test of adaptationism regardless of whether the investigator is convinced either that no optimality model could ever be a sufficient explanation of the trait or that such a model must exist.

The test of adaptationism we advocate need not engender an interminable debate. Forty or 50 appropriately structured studies might well provide a reasonable assessment of adaptationism. For example, if 45 of the 50 tests lead to the conclusion that the trait in question is locally optimal, in our opinion one could conclude that adaptationism is correct. Attainment of some agreed-on number of tests should be a goal of evolutionary biologists. (In these days of the Human Genome Project, it may be appropriate—and perhaps even more useful—to organize a far cheaper Adaptationism Project in order to coordinate quantitative studies of optimality models.) In any case, we hope it is clear why common approaches to data collection and analysis of optimality models do not permit an adequate assessment of adaptationism. There are three reasons to be hopeful that such an assessment is achievable. First, we suspect that many present tests of optimality models may be based on data that are appropriately structured. In particular, we suspect that multiple measurements of individuals or isofemale strains are often made but that this facet of the data has been ignored probably because the importance of assessing the nature of phenotypic heterogeneity in the context of testing an optimality model has been unclear. Second, a number of studies can underwrite a test of (O) because data on individual differences are known to be available, although the present analyses at best allow acceptance of (I) (see, e.g., Waddington and Holden 1979). The final reason is that there are a number of traits for which it is clear that presently available data could be combined with the predictions of an optimality model in order to allow assessment of (O). The test simply needs to be “assembled.” One example is the sex ratio in humans. As is well-known, the observed population sex ratio at the onset of reproduction fits the quantitative prediction of an optimality model, and comparison of families has not revealed genetic variation for this trait. In addition, birth sex ratios are significantly male biased, which indicates that the sex ratio may not be explainable by a competing nonoptimal explanation (i.e., selection for “honest meiosis”). Thus, at present acceptance of (O) seems reasonable.

For all three of these reasons, a test of adaptationism of the size mentioned earlier might even be attainable in the next 10 yr or so.

PRESENTLY UNRESOLVED PROBLEMS WITH THE TEST OF ADAPTATIONISM

Like any exercise in comparative biology, an ensemble test of adaptationism faces some ambiguity as to how instances of biological phenomena are defined. Some of this ambiguity relates to the study of different traits in the same population. For example, we counted the three sex ratio traits of *Nasonia* as one when tallying the studies that contribute to the test of adaptationism. We did so on the basis of data that indicate that first and second sex ratios are significantly correlated (Orzack and Gladstone 1994). To this extent, our counting protocol was conservative. In general, at the present time we have no magical procedure for counting traits other than to suggest that investigators be conservative in the way we have been (except if statistical analyses reveal independence of the traits). Even this approach is problematic in that it is unclear how to assess an investigation of, say, two correlated traits in which one is classified as optimal and one is not.

There is also ambiguity in regard to assessing analyses of the same trait in different populations (or species). Are we to regard such studies as independent contributions to the test of adaptationism? To some extent, the answer depends on how "densely" the traits are sampled. After all, assessments of the optimality of almost any trait in a *Drosophila* species and in, say, a vertebrate can be regarded as independent contributions to the test of adaptationism. In general, the problem of determining independence would appear to be tractable given present techniques (e.g., molecular genetic analyses to determine the extent of gene flow within a species and phylogenetic analyses to determine whether, say, a species could have inherited an optimal trait from an ancestor).

We note that none of these problems appears to be insurmountable, and the degree of their actual effects on the test of adaptationism will be unassessable until more optimality models are tested in a way that allows discrimination between (I) and (O).

THE SIGNIFICANCE OF OPTIMALITY MODELS

Our goal in writing about how optimality models should be tested and about inappropriate testing of these models is not to attack adaptationism. Instead, we wish to promote a valid assessment of this research program. We are motivated by a strong belief in the importance of optimality models for understanding evolution. They are too important to be applied in an inexact way and to have unsubstantiated conclusions drawn about their explanatory power (see also Austad 1984).

In this regard we note that optimality models are important even if they do not explain data. A fundamental contribution of such models is that they describe what organisms *should* do in particular instances (see also Stearns and Schmid-Hempel 1987). Viewed in this way, optimality models of foraging behaviors, for example, clearly are extremely successful. They have led to a much clearer theo-

retical understanding of natural selection in populations experiencing spatial and temporal variability of resources. This success is independent of whether tests of these models lead to the conclusion that optimal foraging behaviors have evolved.

Nonetheless, it is important in evolutionary biology to work toward the integration of theory and data. Perhaps it is the desire to do so that has often led proponents of optimality models to regard any correspondence between theory and data as evidence for the optimality of individual phenotypes. Optimality models deserve better. At stake is our ability to assess the truth of adaptationism.

ACKNOWLEDGMENTS

We thank S. Abrams, S. Arnold, J. Beatty, J. Brown, J. Bull, J. Crow, C. Denniston, S. Frank, J. Kingsolver, M. Kreitman, E. Leigh, R. Lewontin, J. Maynard Smith, D. Parker, M. Rausher, and S. Skinner for advice, critical comments, or discussions and an anonymous reviewer for mentioning flying elephants. S.O. thanks B. Charlesworth, M. Kreitman, and the Department of Ecology and Evolution at the University of Chicago for support.

LITERATURE CITED

- Austad, S. N. 1984. A classification of alternative reproductive behaviors and methods for field-testing ESS models. *American Zoologist* 24:309–319.
- Brockmann, H. J., and R. Dawkins. 1979. Joint nesting in a digger wasp as an evolutionarily stable preadaptation to social life. *Behaviour* 71:203–245.
- Brockmann, H. J., A. Grafen, and R. Dawkins. 1979. Evolutionarily stable nesting strategy in a digger wasp. *Journal of Theoretical Biology* 77:473–496.
- Brooks, D. R., and D. A. McLennan. 1991. *Phylogeny, ecology, and behavior: a research program in comparative biology*. University of Chicago Press, Chicago.
- Cain, A. J. 1989. The perfection of animals. *Biological Journal of the Linnean Society* 36:3–29.
- Cheverton, J., A. Kacelnik, and J. R. Krebs. 1985. Optimal foraging: constraints and currencies. Pages 109–126 in B. Hölldobler and M. Lindauer, eds. *Experimental behavioral ecology and sociobiology*. Sinauer, Sunderland, Mass.
- Crow, J., and M. Kimura. 1970. *An introduction to population genetics theory*. Harper & Row, New York.
- Eshel, I. 1982. Evolutionary stable strategies and viability selection in Mendelian populations. *Theoretical Population Biology* 22:204–217.
- Gillespie, J. H. 1977. Sampling theory for alleles in a random environment. *Nature (London)* 266:443–445.
- Gould, S. J., and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London B, Biological Sciences* 205:581–598.
- Haldane, J. B. S. 1927. A mathematical theory of natural and artificial selection. V. Selection and mutation. *Proceedings of the Cambridge Philosophical Society* 23:838–844.
- Hamilton, W. D. 1967. Extraordinary sex ratios. *Science (Washington, D.C.)* 156:477–488.
- Herre, E. A., E. G. Leigh, Jr., and E. A. Fisher. 1987. Sex allocation in animals. Pages 219–244 in S. C. Stearns, ed. *The evolution of sex and its consequences*. Birkhäuser, Basel.
- Hines, W. G. S. 1980. Strategy stability in complex populations. *Journal of Applied Probability* 17:600–610.
- . 1982. Mutations, perturbations and evolutionarily stable strategies. *Journal of Applied Probability* 19:204–209.
- . 1987. Evolutionarily stable strategies. *Theoretical Population Biology* 31:195–272.
- . 1990. A discussion of evolutionarily stable strategies. Pages 229–267 in S. Lessard, ed.

- Mathematical and statistical developments in evolutionary theory. Kluwer, Norwell, Mass.
- Kimura, M. 1954. Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. *Genetics* 39:280–293.
- Krebs, J. R., D. W. Stephens, and W. J. Sutherland. 1983. Perspectives in optimal foraging. Pages 165–216 in G. A. Clark, Jr., ed. *Perspectives in ornithology: essays presented for the centennial of the American Ornithological Union*. Cambridge University Press, Cambridge.
- Leigh, E. G., Jr. 1986. Ronald Fisher and the development of evolutionary theory. I. The role of selection. *Oxford Surveys in Evolutionary Biology* 3:187–223.
- Lewontin, R. C. 1978. Adaptation. *Scientific American* 239:156–169.
- Lloyd, D. G. 1983. Evolutionarily stable sex ratios and sex allocations. *Journal of Theoretical Biology* 105:525–539.
- Maynard Smith, J. 1978. Optimization theory in evolution. *Annual Review of Ecology and Systematics* 9:31–56.
- . 1982. *Evolution and the theory of games*. Cambridge University Press, Cambridge.
- . 1988. Can a mixed strategy be stable in a finite population? *Journal of Theoretical Biology* 130:247–251.
- Mayr, E. 1983. How to carry out the adaptationist program? *American Naturalist* 121:324–334.
- Mitchell, W. A., and T. J. Valone. 1990. The optimization research program: studying adaptations by their function. *Quarterly Review of Biology* 65:43–52.
- Orzack, S. H. 1990. The comparative biology of second sex ratio evolution within a natural population of a parasitic wasp, *Nasonia vitripennis*. *Genetics* 124:385–396.
- . 1992. Sex ratio evolution in parasitic wasps. Pages 477–511 in D. Wrensch and M. Ebbert, eds. *Evolution and diversity of sex ratio in insects and mites*. Chapman & Hall, New York.
- Orzack, S. H., and J. Gladstone. 1994. Quantitative genetics of sex ratio traits in the parasitic wasp, *Nasonia vitripennis*. *Genetics* (in press).
- Orzack, S. H., and E. D. Parker, Jr. 1990. Genetic variation for sex ratio traits within a natural population of a parasitic wasp, *Nasonia vitripennis*. *Genetics* 124:373–384.
- Orzack, S. H., E. D. Parker, Jr., and J. Gladstone. 1991. The comparative biology of genetic variation for conditional sex ratio adjustment in a parasitic wasp, *Nasonia vitripennis*. *Genetics* 127:583–599.
- Parker, G. A., and J. Maynard Smith. 1990. Optimality theory in evolutionary biology. *Nature* (London) 348:27–33.
- Poethke, H. J. 1988. Sex ratio polymorphism: the impact of mutation and drift on evolution. *Acta Biotheoretica* 37:121–147.
- Sober, E. 1988. What is adaptationism? Pages 105–118 in J. Dupre, ed. *The latest on the best*. MIT Press, Cambridge, Mass.
- . 1993. *Philosophy of biology*. Westview, Boulder, Colo.
- Stearns, S. C., and P. Schmid-Hempel. 1987. Evolutionary insights should not be wasted. *Oikos* 49:118–125.
- Stephens, D. W., and J. R. Krebs. 1986. *Foraging theory*. Princeton University Press, Princeton, N.J.
- Thomas, B. 1984. Evolutionary stability: states and strategies. *Theoretical Population Biology* 26:49–67.
- Thornhill, R., and J. Alcock. 1983. *The evolution of insect mating behavior*. Harvard University Press, Cambridge, Mass.
- Trivers, R. 1985. *Social evolution*. Benjamin/Cummings, Menlo Park, Calif.
- Verner, J., 1965. Selection for sex ratio. *American Naturalist* 99:419–421.
- Vickery, W. L. 1988. Reply to Maynard Smith. *Journal of Theoretical Biology* 132:375–378.
- Waddington, K. D., and L. R. Holden. 1979. Optimal foraging: on flower selection by bees. *American Naturalist* 114:179–196.
- Werren, J. H. 1980. Sex ratio adaptations to local mate competition in a parasitic wasp. *Science* (Washington, D.C.) 208:1157–1159.