



The Causes of Natural Selection

Michael J. Wade; Susan Kalisz

Evolution, Vol. 44, No. 8 (Dec., 1990), 1947-1955.

Stable URL:

<http://links.jstor.org/sici?sici=0014-3820%28199012%2944%3A8%3C1947%3ATCONS%3E2.0.CO%3B2-%23>

Evolution is currently published by Society for the Study of Evolution.

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/ssevol.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.

For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

THE CAUSES OF NATURAL SELECTION

MICHAEL J. WADE

*Department of Ecology and Evolution, The University of Chicago,
915 East 57th Street, Chicago, IL 60637 USA*

SUSAN KALISZ

*Kellogg Biological Station, Michigan State University, 3700 Gull Lake Drive,
Hickory Corners, MI 49060 USA and*

*Department of Botany and Plant Pathology, Department of Zoology,
Michigan State University, East Lansing, MI 48824 USA*

Abstract.—We discuss the necessary and sufficient conditions for identifying the cause of natural selection on a phenotypic trait. We reexamine the observational methods recently proposed for measuring selection in natural populations and illustrate why the multivariate analysis of selection is insufficient for identifying the causal agents of selection. We discuss how the observational approach of multivariate selection analysis can be complemented by experimental manipulations of the phenotypic distribution and the environment to identify not only *how* selection is operating on the phenotypic distribution but also *why* it operates in the observed manner. A significant point of departure of our work from recent discussions is in regard to the role of the environment in the study of natural selection. Instead of viewing the environment as a source of unwanted variation that obscures the relationship between phenotype and fitness, we view fitness as arising from the interaction of the phenotype with the environment. The biotic and abiotic environment is the context that gives rise to the relationship between phenotype and fitness (selection). The analysis of the causes of selection is in essence a problem in ecology. The experimental study of the association between selection gradients and environmental characteristics is necessary to identify the agents of natural selection. We recommend research methods for identifying the agency of selection that depend upon a reciprocity between the observational approach of multivariate selection analysis and the manipulative approach of field experiments in evolutionary ecology.

Received February 28, 1989. Accepted December 22, 1989.

Since the seminal paper by Lande and Arnold (1983), the study of selection in natural populations has been placed in a quantitative framework concordant with evolutionary theory. Recent books by Manley (1985) and Endler (1986) are further evidence of the growing interest in the application of quantitative methods to the study of natural selection. A large number of recent papers (including Arnold, 1986; Koenig and Albano, 1987; Downhower et al., 1987; Hubbell and Johnson, 1987; Crespi and Bookstein, 1989; Schluter, 1988; Mitchell-Olds and Shaw, 1987; Wade, 1987; Wade and Kalisz, 1989) have addressed the statistical advantages and/or disadvantages of the various quantitative approaches to measuring selection. This emphasis on the theoretical methods gives the impression that the quality of an analysis of selection is equated with sophistication in multivariate statistics. It has overshadowed the fact that the analysis of selection is essentially an ecological and not a statistical question. Our purpose is to reexamine recent methods

for measuring selection (Lande and Arnold, 1983; Arnold and Wade, 1984*a*, 1984*b*; Mitchell-Olds and Shaw, 1987; Schluter, 1988) particularly with regard to identifying the conditions necessary and sufficient for inferring the agent of selection acting on a trait.

The most useful measures of natural selection are those that permit field data to be brought to bear incisively on questions in formal ecological and evolutionary theory. The Lande and Arnold (1983) multivariate analysis of selection permits field and laboratory data to be summarized in a way directly applicable to these formal theories. We favor it for this reason although we acknowledge the merit of other approaches for other purposes (see review in Endler, 1986). This method of analysis has been applied to two different kinds of questions: (1) *How* is selection operating to change the phenotypic distribution of traits in nature? and, (2) (Within the ecological context), *why does* selection operate in this way? These two questions are different and a Lande-Arnold

analysis, essentially a study of correlations, is *not sufficient* to provide an answer to the second question. The question of causation or the identification of the agents of selection requires a detailed knowledge of ecology. Since most evolutionary ecologists are as interested in understanding "why" as in describing "how" selection operates, in this paper, we suggest a method for answering this second question. Ecology plays a central role in any evolutionary analysis that seeks to address the question of the agent of natural selection: "... the *reasons* and *mechanisms* for natural selection ... can be obtained only by a detailed knowledge of the ecology and biology of organisms" (Endler, 1986 p. 164).

The procedures we recommend are an empirical complement to the multivariate observational approach to measuring selection. The multivariate approach applied to observational data is not a sufficient method for identifying the cause of selection on a particular trait even when refined for theoretical or statistical purposes in the manner suggested by Lande and Arnold (1983) or Mitchell-Olds and Shaw (1987). The multivariate approach is a correlational rather than a causal analysis. The addition of experimental manipulations to test hypotheses generated by a multivariate analysis is *necessary* to identify agents of selection in the environment and their interaction with the phenotype. (Even well-considered experimental studies may not be *sufficient* in some difficult cases to identify the agents or causes of selection.) In this view, the environment is seen as the selective agent and fitness arises from the interaction of the phenotype with its environment (Levins, 1968).

*Selection: The Covariance between
Fitness and Phenotype*

Changes in the mean, variance, and other moments of the phenotypic distribution are the usual criteria for characterizing the action of evolutionary forces on continuously distributed phenotypic traits. Changes in the phenotypic distribution occur *within* generations owing to the effects of migration, development and natural selection. Changes also occur with the transmission of phenotypes across generations, from parent to offspring, owing to the genetic phenomena

of segregation, recombination, chromosomal rearrangement, and genetic drift. In this paper, we will ignore the effects of migration, and development and identify within-generation change as *selection* and between-generation change as *response to selection* (Johannsen, 1909; Fisher, 1930; Lande and Arnold, 1983; Arnold and Wade, 1984a, 1984b).

When we measure phenotypic selection in natural populations, we compare the phenotypic distribution at one time with the phenotypic distribution at some later time(s) *within the same generation*. It is from this comparison that we draw our inferences regarding the operation of viability and fecundity selection. We deliberately avoid examining changes across generations because, when the process of *transmission* changes the phenotypic distribution, we would confound selection and transmission. We consider the clear separation of these two processes to be conceptually and practically important to the study of selection although we acknowledge that other authors prefer to include inheritance as part of the definition of natural selection (e.g., Endler, 1986).

Discussions in evolutionary ecology have been concerned primarily with changes in the location or mean, \hat{z} , of the phenotypic distribution. Many more discussions have been concerned with simply identifying the direction or sign of the change. The change in the location, \hat{z} , is equal to $\text{Cov}(z, w[z])$, the covariance of z and relative fitness, $w(z)$, taken over the phenotypic distribution, $p(z)$ (Robertson, 1966). The *selection gradient*, β , the partial regression coefficient of relative fitness on character value, has been recommended as a better descriptor of the force of direct selection on a particular trait (cf. Lande and Arnold, 1983 p. 1220). The $\text{Cov}(z, w[z])$ or the selection gradient, β , describes how selection changes the mean of the phenotypic distribution, giving both the direction as well as the magnitude of the change. However, it does not provide any information regarding why an association between fitness, $w(z)$ and character value exists; covariation is not causation. As Endler (1986 p. 164) remarks, "There are few cases in which it is known why natural selection occurs." Several different features of the phenotypic distribution prohibit a direct

interpretation of the causal relationship between character value and fitness. We enumerate several of these features and give examples in the Appendix.

*Observation and Experimentation for
Causal Analysis*

When we ask, Why does selection change the phenotypic distribution? we are asking, What is the agent of selection in nature? or *For what environmental reason does an association between fitness and phenotype exist?* The fitness of an individual is the result of the interaction of the phenotype with the environment and not an intrinsic feature of either one. For example, melanic forms of the moth, *Phigalia pedaria*, are not intrinsically more or less fit than nonmelanic forms; the fitness of each morph depends upon the local environmental conditions (Lees, 1971). In our view, the environment provides the "context" in which there can be a causal relationship between phenotype and fitness. If the environmental context changes, then we expect the relationship between phenotype and fitness to change as well. For this reason, the contribution of a particular trait to overall fitness will often depend jointly upon the value of the phenotype and the value of a local environmental variable. It is this environmental variable that is the selective agent. Thus, identifying those environmental contexts in which there is a relationship between phenotype and fitness is an important first step to identifying the causal basis of fitness.

To understand the relationship of phenotype to fitness we first require knowledge of the life history and ecology to choose those phenotypic traits that may be associated with fitness. Second we estimate the shape of the fitness function by observing changes in the distribution of these traits within a generation in the field (Lande and Arnold, 1983). The selection gradient, β , is one of the shape characteristics of the fitness function, and we wish to understand why it takes the shape that it does. *We expect our estimates of the shape of the fitness function to change spatially and temporally concordant with changes in the agent of selection.* It is this variation in our estimates of the shape of the fitness function that provides the most valuable information for identi-

fying the agent of selection and its quantitative effect on the relationship of the phenotype to fitness. Without such variation in the estimates of β , we cannot find an associated environmental variable that *accounts* (in the statistical sense) for the variation in selection.

Kalish (1986) studied the microspatial variation in β , for timing of germination in the blue-eyed Mary, *Collinsia verna*. She found significant variation in both the sign and magnitude of β within a 10 hectare population. This study is an example that variation in β does exist, and it indicates the scale on which we might look for variation in selective agents to account for the observed variation in β , at least in the statistical sense. Kalish (1986) did not have corresponding microspatial measurements of an environmental variable(s), E. Without a measure of some environmental variable corresponding to each estimate of β , we can only *describe* the local variation in selection. We cannot attempt to account for it by looking for associations at the microspatial level between β and E. (Note that both β and E are characteristics of [microspatial] populations, not individuals.) Differently put, to account for variation in the action of natural selection in terms of variation in the environment, we are examining the covariance between β and an environmental variable, E. This covariance, $\text{Cov}(\beta, E)$, is necessarily zero if there is no variation in β or no variation in E, or if there is no association between β and E. We have identified variation in β as the problem to explain when we ask, Why should there be an association between phenotype and fitness? Because β is a *measure of the association between phenotype and fitness*, variations in β can represent variations in the action of selection.

When $\text{Cov}(\beta, E) \neq 0$, we can hypothesize that the environmental character, E, is the causal agent of selection. Several reasons other than selection, however, could account for this association, including other environmental factors associated with E. To establish a causal relationship between the environment and selection, we need to investigate this hypothesis using experimental manipulations.

Let us illustrate our approach with a simple example. Consider the observation of

selection for increased body size as estimated by the positive selection gradient, $+\beta$, in some natural population. Let the adaptive hypothesis be that there is a fitness advantage to large individuals owing to the relationship between body size and risk of predation. This hypothesis predicts that, in the absence of the predator, our estimate of β is expected to be zero. We could test this prediction by empirically altering the relevant environmental variable, namely, the presence of the predator. Our enclosure experiment would consist of two replicated treatments, one excluding predators and one not, with the same initial distributions of prey body sizes, with replication in each treatment. We would obtain four estimates of β , one in each replicate of each treatment. The experimental estimates of β are expected to be near zero in those treatments from which predators were excluded but significantly positive in those treatments with predators. Breden and Wade (1989) performed such an enclosure experiment in their investigation of the relationship between group size and group fitness in willow leaf beetles. They found that when predators were excluded there was no relationship between group size and larval survivorship (i.e., β was equal to zero in the absence of predators). However, when predation was permitted, a significant positive relationship between group size and fitness was observed.

Breden and Wade (1989) performed several different experimental studies of group size and its relationship to fitness, but not all of the experiments directly address the identification of the causal agent of selection. For example, when they experimentally created groups of different sizes and observed their fitnesses, they were not addressing the *cause* of the relationship between group size and fitness. This experiment demonstrated that group size, rather than some unknown associated phenotypic variable such as initial weight or health, was related to fitness. These results combined with the observations of group defensive displays (Wade and Breden, 1986) strongly suggested the hypothesis that group living was an adaptation for reducing predation. It was the *change* in the relationship between group size and fitness with the exclu-

sion of predators (environmental manipulation) that provided the direct empirical evidence confirming this hypothesis of the agency of selection. The experimental manipulation of the phenotype distribution (here group size) was important to establishing that the phenotype of group size was under selection rather than some other uncontrolled or unmeasured phenotype correlated with group size. However, the entire experiment had to be repeated in a number of different environments (i.e., with and without predators) to demonstrate that predation was the *cause* of the relationship between group size and fitness.

Not all studies of the agency of selection will be this straightforward. In particular, there is a growing interest in frequency dependent selection especially in behavioral ecology. In this case, it is some feature of the phenotypic distribution itself that is the causal agent of selection. The experimental approach here involves two aspects. First, one must manipulate the phenotypic distribution of one trait in relation to other traits to demonstrate that it, rather than some unmeasured correlated trait, is the focus of selection. This experimental manipulation of the phenotype must be conducted so that the frequency dependent context remains unaltered. Once it is established that there is selection on the trait, then one would manipulate the environmental context to investigate the agency of selection. In the case of frequency dependent selection, we hypothesize that the shape characteristics of the phenotypic distribution itself, such as the variance or the skewness, are the causes of the relationship between trait and fitness. To proceed further, we must be able to experimentally establish populations with different variances (or different skewnesses) and then measure the fitness of individuals within each of these contexts. Clearly, the empirical task is greatly complicated. Unlike the case of predation discussed above where we introduced an entire distribution of phenotypes into the two contexts, here we can introduce only one or at most a few individuals into any experimental context. When the context is itself a phenotypic distribution with, say, a controlled variance, adding an entire population would change the environmental treatment, i.e., the variance.

Hence, *each* context must be replicated to examine the fitnesses of individuals and obtain estimates of β and there must furthermore be *several* contexts to investigate the change in β with context. For these reasons, the empirical study of the causal basis of frequency dependent selection will always be difficult.

For example, consider a situation in which variation in the date of emergence causes a hierarchy of plant sizes that subsequently affects the degree of self-thinning, frequency dependent viability selection. If we control emergence date, either by studying cohorts with the same date or statistically by covariance analysis, we then remove the important effect of size variation in the selective environment because uniform emergence times do not give rise (in this example) to variations in plant size. Whatever the cause of variation in plant size, genetic, environmental, or historical, it is this feature of the phenotypic distribution, the variance in size, that is the *agency* of viability selection through competition and self-thinning. This example underscores the context dependent nature of the measurement of natural selection and the need for a careful choice of ecologically important characters, fitness components, and field experiments. Some other examples include genotype-frequency dependent fitnesses (e.g., Lewontin, 1955; Wade, 1985), age-structure dependent fitnesses (e.g., McCauley, 1978; Mertz et al., 1948), and the effects of the variance (McCauley, 1979) and the mean (Lande, 1976) of continuously distributed characters. These effects of context make the study of selection more difficult in practice if not in principle but they also provide some of the most interesting examples of agency.

The Complex Role of Environment

“... many of the most intriguing questions about selection in nature concern situations in which selection operates differently within different parts of a population” (Heisler and Damuth, 1987 p. 582). Consider the case in which variation in the environment results in phenotypic variation and that phenotypic variation subsequently results in selection. For example, Mitchell-Olds and Shaw (1987 p. 1154) hypothesize a situation in which local variation in patch

quality gives rise to local variation in plant size. Plants grow large in high quality patches, but plants in low quality patches are smaller. In this scenario, competition between individuals is more intense in the high quality patches and leads to mortality as a result of self-thinning (Fig. 1, upper). The conclusion that large size is negatively correlated with viability (i.e., large plants are selected against) cannot be directly interpreted as selection for small size because the variation in fitness is environment-specific (i.e., between-patches). Mitchell-Olds and Shaw (1987 p. 1155) recommend statistical methods to remove the variations in environmental quality or experimental methods to randomize them across individuals to achieve a more accurate depiction of selection.

On average there *is* selection against large plant size because there is a negative statistical relationship between size and mortality in this population as a whole (Fig. 1, lower). However, it is difficult to identify the cause of selection, namely, density-dependent competition or self-thinning and/or environmental tolerance. In Figure 1, upper, environmental quality and self-thinning through competition are confounded. To separate these agents, we would use the experimental design illustrated in Figure 2. This design is motivated by the observed fitness function (Fig. 1, lower), the bimodal phenotypic distribution of plant sizes, and the observed local variation in self-thinning. The proposed experiment (Fig. 2) is designed to examine the separate effects of density and resource quality (and with replication within cells, their potential interaction as well). Differences observed among columns in the statistical relationship between fitness and plant size indicate that density is the agent of selection. Differences observed among rows indicate that environmental quality is the agent of selection. The experiment provides knowledge of β as a function of initial density (N) and environmental quality (E). We can apply this fitness function to any plant population in nature if we know the size distribution and have a priori knowledge of N and E.

In this example, there are several different ways in which the environment might be “controlled” either statistically or experi-

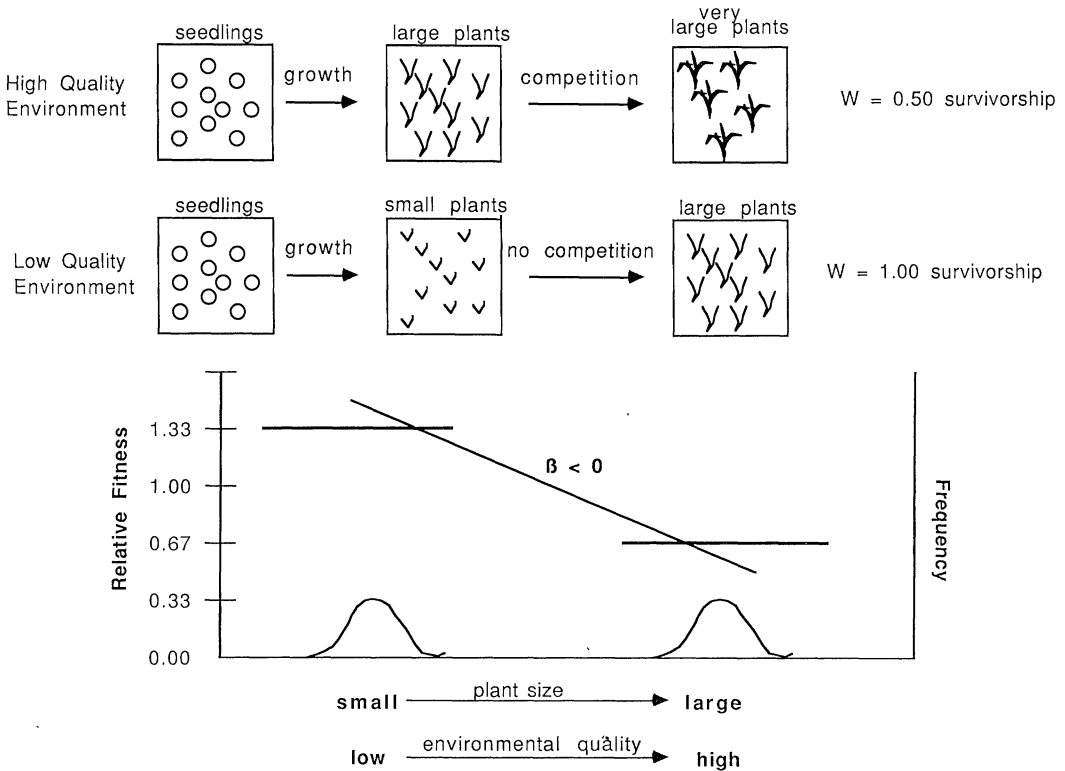


FIG. 1. (upper) A schematic illustration of a situation in which environmental variation induces phenotypic variation and that phenotypic variation subsequently results in selection. Note that initially there are the same numbers of seedlings in each environment. The symbol W represents absolute viability fitness in each environment. (See text for further discussion.) FIG. 1. (lower) Fitness as a function of the plant size distribution above. The slope of the regression of relative fitness (the ratio of absolute to average fitness) on plant size is negative indicating selection against large size. Note the bimodality in the plant size frequency distribution. (See text for further discussion.)

mentally. All are expected to influence the measurement of selection but not all of them can distinguish the environment as context from environmental variation as an independent cause of phenotypic and fitness variation. We emphasize this example to influence the planning of experimental manipulations.

DISCUSSION

When selection is defined as the covariance between phenotype and relative fitness, it is important to recognize that relative fitness itself is a context dependent measure because its denominator, mean fitness, is fitness averaged over the phenotypic distribution. The mean fitness of a manipulated study population can be different from that of the natural population; consequently, the effects of selection on the phenotypic

distribution of the study population will be different from the effects of the same agent of selection on the naturally occurring distribution. We suggest using this difference to our experimental advantage in estimating selection but we emphasize the effects in regard to inferring agency.

In a recent paper, Mitchell-Olds and Shaw (1987) discuss the statistical problems with the multivariate approach as well as problems of biological inference. They are concerned with those problems of unmeasured characters and extrinsic environmental factors that complicate or prohibit the unbiased estimation of selection gradients operating on focal characters and the establishment of confidence intervals on those estimates. They explicitly discuss the environment and its effects on fitness, either directly or through phenotype \times environ-

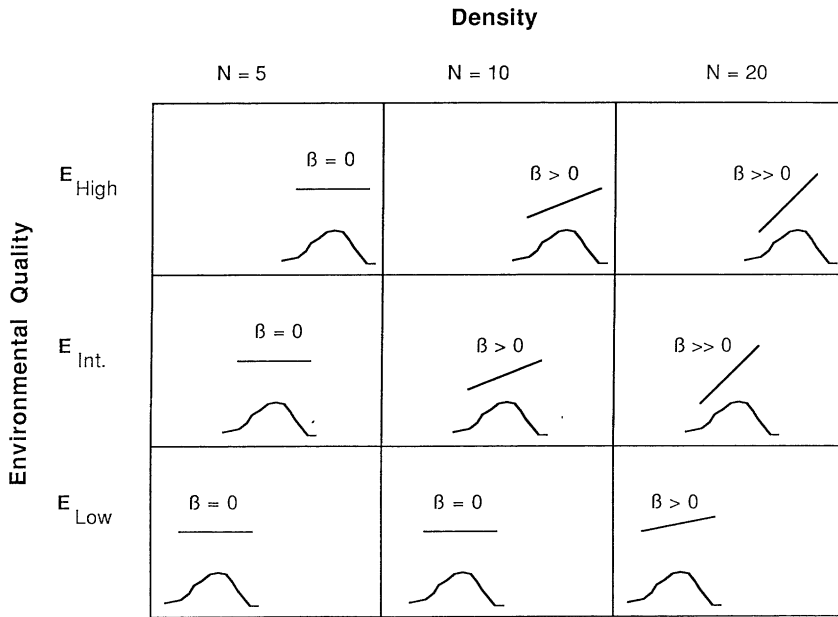


FIG. 2. Experimental design for investigating the separate effects of density (N) and environmental quality (E) as agents of selection. Within cells, the bell shaped curves represent the resulting plant size distributions and the straight lines represent the regressions of relative viability fitness on plant size *within each N-E treatment*; i.e., the hypothetical average relationships between fitness and plant size phenotype for each combination of density and environment. (For pedagogical reasons, we portray only the average fitness function per cell, rather than each replicate separately. Replication within cells is necessary for detecting an interaction between density and environmental quality.)

ment interactions. They also discuss the way in which such environmental effects could interfere with the accurate detection and estimation of the effects of phenotype on fitness.

We are concerned with the research program *after* this estimate of the fitness function is obtained. We suggest experimental manipulations of the environment to identify the agents of selection in terms of associations between β and E. We too are interested in the best unbiased estimate of the selection gradient and recognize the importance of obtaining confidence limits on these estimates for comparative purposes. However, we treat even the best estimate of the selection gradient through multivariate analysis as a "hypothesis," itself subject to further experimental investigation. That is, no matter how one refines the observational approach of multivariate regression analysis, whether by statistically removing environmental factors as covariates or through the experimental identification and control

of environmental factors, it remains essentially a correlational analysis. It can suggest causation only.

The environmental manipulations that we recommend are concordant with or suggested by the prior observational analyses. When we manipulate the environment in a particular dimension and observe a change in the estimate of selection, then we can infer that the agency of selection is associated with that environmental feature. Differently put, we believe that there is an essential reciprocity between observational and experimental approaches to the study of evolutionary processes as opposed to viewing experimental manipulation primarily as a means of improving the observational analysis. In our view, the experimental study of the association between selection gradients and environmental parameters is the next and necessary step in identifying the agents of selection and in understanding the functional interaction of the phenotype and the environment that

causes fitness. The study of the regression of selection gradients on environmental parameters is the study of the causes of fitness.

ACKNOWLEDGMENTS

We thank C. Goodnight, M. McPeck, L. Stevens, S. Tonsor, P. Phillips, C. Boake, S. Arnold, L. Heisler, and J. Travis for helpful discussions during the development of this paper. S. Tonsor generously provided us with an example from his unpublished data. We also thank B. Brandon and an external reviewer who we suspect is fond of hats for comments on the manuscript. This research was supported in part by grant GM22523 to M.J.W. and grant NSF BSR-8713967 to S.K. This is the Kellogg Biological Station contribution number 665.

LITERATURE CITED

- ARNOLD, S. J. 1986. Limits on stabilizing, disruptive, and correlational selection set by the opportunity for selection. *Am. Nat.* 128:143-146.
- ARNOLD, S. J., AND M. J. WADE. 1984a. On the measurement of natural and sexual selection: Theory. *Evolution* 38:709-719.
- . 1984b. On the measurement of natural and sexual selection: Applications. *Evolution* 38:720-734.
- BREDEN, F. J., AND M. J. WADE. 1989. Selection within and between kin groups of the imported willow leaf beetle. *Am. Nat.* 134:35-50.
- BULMER, M. G. 1980. *The Mathematical Theory of Quantitative Genetics*. Clarendon Press, Oxford.
- CRESPI, B. J., AND F. L. BOOKSTEIN. 1989. A path-analytical model for the measurement of selection on morphology. *Evolution* 43:18-28.
- DOWNHOWER, J. F., L. S. BLUMER, AND L. BROWN. 1987. Opportunity for selection: An appropriate measure for evaluating variation in the potential for selection? *Evolution* 41:1395-1400.
- ENDLER, J. 1986. *Natural Selection in the Wild*. Princeton Univ. Press, Princeton, NJ.
- FISHER, R. A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- HEISLER, I. L., AND J. DAMUTH. 1987. A method for analyzing selection in hierarchically structured populations. *Am. Nat.* 130:582-602.
- HUBBELL, S. P., AND L. K. JOHNSON. 1987. Environmental variance in lifetime mating success, mate choice, and sexual selection. *Am. Nat.* 129:91-112.
- JOHANNSEN, W. 1909. *Elemente der exakten Erblichkeitslehre*. Gustav Fisher, Gena.
- KALISZ, S. 1986. Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). *Evolution* 40:479-491.
- KOENIG, W. D., AND S. S. ALBANO. 1987. Lifetime reproductive success, selection and the opportunity for selection in the white-tailed skimmer, *Plathemis lydia* (Odonata: Libellulidae). *Evolution* 41:22-36.
- LANDE, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314-334.
- LANDE, R., AND S. J. ARNOLD. 1983. The measurement of selection on correlated characters. *Evolution* 36:1210-1226.
- LEES, D. R. 1971. The distribution of melanism in the pale brindled beauty moth, *Phigalia pedaria*, in Great Britain, pp. 153-174. *In* R. Creed (ed.), *Ecological Genetics and Evolution*. Blackwell Scientific, Oxford.
- LEVINS, R. 1968. *Evolution in Changing Environments*. Princeton University Press, Princeton, NJ.
- LEWONTIN, R. C. 1955. The effects of population density and composition on viability in *Drosophila melanogaster*. *Evolution* 9:27-41.
- MANLY, B. F. J. 1985. *The Statistics of Natural Selection on Animal Populations*. Chapman and Hall, N.Y.
- MCCAULEY, D. E. 1978. Demographic and genetic responses of two strains of *Tribolium castaneum* to a novel environment. *Evolution* 32:398-415.
- . 1979. Geographic variation in body size and its relation to mating structure of *Tetraopes* populations. *Heredity* 42:143-50.
- MERTZ, D. B., D. M. CRAIG, M. J. WADE, AND J. F. BOYER. 1984. Cohort selection. *Evolution* 38:560-570.
- MITCHELL-OLDS, T., AND R. G. SHAW. 1987. Regression analysis of natural selection: Statistical inference and biological interpretation. *Evolution* 41:1149-1161.
- NEYMAN, J. 1935. Comments on Mr. Yates's paper. *Suppl. J. of Royal Stat. Soc.* 2:235-241.
- PRICE, G. R. 1970. Selection and covariance. *Nature* 227:520-521.
- ROBERTSON, A. 1966. A mathematical model of the culling process in dairy cattle. *Anim. Prod.* 8:95-108.
- SCHLUTER, D. 1988. Estimating the form of natural selection on a quantitative trait. *Evolution* 42:849-861.
- WADE, M. J. 1985. The effects of genotypic interactions on evolution in structured populations, pp. 283-290. *In* Proceedings of the XV International Congress on Genetics. Oxford & IBH Publishing Co., New Delhi.
- . 1987. Measuring sexual selection, pp. 197-207. *In* J. W. Bradbury and M. B. Andersson (eds.), *Sexual Selection: Testing the Alternatives*. John Wiley & Sons, Berlin.
- WADE, M. J., AND F. BREDEN. 1986. Life history of natural populations of the imported willow leaf beetle, *Plagioderma versicolora* (Coleoptera: Chrysomelidae). *Ann. Entomol. Soc. Am.* 79:73-79.
- WADE, M. J., AND S. KALISZ. 1989. The additive partitioning of selection gradients. *Evolution* 43:1567-1569.

Corresponding Editor: J. A. Endler

APPENDIX

Below we discuss and illustrate by example, four of the most common distributional problems encountered in the application of multivariate techniques to the analysis of real data:

1. *The association between character and fitness may be indirect or mediated by other traits.* Let the trait, z , be phenotypically correlated with another trait, y ; i.e., $\text{Cov}(z, y) \neq 0$. If selection acts directly on y but not z , say, $w(z, y) = ay$, it is still possible to express fitness as a function of z ,

$$\text{Cov}(z, w[z, y]) = \text{Cov}(z, ay) = a \text{Cov}(z, y) \neq 0. \quad (1)$$

This is the functional relationship between z and fitness, that we would infer if we measured z but not y . In this example, the mean of the phenotypic distribution of z changes not owing to direct selection on z but rather to selection on the correlated character, y . This kind of indirect association between fitness and character value was defined as "indirect selection" by Lande and Arnold (1983). (Note that this occurs owing to the phenotypic correlation of the two traits and that genetic correlations can play a similar role in the process of transmission across generations.)

2. *The fitness function may involve a linear effect of both phenotypes, say, $w(z, y) = ay + bz$, but one (y) is unmeasured.* Again, the phenotypic correlation between the two traits distorts our inference of z in relation to fitness:

$$\begin{aligned} \text{Cov}(z, w[z, y]) &= \text{Cov}(z, ay + bz) \\ &= b \text{Var}(z) + a \text{Cov}(z, y). \end{aligned} \quad (2)$$

Thus, the regression of fitness on z , obtained by dividing Equation (2) by the $\text{Var}(z)$, is found to be $(b + aK)$, the sum of the direct effect (b) and the indirect effect (aK), instead of simply b . Here, K is the regression of y on z . If $aK > 0$, then the effect of z on fitness will be overestimated and conversely if $aK < 0$. (Clearly, the magnitude of the estimation error depends upon the relative variances of z and y and the covariance between them.)

For example, in studies of the determinants of fecundity in the blue-eyed Mary, *Collinsia verna*, in the greenhouse, Kalisz (unpubl. data) found that seed weight and emergence date were positively correlated but only emergence date influenced fecundity when analyzed in a bivariate partial regression. A regression of fecundity on seed weight alone, however, produced the appearance of selection on seed weight owing to the correlation between the "unmeasured" emergence date and seed weight although no direct selection on seed weight was occurring.

3. *The shape of the phenotypic distribution of z itself may lead to an association between z and $w(z)$ for reasons other than direct selection on the linear value of z .* Lande and Arnold (1983 p. 1218) discuss the "lack of stability" of the estimates of the directional selection gradient in the presence of skewness in the phenotypic distribution. By "lack of stability" they mean that the

linear regression estimate of the direct selection gradient on the phenotype z will be different from the same estimate from a quadratic regression *unless there is no skewness*. To illustrate this, let the distribution, $p(z)$, be skewed. This is algebraically equivalent to permitting a nonzero covariance between the linear and quadratic terms of the phenotypic distribution; i.e., saying that the skewness is nonzero is algebraically equivalent to saying that the $\text{Cov}(z, z^2) \neq 0$. If fitness were a function of the linear and quadratic values of z , say, $w(z) = bz + az^2$, then

$$\begin{aligned} \text{Cov}(z, w[z, z^2]) &= \text{Cov}(z, bz + az^2) \\ &= b \text{Var}(z) + a \text{Cov}(z, z^2). \end{aligned} \quad (3)$$

When the phenotypic distribution is skewed, the second moment about zero acts like another character; i.e., z^2 is functioning in the above regression [Equation (3)] just as the correlated trait y did in the example above [Equation (2)].

As an example of the effects of skewness on the estimates of beta and gamma we use data from a long-term study of gene flow and life history in *P. lanceolata* (Tonsor, unpubl. data). Tonsor defined the reproductive effort of 1,321 plants in 1986 as the total number of infructescences produced divided by the product of corm size and the number of vegetative shoots. This measure of reproductive effort was the phenotypic character chosen for analysis and survivorship through 1987 was the measure of fitness. The log of reproductive effort was highly skewed (skewness = 1.40): the majority of plants exhibited little or no reproduction while a few plants exhibited high reproduction. The estimate of directional selection on log reproductive effort changes from 0.61 in a linear regression to 0.20 for a quadratic regression of viability on reproductive effort. This is a three-fold difference in the estimate of the strength of direct selection on reproductive effort and an example of the "instability" of the estimates of directional selection in the presence of skewness. (When we discuss the effect of transformation on our ability to identify the agents of selection [below], we will return to this same example.)

4. *Associations between z and quadratic values of another trait, y^2 , such as, $\text{Cov}(z, y^2) \neq 0$, can give rise to an association between z and fitness, $w(z)$.* For example, when $w(z, y) = ay^2$. Because the precision of the estimates of quadratic and interaction terms is always less than that of linear terms (Neyman, 1935), skewness always has the potential of plaguing our estimates of directional selection. (Note that Lande and Arnold (1983) did not investigate the existence of skewness in *Bumpus* data or in the pentatomid data that they used to illustrate the multivariate selection analysis.)