

## A PATH-ANALYTIC MODEL FOR THE MEASUREMENT OF SELECTION ON MORPHOLOGY

BERNARD J. CRESPI<sup>1</sup>

*Museum of Zoology and Department of Biology, University of Michigan, Ann Arbor, MI 48109*

AND

FRED L. BOOKSTEIN

*Center for Human Growth and Development, University of Michigan, Ann Arbor, MI 48109*

*Abstract.*—This paper describes a path model for the analysis of phenotypic selection upon continuous morphological characters. The path-analysis model assumes that selection occurs on unmeasured general size and shape allometry factors that summarize linear relations among sets of ontogenetically, phylogenetically, or functionally related traits. An unmeasured factor for general size is considered the only aspect of morphometric covariance matrices for which there is an a priori biological explanation. Consequently, selection coefficients are derived for each measured character by holding constant only a general size factor, rather than by using multiple regression to adjust for the full covariance matrix. Fitness is treated as an unmeasured factor with loadings, representing directional selection coefficients, computed as the covariances of the size-adjusted characters with the measured fitness indicator. The magnitudes and signs of the selection coefficients, combined with biological insight, may suggest hypotheses of selection on one or more shape allometry factors. Hypotheses of selection on general size and shape allometry factors are evaluated through cycles of measurement, analysis, and experimentation, designed to refine the path diagram depicting the covariances among the measured characters, the measured indicator of fitness, and unmeasured factors for morphology and fitness.

The path-analysis and multiple-regression models were applied to data from remeasurement of Lande and Arnold's (1983) pentatomid bugs and to Bumpus's (1899) data on house sparrows. The path analysis suggested the hypothesis that variation in bug survivorship was an expression of directional selection on wing loading. Bumpus's data are consistent with a hypothesis of stabilizing selection on general size in females and directional selection for small wing size relative to body size in males.

Received December 7, 1987. Accepted July 18, 1988

Understanding the evolution of morphology requires the development of robust methods for the measurement of phenotypic selection on morphological characters (Endler, 1986; Mitchell-Olds and Shaw, 1987). Because morphological characters are usually correlated among themselves, multivariate analyses of selection must consider not only the covariances between the characters and some observed measure of relative fitness but also the covariances among the characters. Lande and Arnold (1983) modeled the first of these two sets of covariances by multiple linear regression. In their adaptation of Pearson's (1896, 1903, 1911) method, each regression coefficient is interpreted as the "force" of directional or stabilizing selection upon that character, once the effects of the other characters have

been adjusted away by linear statistical methods.

The multiple-regression model for measuring selection has two advantages over univariate analyses (Lande and Arnold, 1983; Endler, 1986 pp. 189–190). First, the regression equation provides a simple method for predicting fitness from multiple characters, by computing a function for shifts in the multivariate phenotypic distribution. Second, combining the regression with genetic data allows an extrapolation of the trajectory of the population's morphometric mean with respect to the fitness function. This statistical combination of selection and inheritance is embodied in the equation  $\Delta\bar{z} = \mathbf{G}\mathbf{P}^{-1}\mathbf{s}$ , where  $\Delta\bar{z}$  is the column vector representing change in a set of characters,  $\mathbf{G}$  and  $\mathbf{P}$  are the genotypic and phenotypic variance-covariance matrices, and  $\mathbf{s}$  is a column vector of selection differentials (see Lande, 1979; Arnold, 1983).

The efficacy of multiple regression as an exploratory data analysis depends upon the

<sup>1</sup> Current address: School of Zoology, P.O. Box 1, University of New South Wales, Kensington, N.S.W., Australia 2033.

extent to which its assumptions are met in nature. For the coefficients in a multiple linear regression to be interpreted in causal terms, the measured characters must affect fitness linearly, independently, and completely, with an error term unattributable to any other measurable character correlated with those that are included in the analysis (Lande and Arnold, 1983; Endler, 1986 p. 191; Mitchell-Olds and Shaw, 1987). In accordance with these assumptions, the regression analysis inverts the full covariance matrix of the characters and predicts survivorship "optimally" by a Fisherian discriminant function. To the extent that the assumption of independence is violated, multiple-regression coefficients are unstable against variation in choice of characters. Moreover, high positive correlation among the characters, such as is caused by size factors in most morphometric data sets (Wright, 1918, 1932, 1954; Bookstein et al., 1985 pp. 85–101), induces artificial divergence of the multiple-regression coefficients. The differences among the coefficients therefore incorporate a large amount of noise, even when the covariance matrix can be inverted.

These problems with multiple regression analysis—"missing" characters and high correlation among characters—may be reduced in practice by judicious character choice and use of principal-components analysis for highly correlated data sets (Lande and Arnold, 1983; Endler, 1986 pp. 191–192). However, such adjustments a posteriori interfere with one of the primary goals of multivariate selection analysis, which is to determine how selection acts on the complex, integrated phenotypes expressed in the measured characters. More importantly, violation of the assumptions of multiple-regression analysis may be interpreted as evidence that, for morphometric data, the underlying linear model is inappropriate.

In this paper, we describe an alternative model for the measurement of selection on morphological characters. The model follows from biological and statistical assumptions different from those of multiple regression. We assume that selection occurs on two unmeasured factors, size and shape allometry, that summarize linear relations among sets of ontogenetically, phylogenet-

ically or functionally related traits. Measured characters are considered to contain biological information only insofar as they are correlated with these factors, and correlations among characters are considered to indicate their shared dependence upon factors. To estimate correlations of observed characters with unobserved factors from character data, we need a statistical method that explains covariances (rather than one that maximizes variance) and that recognizes the importance of unmeasured factors in biological explanation.

Sewall Wright developed such a method, path analysis, as an alternative to the ambiguities inherent in multiple regression analysis when applied to interactive biological systems (Wright, 1921, 1934; Li, 1981 p. 170; Provine, 1986 pp. 127–132). Path analysis is the decomposition of observed correlations to correspond to a multiplicity of causal schemes posited simultaneously. In a general sense, path analysis includes the depiction of any statistical relationship using standardized regression coefficients (Wright, 1934, 1954, 1984). However, as conceived in biological application, the method is designed to explain covariances in causal terms, often through the use of unmeasured factors that require estimation from the data (see Wright, 1921, 1932, 1934, 1984). Recently, Wright (1984 p. 30) noted, "In the first two broad classes of application of causal path analysis . . . , arrays of coordinate variables [measured characters] and phenotype-genotype relations, *all* the causal variables are hypothetical and unmeasurable. Causal multiple regression could not be applied at all."

In the first section of this paper, we explain a path-analytic model for the measurement of phenotypic selection on a set of measured characters when fitness is measured dichotomously, as by survivorship. We apply the path-analysis and multiple-regression models to a remeasured version of Lande and Arnold's (1983) data on pentatomid bugs and to Bumpus's (1899) data on house sparrows, and we compare the analyses and their interpretations. Finally, we show how the differences in assumption and interpretation between the two statistical models are based on two views of the purpose of measuring selection, which rep-

resent two approaches to the study of morphological evolution.

*A Path-Analytic Approach to the Measurement of Selection on Morphology*

*Prediction, Description, and Causation in Selection Models.*—Measuring selection begins with a set of phenotypic measurements, referred to here as “characters,” and a measurement of an indicator of relative fitness. In the examples discussed below, the characters are morphological measurements, and the indicator of fitness is survivorship. We note, however, that any type of phenotypic character could be analyzed using either multiple-regression or path-analysis models and that any indicator of fitness could be used.

The multiple-regression model for measuring selection assumes that the measured characters cause survivorship (Fig. 1a). Thus, the directional-selection coefficients are computed for each character as the predicted effects on survivorship of a unit change in that character, with all of the other characters remaining constant (see Mitchell-Olds and Shaw, 1987). The multiple-regression model provides a statistically optimal prediction, but the coefficients can be interpreted in causal biological terms only if the characters independently and completely determine survivorship. This assumption is unwarranted for virtually all selection studies.

The path-analysis model assumes: 1) that selection occurs on two unmeasured morphological factors, general size and shape allometry, for which the measured characters serve as more or less imperfect indicators; 2) that measured morphological characters are usually correlated due to an unmeasured factor for general size; and 3) that survivorship may be correlated with unobserved relative fitness. The selection coefficients produced by the path model therefore account for the covariances among the characters and between the characters and survivorship using two unmeasured factors: fitness (for which survivorship is the indicator) and general size (for which the measured characters are the joint indicators). This general size factor is considered the only aspect of morphometric covariance matrices for which there is an a priori

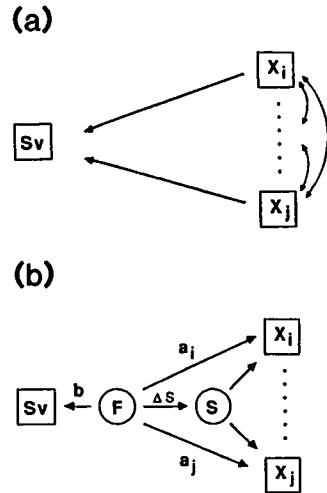


FIG. 1. Models for the measurement of selection on morphology. The measured variables (the characters  $X_i$  through  $X_j$  and survivorship  $Sv$ ) are in squares and the unmeasured factors (general size  $S$  and fitness  $F$ ) are in circles. a) Statistical prediction; the multiple-regression model predicts survivorship optimally from the measured characters. b) Statistical description; an unmeasured factor for general size explains most of the covariance among the measured characters. Fitness explains the covariance of the size-adjusted characters with survivorship, the indicator of fitness in the model. Data on causation allow reversal of the direction of the paths from fitness to the measured characters and from fitness to general size.

biological explanation. Consequently, in the path model, selection coefficients are computed by adjusting each character only for general size, rather than by using multiple regression to adjust for the full covariance matrix.

The path-analysis model constitutes a statistical description of the covariances among the measured characters, the unmeasured size factor, survivorship, and unmeasured fitness (Fig. 1b). No biological causation is implied or assumed. Although we believe that phenotypes cause fitness in time and space, we have no evidence whatsoever, without biological information external to this description, that the measured characters indicate morphological factors that affect survivorship or that survivorship is correlated with fitness. In the absence of information about the causes and objects of selection, we prefer this purely descriptive statistical model, because it prevents the inference of causation from correlation and because it explicitly recognizes the errors involved in the estimation of unmeasured

variables from measured ones. When biological information about the causes of selection on morphological factors is available, the direction of the paths from fitness to the measured characters and general size may be reversed, in order to create a model that represents a causal hypothesis of selection.

In most selection studies and for the examples discussed here, data regarding the causes and objects of selection are lacking. We therefore derive selection coefficients for the path model depicted in Figure 1b and note that the coefficients are derived for an explicitly causal model in a statistically identical way.

*Statistical Model for Measuring Selection.*—Assume first that one variable,  $F$  (which will be fitness below), is related to the values of many other variables,  $X_1 \dots X_k$ , according to separate simple regressions with independent error terms. That is, assume, for each variable  $X_i$ , that

$$X_i = a_i F + e_i$$

or, for each  $i$  and each case  $j$ ,

$$X_{ij} = a_i F_j + e_{ij},$$

where  $X_i$  is the  $i$ th measured variable,  $a_i$  is the  $i$ th regression coefficient,  $e_i$  is the error about the  $i$ th regression, and  $X_{ij}$ ,  $F_j$ , and  $e_{ij}$  are the true values of  $X_i$ ,  $F$ , and  $e_i$  for the  $j$ th case. The error terms,  $e_i$ , are assumed to be independent of each other and of  $F$ . Each observed variable,  $X_i$ , is normalized to mean zero and variance 1, as is  $F$ , so the variance of the error term ( $e_i$ ) is  $1 - a_i^2$ . By this model, the quantities  $a_i$  also serve as the correlations of the  $X_i$  with  $F$ , whereas the correlation between  $X_i$  and  $X_j$  is  $a_i a_j$ .

Suppose that this model is true but that the factor  $F$  is unobserved. How might an estimate of the value of  $F$  be constructed, case by case, from the observed variables  $X_1 \dots X_k$ ? The usual multiple-regression estimate of  $F$  from the  $X$ 's can exploit the covariances specified by the model in place of those observed. The predicted value for this regression can be shown to be a multiple of the expression

$$\sum_{i=1}^k \frac{a_i X_i}{1 - a_i^2}.$$

Up to a constant of proportionality, this is an ordinary weighted average: each regression coefficient equals its correlation ( $a_i$ ) inversely weighted according to the error variance ( $1 - a_i^2$ ) of its own separate simple regression. When all  $a$ 's are small (as they are for the data to be analyzed below) or when the  $a$ 's vary over a narrow range, one can ignore the denominators of the coefficients in this regression. The factor estimate is then proportional nearly to  $\sum a_i X_i$ , the simple sum of all the univariate predictors of  $F$  by the observed  $X$ 's separately. Bookstein (1986) calls this the "net partial predictor," in contrast to the optimum linear predictor supplied by a multiple regression. It corresponds to the output of a ridge regression (a regression with diagonal elements of the correlation matrix increased) with a very large ridge parameter, and it is very stable against substitution of characters.

The net partial predictor used for the estimation of fitness from the measured characters is also used for the prediction of other measurable characters that load on the same factor for  $F$ . Assume now that Sv (survivorship) is an additional indicator of fitness not among the variables  $X_i$ ; that is, let

$$Sv = bF + e_{sv}$$

where  $e_{sv}$  is independent of  $F$  and all of the other error terms ( $e_i$ ) in the prediction of the  $X$ 's from the same  $F$ . By this model, the correlations between Sv and the  $X$ 's are equal to the products  $ba_i$ , which are proportional to the coefficients,  $a_i$ , that relate the characters to fitness. Thus the appropriate indicator of Sv is proportional to the estimate,  $\sum a_i X_i$ , of  $F$ .

Morphometric data hardly ever meet the assumptions of the single-factor model for fitness described above. In particular, the error terms for the regression of the characters on any quantity estimating Sv will be strongly correlated. This finding indicates the existence of another factor, general size, which represents the best single explanation for the pooled within-group covariance matrix of the characters. General size is estimated by the linear combination of characters  $X_i$  having as coefficients the first eigenvector of the characters after they are centered at zero separately by survivorship group (see Bookstein et al., 1985 pp. 101–

109). That  $X_i$  contributes to the estimation of general size affects the computed significance levels of allometric coefficients but has hardly any effect on the factor loadings for fitness.

The path model for survivorship studies therefore includes at least two factors: fitness ( $F$ ) and general size ( $S$ ) (Fig. 1b). General size is computed as the single factor best explaining the correlations among the characters after an estimate of size-free fitness (survivorship, in this case), is partialled out. Symmetrically, fitness is characterized as the single factor best explaining the correlations among the characters after fitness-free general size has been partialled out. The complete model is

$$X_i = s_i S + a_i F + e_i$$

where each error term ( $e_i$ ) is independent of general size,  $F$ , and the other error terms and the  $s_i$  are the allometric coefficients relating the measured characters to general size. The model is identified by equating  $F$  to one of its indicators,  $S_v$ , for the purpose of estimating general size, and then equating general size to its estimate via the first within-group principal component for estimating loadings of  $F$  (see Rohlf and Bookstein, 1987).

*Computation and Interpretation of Selection Coefficients.*—Because survivorship is a dichotomous variable, the selection coefficients  $a_i$  that relate survivorship to any size-adjusted variable  $X_i$  are computed as the differences in adjusted mean between live and dead groups in separate analyses of covariance adjusting for general size. When fitness is measured as a continuous variable, the selection coefficients are the generalization of size-adjusted mean differences: the partial covariances of fitness with the characters adjusted for general size.

The meaning of these selection coefficients  $a_i$ , referred to here as “shape coefficients,” is shown in Figure 2. The path model asserts that each variable  $X_i$  is derived from the appropriate regression line (live or dead) in this figure by the addition of a suitable error term, independent of the covariate and of all the error terms of the other characters. These selection coefficients can be considered intermediate between selection differentials, which correspond to sim-

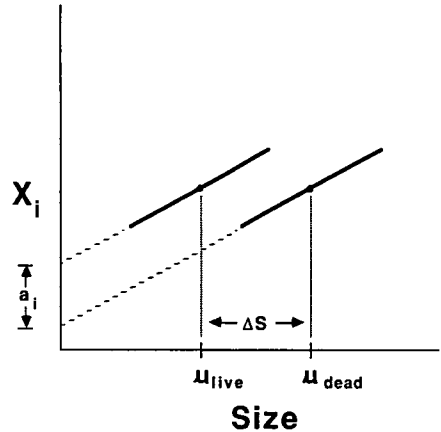


FIG. 2. The path coefficients  $a_i$  of Figure 1b are adjusted mean differences in an analysis of covariance of the characters upon general size. The usual assumptions of covariance analysis, particularly the equality of regression slopes, must be satisfied for the coefficients to be interpreted one by one.

ple mean differences, and selection “gradients” (partial regression coefficients; Lande and Arnold, 1983), which correspond to mean differences as adjusted for all of the other characters separately, rather than for the common factor, general size, that is their joint explanation.

The analysis of covariance that generates the loadings on fitness offers a partition of covariance incorporating four separate aspects of selection. Directional selection on general size is observed as a mean difference between survivorship groups. Stabilizing selection on general size is observed in differences between the two within-group variances of size whenever the mean sizes are effectively the same. Directional selection on shape allometry is observed as the vector of size-adjusted mean differences between survivorship groups. Finally, stabilizing selection upon shape allometry is observed in any decrease of the residual variance of the analysis of covariance from the dead group to the live group.

The path analysis continues with an inspection of the resulting shape coefficients and variances. Similarities among the characters in the signs and magnitudes of the size-adjusted mean differences, combined with biological information about possible causes of covariance between aspects of morphology and fitness, may suggest hy-

potheses of selection on shape allometry factors. These are groups of characters (i.e., the "group size" factors of Wright [1932, 1934]) that reflect a single underlying factor subjected more or less independently to selection or artifacts. These shape allometry factors are constructed from appropriate subsets of the measured characters (as the general size factor was constructed above), and their biological significance is evaluated in two ways. First, additional characters are measured, some of which are expected to serve as indicators of the hypothesized shape allometry factors. This augmented data set is then subjected to the original path analysis, to assess the consistency of the shape coefficients in terms of the proposed factors. Second, experiments involving manipulation of the morphological factors in the field or observations under controlled conditions may reveal the causes of selection. The results of such analyses of selection on shape allometry and general size factors are used to construct a new path diagram consistent with the available data on causation and correlation.

*Analysis of Lande and Arnold's  
Pentatomid Bugs*

*Data.*—Lande and Arnold (1983) describe their collection of 39 live and 55 dead pentatomid bugs (*Euschistus variolarius* PB) washed onto the shoreline of Lake Michigan after a storm. They provided us with their specimens, which we remeasured using digitized landmark coordinates. We computed linear distances that corresponded closely to their characters: head width, pronotum width, scutellum length, and forewing length (Table 1). Because some specimens were damaged, the new data set was obtained by remeasuring 85 complete specimens, including 55 males (28 "live") and 30 females (9 "live").

The new data differed from the original data (Lande and Arnold, 1983) in showing higher correlations between the characters, almost no difference between live and dead bugs in scutellum length, and a small, non-significant regression coefficient for pronotum width (Table 1). However, both the original and new data assigned a substantial regression coefficient to wing length. Path

analysis of the new data yielded a significant shape coefficient only for wing length, and there was no general size difference between survivorship groups.

*Interpretation.*—The differences between the original and new data sets in the correlations among the characters and between the characters and survivorship are attributable to accidents of measurement and, perhaps, to the elimination from the sample of the nine bugs not remeasured. Lande and Arnold (1983) interpret the significant regression coefficients of opposite sign for wing length and pronotum width in their data as indicating selection for small wings and a wide pronotum. The path analysis of the remeasured data suggests a somewhat different interpretation. In these data, the shape coefficients for body dimensions are either positive or close to zero. Thus, head width, pronotum width, and scutellum length may be interpreted as indicators of a shape allometry factor for body size. A variable for body size was constructed from these three distances, and comparison of the plots of  $\log(\text{wing length})$  vs.  $\log(\text{body size})$  showed that live bugs had smaller wings relative to body size than did dead ones (Fig. 3). This finding is consistent with an hypothesis of directional selection on wing loading, the allometric relationship between body and wing size. This hypothesis could be evaluated by the experimental manipulation of wing loading (measured properly as body weight divided by wing surface area), followed by observations about flying ability, or by measuring the dry weights and wing surface areas of the live and dead bugs.

If the wing-loading hypothesis is correct, another set of measured characters reflecting the same factors should yield similar findings. A larger set of characters from the remeasured data was subjected to both multiple regression and the path analysis (Table 2). The multiple regression, although significant overall, gave no indication of selection for relatively small wings. By contrast, the path analysis yielded significant shape coefficients for wing length and wing width. Covariance analysis of shape allometry factors for wing size (wing length and width) and body size (head, pronotum, and scutellum characters) was consistent with the wing-loading hypothesis.

TABLE 1. Summary of analyses of pentatomid bugs. Lande and Arnold (1983) measured head width at the eyes, pronotum width posteriorly (at its widest point), scutellum length along the dorsal midline, and forewing length from insertion to the apex. We measured identical head and pronotum characters, scutellum length as the midline of the triangle defined by the apices of the scutellum, and wing length along the subcostal (anterior) margin of the corium (the basal portion of the forewing). All data are transformed to natural logarithms.

Statistics	Head	Pronotum	Scutellum	Wing
Descriptive statistics:				
Mean	0.892	2.045	1.476	1.933
Pooled SD	0.027	0.041	0.048	0.039
Correlations				
Head	1.000	—	—	—
Pronotum	0.76	1.000	—	—
Scutellum	0.72	0.82	1.000	—
Wing	0.67	0.85	0.82	1.000
Analyses of survivorship:				
Regression coefficient	4.025	0.158	3.628	-8.198**
Mean, live	0.894	2.041	1.475	1.925
Mean, dead	0.891	2.048	1.476	1.940
Mean difference (live minus dead)	0.003	-0.007	-0.001	-0.015
Mean difference, size-adjusted	0.0062	-0.0011	0.0059	-0.0103**

\*  $P < 0.05$ ; \*\*  $P < 0.01$ .

#### *Analysis of Bumpus's House Sparrows*

**Data.**—The data collected by Hermon Bumpus (1899) on house sparrow (*Passer domesticus*) survivorship following a prolonged winter storm have become a favorite for demonstrating the measurement of selection (Harris, 1911; Calhoun, 1947; Grant, 1972; Johnson et al., 1972; O'Donald, 1973; Manly, 1976). Bumpus's measurements included total body length, wing length (alar extent), body weight, head length, humerus length, femur length, tibiotarsus length, skull width, and sternum length on 49 females (21 live) and 87 males (51 live).

There are serious methodological problems with Bumpus's measurements of weight and total length (Grant, 1972; Johnston et al., 1972). The surviving birds had obviously been alive more recently than the dead ones and so had continued to lose weight by respiration and excretion; in small birds such weight losses may be substantial (Packard, 1967). Total length is also likely to be spuriously related to survivorship, because the vertebral column of birds is S-shaped and therefore is affected by both rigor mortis and the possibility of incomplete thawing at the time of measurement (Grant, 1972; R. Storer, pers. comm.). We will describe the multiple-regression analysis and the path analysis for the original nine characters and then for the data set with both weight and total length excluded.

Multiple-regression analysis of the nine-character data set (Lande and Arnold, 1983) yielded significant coefficients for weight and total length in males; in females, the analysis yielded a significant coefficient for weight but lack of significance overall (Table 3). Lande and Arnold (1983) also detected significantly lower variance in the first prin-

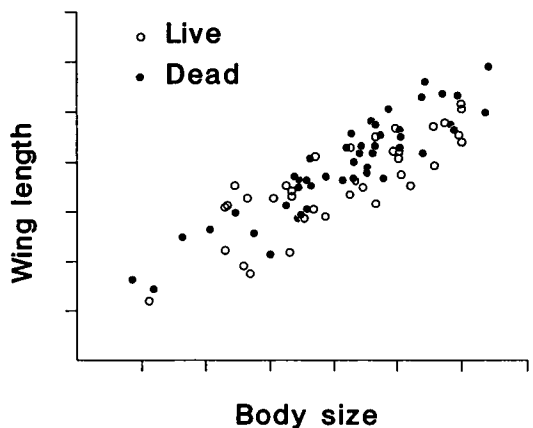


FIG. 3. The relationship between wing length and body size (a factor indicated by head width, pronotum width, and scutellum length) in pentatomid bugs *Euschistus variolarius* collected while they were either live or dead. Live bugs averaged smaller wings relative to body size (one-way analysis of covariance,  $P < 0.05$ , with either variable as dependent). The ANCOVA significance test is used only heuristically; when the variables are highly correlated (as they are here), computations using vertical and orthogonal deviations yield similar results (see Gould, 1966).

TABLE 2. Multiple regression and path analyses of an augmented character set for the Crespi-Bookstein data. Head, pronotum 1, scutellum 1, and wing 1 are as described in Table 1. Pronotum 2 is pronotum width measured anteriorly, scutellum 2 is the anterior width of the scutellum, and wing 2 is wing width along the distal margin of the corium. The multiple regression is highly significant ( $R^2 = 0.22$ ). Analysis of covariance using wing size and body size as either dependent or independent variables gave significant differences between live and dead bugs.

Statistic	Head	Pronotum 1	Pronotum 2	Scutellum 1	Scutellum 2	Wing 1	Wing 2
Mean	0.892	2.045	0.970	1.476	1.311	1.933	1.223
Pooled SD	0.027	0.041	0.039	0.048	0.041	0.039	0.041
Mean difference, live minus dead	0.003	-0.007	-0.005	-0.001	-0.014	-0.015	-0.020*
Multiple regression coefficient	5.366	2.216	-1.341	5.544*	-5.258*	-4.337	-3.884
Adjusted mean differ- ence (shape coef- ficient)	0.0085*	0.0028	0.0023	0.0105**	-0.0057	-0.0066*	-0.0120**

\*  $P < 0.05$ ; \*\*  $P < 0.01$ .

cipal component of the pooled covariance matrix (a general size factor) in live females. Elimination of total length and weight from the multiple-regression equations results in a significant coefficient only for wing length in males. Note also the instability of the coefficients in Table 3; for example, the coefficient for wing length changed from -3.16 to 10.38 when weight and total length were omitted.

Path analysis of the nine-character data set revealed significant "shape" coefficients for weight and total length in males. In fe-

males, the shape coefficient for weight was significant; also, there was lower variance in live females for general size. In the path analysis of the seven-character data set, the shape coefficients remained generally stable (Table 3). The coefficient for male wing length, however, was somewhat higher, indicating that live males had smaller wings relative to body size, or, more precisely, larger bodies relative to wing size. Live females still exhibited lower variance for general size.

*Interpretation.*—The apparent selection

TABLE 3. Analysis of Bumpus's house sparrows, using his original nine characters and a subset of seven characters not suspected a priori to be subject to methodological artifacts.

Character	Regression Coefficient		Shape Coefficient	
	Nine characters	Seven characters	Nine characters	Seven characters
<b>Females (<math>N = 49</math>):</b>				
Total length	3.813	—	0.0074	—
Wing length	2.201	5.877	0.0031	0.0039
Weight	11.881*	—	0.0100*	—
Head length	-0.600	2.810	0.0024	0.0035
Humerus	-3.368	-0.188	-0.0021	-0.0007
Femur	1.801	-0.671	-0.0059	-0.0043
Tarsus	-5.354	-6.110	-0.0100	-0.0084
Skull width	0.058	2.480	0.0033	0.0044
Sternum	-1.521	-0.398	0.0029	0.0048
<b>Males (<math>N = 87</math>):</b>				
Total length	15.144**	—	0.0204**	—
Wing length	-3.161	10.376*	0.0040	0.0052*
Weight	8.785**	—	0.0129**	—
Head length	-1.452	0.772	-0.0012	0.0000
Humerus	-2.998	-6.107	-0.0054	-0.0027
Femur	-2.020	-0.958	-0.0032	-0.0003
Tarsus	0.275	0.762	-0.0003	0.0027
Skull width	-3.368	-0.512	-0.0006	0.0005
Sternum	-2.301	-1.532	-0.0055	-0.0028

\*  $P < 0.05$ ; \*\*  $P < 0.01$ .



on weight and total length detected by both analyses of the nine-character data set is consistent with the hypothesis that these variables are spuriously correlated with survivorship because of methodological artifacts. The existence of these artifacts could be tested experimentally by measuring total length in relation to sparrow thawing and rigor mortis and by measuring the effects of winter storm conditions and recent death on weight. Without such tests, Lande and Arnold's (1983) interpretation of the nine-character data set—that the winter storm apparently favored small birds—should be viewed with caution.

Interpretation of the hypothesized directional selection on the allometry of male wing and body size in the two analyses of the seven-character data set is, as for the pentatomid bugs, clouded by ignorance of the relevant causal factors. (Grant [1972] discusses the apparent stabilizing selection on female size.) Relatively small wings in surviving males may result from phenotypic selection or from an artifact involving contraction or partial freezing of muscle in one survivorship group (as for total length), or wing wear in caged survivors after the storm (see Calhoun, 1947).

#### DISCUSSION

The path-analysis and multiple-regression models differ in assumptions, scope, and purpose. The purpose of the multiple-regression model is to provide an optimal statistical prediction of fitness from a set of measured characters that are assumed to affect survivorship independently and completely. By contrast, the path-analysis model explains the covariance among the measured characters and between the characters and survivorship in terms of unmeasured factors for fitness, general size, and shape allometry and seeks to uncover the morphological factors subject to selection through cycles of measurement, analysis, and experimentation. The purpose of the path model is to combine biological insight with statistical evaluation in developing a path-diagram hypothesis of selection consistent with the data at hand and to test this hypothesis to whatever extent is possible.

The path model avoids the problems of character choice and character correlation

in several ways. First, because the model adjusts the measured characters only for a general size factor, rather than for the full covariance matrix, the selection coefficients remain relatively stable to inclusion and exclusion of characters. The model therefore rewards rather than punishes measurement of large numbers of correlated characters. Second, the path method provides a non-arbitrary criterion for character choice, whereby the signs and magnitudes of the shape coefficients are combined with biological knowledge to reveal hypotheses of selection on shape allometry factors. Indeed, identification of morphological factors subject to selection is one of the primary purposes of the path model. Third, the path model reduces the conceptual and empirical problem of "missing" characters, because it assumes that all measured characters more or less imperfectly reflect unmeasurable general size and shape allometry factors. Because of its robustness, the model suggests new characters to measure better in the next data set and provides an interactive method, driven by hypotheses, which may eventually reveal the causes of selection. Endler (1986 pp. 162–165) and Mitchell-Olds and Shaw (1987) argue (and we concur) that biological understanding of the causes of selection is ultimately the only way to determine how selection occurs on phenotypes.

The path model described here has two limitations. First, the model assumes that a general size factor exists in the data. The existence of a general size factor can be tested by inspection of a correlation matrix for a large majority of positive coefficients or by inspecting the loadings of the characters on the first principal component of the covariance matrix. For some data sets, the measured characters will not reflect general size variation (e.g., Wright, 1968 pp. 331–335). However, an explicable part of the morphometric covariance matrix need not be limited to a factor for general size. Any factor for which there is biological evidence can be constructed and adjusted for, through simple changes in the path model, so long as the statistical adjustment is biologically plausible (Endler, pers. comm.; see Bookstein et al., 1985 pp. 123–126). Moreover, any characters that are suspected a priori to

be independent of general size can be treated as independent net partial predictors of fitness in the model. Such flexibility may be the greatest strength of Wright's path-analytic techniques.

Second, the selection coefficients computed by the path model cannot be combined directly with genetic data to yield expected evolutionary trajectories, as can the partial coefficients of multiple regression (Lande and Arnold, 1983). The importance of this limitation depends on whether one's purpose is the optimal statistical prediction of selection response from more or less arbitrarily chosen characters or is instead the explanation of the covariance among the characters, morphological factors, survivorship, and fitness, with the ultimate goal being to determine the causes and objects of selection. In this latter context, the approaches to analyzing phenotypic selection that multiple regression and path analysis represent may be viewed as complementary. The multiple-regression function may, in principle, be used with characters derived by path analysis to yield predictions of morphological change. The accuracy and usefulness of these predictions depend upon the critical assumption of multiple-regression analysis: that the included variables affect fitness linearly, independently, and completely. Whether these assumptions are met by real data must be examined case by case. However, we agree with Wright's (1935 p. 264) view: "The system of peaks relative to one character is not independent of that relative to another. Moreover, it is the harmonious adjustment of all characteristics of the organism that is the object of selection, not the separate metrical 'characters.'"

The pentatomid bug and house sparrow examples illustrate the differences between the path and multiple-regression models in data analysis and interpretation. First, in both the bug and sparrow analyses, the multiple-regression coefficients are less stable than the shape coefficients as the list of characters is varied. This instability may be interpreted to mean that the measured characters are mainly expressing size and allometry factors subject to selection, so that the assumptions of the multiple-regression model are necessarily violated. Second, in the bug analyses, the path model suggested

a new hypothesis for interpretation of the data: directional selection on an unmeasured variable, wing loading. Such hypotheses are difficult to infer from multiple-regression analysis, which assumes that selection occurs on the measured characters separately. Third, the sparrow data set and one of the two bug data sets show apparently spurious correlations between the characters and survivorship. Such ambiguities emphasize the importance of a thorough understanding of all the covariance in a data set, not just that attributable to the effects of selection. Indeed, the bug and sparrow examples indicate that, in the absence of causal information, "selection" on the measured "characters" should be given no greater a priori import than any other hypothesis inferred from a matrix of numbers.

The partition of shape from size provided by the path model allows novel tests of several important hypotheses in the study of morphological evolution. We may ask, for example, whether stabilizing selection for particular allometric relationships is common (e.g., Schluter and Smith, 1986) or whether developmental programs regulating allometry are usually conserved under selection. Similarly, the model encourages tests of the hypothesis that directional selection often occurs purely on general size, which could lead to allometric maladaptation (e.g., Gould, 1974), in comparison with selection jointly on general size and shape allometry.

Path-analysis models are not restricted to static analyses of selection on morphological characters. Selection for ontogenetic size and shape trajectories can also be analyzed by path methods (see Bookstein et al., 1985 pp. 192–205), which may lead to hypotheses of selection for changes in developmental timing. When other types of characters, such as physiological or behavioral traits, are analyzed, path diagrams can be constructed in accordance with the requisite biological knowledge and assumptions (Wright, 1968 pp. 299–372; Li, 1981).

#### ACKNOWLEDGMENTS

We are grateful to P. Buckup and B. Fink for help with digitizing, W. R. Dawson, R. Storer, R. B. Payne, and T. Van't Hof for information about sparrows, S. Frank, L.

Vawter, and the members of the University of Michigan Museum of Zoology Graduate Student Research Workshop for clarifying our thinking about selection and causality, and S. Arnold, J. Crow, N. MacLeod, and D. Schluter for thoughtful comments and discussion. We also thank D. Schultz, who was instrumental in the conception of this project, S. Arnold, for loaning us the bugs and allowing us to analyze his data, and M. O'Brien, for expert graphics assistance. We are very grateful to J. Endler for helpful comments and ideas, and we thank two anonymous reviewers for clarifying the differences between the models.

## LITERATURE CITED

- ARNOLD, S. J. 1983. Morphology, performance and fitness. *Amer. Zool.* 23:347-361.
- BOOKSTEIN, F. L. 1986. The elements of latent variable models: A cautionary lecture, pp. 203-230. *In* M. E. Lamb, A. L. Brown and B. Rogoff (eds.), *Advances in Developmental Psychology*, Vol. 4. Erlbaum, Hillsdale, NJ.
- BOOKSTEIN, F. L., B. CHERNOFF, R. ELDER, J. HUMPHRIES, G. SMITH, AND R. STRAUSS. 1985. Morphometrics in Evolutionary Biology. The Geometry of Size and Shape Change, with Examples from Fishes. *Acad. Nat. Sci. Philadelphia*, Philadelphia, PA.
- BUMPUS, H. C. 1899. The elimination of the unfit as illustrated by the introduced sparrow, *Passer domesticus*. *Biol. Lect. Woods Hole Mar. Biol. Sta.* 6: 209-226.
- CALHOUN, J. B. 1947. The role of temperature and natural selection in relation to the variations in the size of the English Sparrow in the United States. *Amer. Natur.* 81:203-228.
- ENDLER, J. A. 1986. *Natural Selection in the Wild*. Princeton Univ. Press, Princeton, NJ.
- GOULD, S. J. 1966. Allometry and size in ontogeny and phylogeny. *Biol. Rev.* 41:587-640.
- . 1974. The origin and function of "bizarre" structures: Antler size and skull size in the "Irish elk" *Megaloceros giganteus*. *Evolution* 28:191-220.
- GRANT, P. R. 1972. Centripetal selection and the House Sparrow. *Syst. Zool.* 21:23-30.
- HARRIS, J. A. 1911. A neglected paper on natural selection in the English Sparrow. *Amer. Natur.* 45: 314-318.
- JOHNSTON, R. K., D. M. NILES, AND S. A. ROHWER. 1972. Hermon Bumpus and natural selection in the house sparrow *Passer domesticus*. *Evolution* 26: 22-31.
- LANDE, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* 33:402-416.
- LANDE, R., AND S. J. ARNOLD. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210-1226.
- LI, C. C. 1981. *Path Analysis—A Primer*. Boxwood, Pacific Grove, CA.
- MANLY, B. F. J. 1976. Some examples of double exponential fitness functions. *Heredity* 36:229-234.
- MITCHELL-OLDS, T., AND R. G. SHAW. 1987. Regression analysis of natural selection: Statistical inference and biological interpretation. *Evolution* 41: 1149-1161.
- O'DONALD, P. 1973. A further analysis of Bumpus' data: The intensity of natural selection. *Evolution* 27:398-404.
- PACKARD, G. C. 1967. House Sparrows: Evolution of populations from the Great Plains and Canadian Rockies. *Syst. Zool.* 16:73-89.
- PEARSON, K. 1896. Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. *Phil. Trans. Roy. Soc. London Ser. A* 187:253-318.
- . 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection of the variability and correlation of organs. *Phil. Trans. Roy. Soc. London Ser. A* 200:1-66.
- . 1911. On the general theory of the influence of selection on correlation and variation. *Biometrika* 8:437-443.
- PROVINE, W. 1986. *Sewall Wright and Evolutionary Biology*. Univ. Chicago Press, Chicago, IL.
- ROHLF, F., AND F. L. BOOKSTEIN. 1987. A comment on shearing as a method for "size correction." *Syst. Zool.* 36:356-367.
- SCHLUTER, D., AND J. N. M. SMITH. 1986. Genetic and phenotypic correlations in a natural population of song sparrows. *Biol. J. Linn. Soc.* 29:23-36.
- WRIGHT, S. 1918. On the nature of size factors. *Genetics* 3:367-374.
- . 1921. Correlation and causation. *J. Agr. Res.* 20:557-585.
- . 1932. General, group and special size factors. *Genetics* 17:603-619.
- . 1934. The method of path coefficients. *Ann. Math. Stat.* 5:161-215.
- . 1935. Evolution in populations in approximate equilibrium. *J. Genet.* 30:257-266.
- . 1954. The interpretation of multivariate systems, pp. 11-34. *In* O. Kempthorne, T. A. Bancroft, J. W. Gowen, J. L. Lush (eds.), *Statistics and Mathematics in Biology*. Iowa State Univ. Press, Ames.
- . 1968. *Evolution and the Genetics of Populations*, Vol. I. Genetic and Biometric Foundations. Univ. Chicago Press, Chicago, IL.
- . 1984. Diverse uses of path analysis, pp. 1-34. *In* A. Chakravarti (ed.), *Human Population Genetics*. Van Nostrand Reinhold, N.Y.