UNIVARIATE METRICS ARE NOT ADEQUATE TO MEASURE AVIAN BODY SIZE

SCOTT FREEMAN AND WENDY M. JACKSON
Department of Zoology, NJ-15, University of Washington, Seattle, Washington 98195 USA

ABSTRACT.—In the literature on avian contests and energetics, a single external measure is frequently used to represent overall body size. In an attempt to assess the quality of information available in such external measurements of birds, we measured internal and external elements from museum skeletal plus flat-skin specimens for both sexes of three passerine species. We estimated the "true" overall size of each individual as a factor score computed from the first principal component extracted from a large matrix of skeletal measurements. Bivariate correlations and stepwise regressions indicate that mass or tarsus length, or a principal component factor which combines mass and tarsus length, is the best predictor of overall body size as estimated from bone measurements. Multiple regressions, however, suggest that several external measurements combined often explain only 40–60% of the total variance in overall body size. We suggest that fieldworkers be cautious in their use of single external metrics to represent overall size in small birds. When a single metric for the body size of small passerines is required, fieldworkers should prefer tarsus length or mass to represent overall size. *Received 6 October 1988, accepted 18 July 1989.*

BEHAVIORAL and physiological studies that seek correlations between avian body size and other variables are faced with the problem of how to assess body size. Often the variable of interest is body mass, for example, because of the importance of mass in energetics (metabolic rate, daily caloric demands) or dominance (ability to win escalated contests). Measuring body mass accurately is difficult because mass fluctuates seasonally, daily, or even hourly, depending upon variables such as time since feeding, weather, and activity. Investigators who work with birds in captivity can reduce the effects of these confounding factors by measuring mass at a standardized time of day (often early morning) and by taking repeated measures. Field-workers seldom have these options. Our goals in this study were to determine which (or which combination) of the external measures commonly employed by field biologists best predicts actual body size, and to determine the efficiency of this predictor.

We estimated actual body size from a principal component analysis (PCA) of a large number of skeletal measures. The first factor extracted from a PCA of various size measurements has been interpreted by many researchers as an index of overall body size (e.g. Robins and Schnell 1971, Rohwer 1972, Niles 1973, Zink 1982, Schluter 1984, Rising 1987). There are also three theoretical grounds on which to defend our assumption that this internal metric is the best possible measure of "true" body size. First, the internal metric is composed of many different characters, and it should therefore be less prone to variance caused by developmental abnormalities or measurement errors in single characters. Second, the repeatability of bone measurements is greater than the repeatability of most external measures because bone measurements can be made more precisely, and because many external measures vary through time. For example, wing chord and tail length vary both with the individual's age and the degree of feather wear, and beak length varies through time due to rhamphotheca wear. Third, and most importantly, a metric composed of size measurements from many bones summarizes the amount of total attachment surface available for muscle and connective tissue as well as the amount of support structure for internal organs. This metric should be an accurate measure of "structural size" and a biologically sensible predictor of "average massiveness" within populations.

METHODS

We measured internal and external features of male and female Red-winged Blackbirds (*Agelaius phoeniceus*), Red Crossbills (*Loxia curvirostra*), and Harris' Sparrows (*Zonotrichia querula*). The crossbills were all collected near Mt. Rainier in Pierce County, Washington, in August 1974, the Blackbirds during the springs of 1977 and 1978 in Grant County, Washing-

TABLE 1. Measurements taken in this study.

Element	Description/reference
Skeletal measurements	
Skull length	Rohwer 1972
Skull width	Rohwer 1972
Beak length	Cranio-facial hinge to tip of upper mandible
Beak width	Rohwer 1972
Sternum length	Robins and Schnell 1971, Rohwer 1972
Sternum width	Rohwer 1972
Synsacrum length	Rohwer 1972
Synsacrum width	Rohwer 1972
Humerus length	Robins and Schnell 1971, Rohwer 1972
Ulna length	Robins and Schnell 1971, Rohwer 1972
Femur length	Robins and Schnell 1971, Rohwer 1972
Tibiotarsus length	Robins and Schnell 1971, Rohwer 1972
Tarsometatarsus length	Robins and Schnell 1971, Rohwer 1972
Flat-skin measurements	
Mass	To nearest 0.5 gram, measured in lab by preparator of flat-skin specimen
Wing chord	Rohwer 1972
Tail length	Rohwer 1972
Tarsus length	From pit at junction of tibiotarsus and tarsometatarsus to last undivided scute

ton, and the sparrows in Marshall County, Kansas, in November 1982. Our criteria for selecting species to study were to include a range of passerine body sizes and to utilize species with large sample sizes in the skeleton and flat-skin collections at the University of Washington Burke Museum. We analyzed the sexes separately throughout the study; all individuals measured had completely pneumatized skulls. We eliminated first-year Red-winged Blackbirds from the analysis (using epaulet color for females and general plumage color for males) and first-year Harris' Sparrows (using plumage characters, after Rohwer et al. 1981). We were not able to age the crossbills by any external character.

All measurements from skeletal and flat-skin specimens (Table 1) were made by a single observer for each data set except for the female Red-winged Blackbirds, where some skeletal measurements were taken by a third observer. A *data set* here refers to a series of measurements from one sex of one species. Our goal in selecting bone measurements was to obtain information on the size of head, trunk, and limbs to include in the PCA. We performed all PCAs of the variables involved on the Pearson correlation matrix, which weights all variables equally, rather than the variance-covariance matrix, which would have weighted variables by their variances. Such an analysis would have overemphasized the importance of long bones because they have larger means, and thus larger variances, and the factor scores would have been less a general body-size metric and more a limb metric. After performing a PCA, we computed factor scores for each individual from factor loadings of either the unrotated principal axes or from a varimax (orthogonal) rotation of the axes, depending upon which solution produced higher loadings for most of the variables and was thus more readily interpretable as an index of overall size (see Robins and Schnell 1971).

We excluded from the analysis all individuals that had more than two broken bones, and those that were missing any external element (e.g. highly worn tail, broken tarsus). If an individual had one or two broken bones, we substituted an estimate for the missing element(s) by finding the variable most highly correlated with the missing bone measure, calculating the percentage difference between the individual's correlated measure and the population mean for that measure, and adding (or subtracting) the same percentage difference to the population mean for the missing element. We compared the correlations between factor scores for individuals and their external measurements for the sample of birds with no missing bones with those computed for the larger sample, and we found that they were virtually identical (analysis not shown).

We checked all variables for deviations from Gaussian expectations for skewness and kurtosis before performing bivariate correlations or PCAs. There were no significant deviates from Gaussian expectations (data not shown). We also made histograms of all residuals, and we plotted residuals versus fitted values during our regression analyses to check the assumptions underlying the regression statistics. We performed both stepwise and multiple regressions for each data set because we sought to ascertain which external variable or variables best predicted an index of body size computed from internal measurements. We also wanted to identify how much of the total variance in body size could be explained by external metrics (i.e. how efficiently traditional external measures predicted an internal metric of body size).

RESULTS

Principal component analyses.—As an index of "true" body size, the quality of the first principal component extracted from the correlation matrix of bone measurements (hereafter PC1-bones; Table 2) varied among species and between sexes. Factor loadings (the contribution each skeletal element made to the first principal

TABLE 2. Factor score loadings for the first principal component (unrotated solution) extracted from a Pearson correlation matrix of bone measurements. NM indicates elements that were not measured. Sample sizes are in parentheses.

	Blackbirds		Crossbills		Sparrows	
Variable	Male (64)	Female (46)	Male (36)	Female (32)	Male (16)	Female (17)
Skull length	0.522	0.576	0.562	0.600	0.515	0.787
Skull width	0.455	0.126	0.617	0.436	0.582	0.731
Beak length	0.641	0.242	0.639	0.257	NM	NM
Beak width	NM	NM	0.447	0.522	0.583	0.595
Sternum length	0.452	0.480	0.499	0.573	0.240	0.762
Sternum width	0.328	−0.044	0.079	0.299	0.355	0.332
Synsacrum length	0.708	0.507	0.185	0.729	0.544	0.849
Synsacrum width	0.601	0.440	0.069	0.600	0.425	0.770
Humerus length	0.846	0.921	0.859	0.730	0.790	0.850
Ulna length	0.851	0.912	0.873	0.837	0.906	0.889
Femur length	0.831	0.873	0.819	0.804	0.810	0.859
Tibiotarsus length	0.854	0.853	0.861	0.806	0.788	0.895
Tarsometatarsus length	0.874	0.816	0.696	0.807	0.749	0.919

axis extracted) were highest for male blackbirds, female crossbills, and female sparrows, and they were lowest for male crossbills. We therefore have the most confidence that our PC1-bones factor scores represent "true" body size for male blackbirds, and female crossbills and sparrows.

We found (Table 3) that the PCAs for the four flat-skin (external) elements separated wing and tail from tarsus and mass in three of the data sets. There are several possible explanations for this result. If these phenotypic correlations reflect underlying genetic correlations (Boag 1983), we could interpret this result as an indication that wing and tail length may be controlled by genes independent of those that influence tarsus length and body mass. Alternatively, the PCA might separate the feather measurements because wear contributes error variance that is independent of mass and tarsus size. The wing-tail vs. tarsus-mass patterns observed in some data sets, whether explained by genetics or feather wear, could be disrupted by heterogeneity in time-since-molt or the individual's age (because wing length increases from the first to second year in passerines). This may be an explanation for the lack of the pattern in crossbills, the only species which we could not age by external characters.

Bivariate correlations.—In general, the highest bivariate correlations were between PC1-bones and mass or tarsus (Table 4).

Stepwise regressions.—The stepwise regressions (Table 5) always selected mass or tarsus as the most important variable in explaining PC1-bones. When the PCA for external variables had separated tarsus and mass from wing and tail, the stepwise regression always found the factor summarizing tarsus and mass to be the most important variable in explaining PC1-bones.

TABLE 3. Factor score loadings for the first principal component (solution type indicated) extracted from a matrix of external measurements from flat-skin specimens.

	Blackbirds			Crossbills				Sparrows			
	Male[a]	Female[b]		Male[b]		Female[b]		Male[b]		Female[b]	
Variable	(1)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Mass	0.76	0.04	0.79	0.88	0.49	0.61	−0.54	0.09	−0.65	0.81	0.05
Wing chord	0.81	0.92	0.16	0.86	−0.05	0.75	−0.05	0.83	0.16	−0.31	0.82
Tail length	0.84	0.93	0.04	0.45	0.83	0.12	0.92	0.79	−0.20	0.26	0.85
Tarsus length	0.70	0.12	0.75	−0.06	0.82	0.82	0.23	0.07	0.81	0.83	−0.07

[a] PCA Solution: unrotated. Factor number is in parentheses.

[b] PCA Solution: orthogonal. Factor number is in parentheses.

TABLE 4. Bivariate correlations between internal and external measurements. NA indicates correlations between morphological elements and PCA factors which included that element. Males, lower left; females upper right. * = $P < 0.05$; ** = $P < 0.01$.

	Mass	Wing chord	Tail length	Tarsus length	PC1-bones	PC1-flats	PC2-flats
			Red-winged Blackbirds[a,b]				
Mass	—	0.16 (43)	0.11 (36)	0.15 (43)	0.28 (43)	NA	NA
Wing chord	0.46** (53)	—	0.73** (39)	0.16 (46)	0.35* (46)	NA	NA
Tail length	0.47** (53)	0.67** (53)	—	0.18 (39)	0.41* (39)	NA	NA
Tarsus length	0.46** (53)	0.37** (53)	0.43** (53)	—	0.48** (46)	NA	NA
PC1-bones	0.50** (60)	0.48** (54)	0.37** (54)	0.72** (54)	—	0.44** (36)	0.54** (36)
PC1-flats	NA	NA	NA	NA	0.69** (51)	—	NA
			Red Crossbills[a,c]				
Mass	—	0.31	0.39*	0.18	0.48**	NA	0.18
Wing chord	0.12	—	0.26	−0.21	0.32	NA	−0.57**
Tail length	−0.11	0.44*	—	0.00	0.43*	NA	−0.10
Tarsus length	0.30	0.41*	0.08	—	0.32	0.09	NA
PC1-bones	0.40*	0.17	−0.20	0.57**	—	0.57**	0.18
PC1-flats	−0.06	NA	NA	0.43**	0.05	—	−0.10
PC2-flats	NA	0.47**	−0.05	NA	0.60**	0.30	—
			Harris' Sparrows[a,d]				
Mass	—	0.14	−0.14	0.39	0.51*	NA	0.05
Wing chord	0.03	—	0.40	0.12	0.27	0.26	NA
Tail length	0.04	0.33	—	−0.23	−0.20	−0.31	NA
Tarsus length	−0.12	−0.10	0.11	—	0.81**	NA	−0.07
PC1-bones	0.02	0.19	−0.03	0.60*	—	0.84	0.05
PC1-flats	0.09	NA	NA	0.07	0.14	—	0.00
PC2-flats	NA	0.20	−0.16	NA	0.38	0.00	—

[a] Blackbird sample sizes are in parentheses; crossbill sample sizes are 36 males, 32 females; sparrow sample sizes are 16 males, 17 females.

[b] For males, PC1-flats is "body size"; for females, PC1-flats is "wing/tail" and PC2-flats is "mass/tarsus."

[c] For males PC1-flats is "wing/tail" and PC2-flats is "mass/tarsus"; for females, PC1-flats is "mass/tail/wing" and PC2-flats is "tarsus."

[d] For males PC1-flats is "wing/tail" and PC2-flats is "mass/tarsus"; for females, PC1-flats is "mass/tarsus" and PC2-flats is "wing/tail."

TABLE 5. Stepwise regressions with PC1-bones as the dependent variable and external metrics as predictor.

Sex	Predictor variables	1st to enter	r^2	2nd to enter	r^2
		Red-winged Blackbirds			
Male	Flat variables: mass, wing, tail, PC1-flat, tarsus	Tarsus	0.53	Mass	0.60
Female	Flat variables: mass, wing, tail, tarsus	Tarsus	0.35	Wing	0.50
Female	PCA variables from flats: tarsus/mass, wing/tail	Tarsus/Mass	0.32	Wing/Tail	0.52
		Red Crossbills			
Male	Flat variables: mass, wing, tail, tarsus	Tarsus	0.32	none	—
Male	PCA variables from flats: tarsus/mass, wing/tail	Tarsus/Mass	0.36	none	—
Female	Flat variables: mass, tarsus, wing, tail	Mass	0.23	none	—
Female	PCA variables from flats: mass/wing/tail, tarsus	Mass/Wing/Tail	0.33	none	—
		Harris' Sparrows			
Male	Flat variables: mass, tarsus, wing, tail	Tarsus	0.35	none	—
Male	PCA variables from flats: mass/tarsus, wing/tail	none	—	none	—
Female	Flat variables: mass, tarsus, wing, tail	Tarsus	0.75	none	—
Female	PCA variables from flats: tarsus/mass, wing/tail	Tarsus/Mass	0.71	none	—

TABLE 6. Multiple regressions with PC1-bones as the dependent variable and external metrics as independent variables.

Sex	Independent variables	r^2	Var	P	Var	P
			Red-winged Blackbirds			
Male	Flat variables: mass, wing, tail, tarsus	0.64	Tarsus	0.0001	Wing	0.02
Female	Flat variables: mass, wing, tail, tarsus	0.54	Tarsus	0.0010	Wing	0.07
Female	PCA-flat variables	0.48	Tarsus/Mass	0.0001	Wing/Tail	0.001
			Red Crossbills			
Male	Flat variables: mass, wing, tail, tarsus	0.43	Tarsus	0.0030	none	—
Male	PCA-flat variables	0.38	Tarsus/Mass	0.0001	none	—
Female	Flat variables: mass, wing, tail, tarsus	0.40	Tarsus	0.0004	none	—
Female	PCA variables	0.38	Mass/Tail/Wing	0.0004	none	—
			Harris' Sparrows			
Male	Flat variables: mass, wing, tail, tarsus	0.46	Tarsus	0.0100	none	—
Male	PCA variables	0.40	none	—		
Female	Flat variables: mass, wing, tail, tarsus	0.81	Tarsus	0.0002	none	—
Female	PCA variables	0.71	Tarsus/Mass	0.0001	none	—

Multiple regressions.—The multiple regressions (Table 6) indicate how much of the overall variance in PC1-bones can be explained by external measurements. Although the four flat-skin elements explain 81% of the variation in PC1-bones for female Harris' Sparrows, most of the regressions (Table 6) show that external measures can explain only 40–60% of the variation in PC1-bones. These results suggest that field-workers should exercise caution in interpreting any one external variable as a reliable predictor of overall body size.

DISCUSSION

The essence of the body-size metric problem in ornithology is that field-workers often need to know an individual's mass at a time when it is impossible to accurately measure mass. The best, and often the only, alternative is to measure an external morphological element that the investigator thinks correlates well with the "true" body size of an individual and hope that this external measure serves as a good general predictor of ranks in average body mass among individuals, if not of actual mass at the specific time in question.

Clearly, having a reliable and easy-to-measure predictor of average body mass would be an enormous help to ornithologists. We have shown that wing chord, a commonly used predictor of average mass, is probably not the best metric to use within populations of small passerines, although it may be a meaningful index for inter-populational studies (James 1970; but note that, for a single population of Downy Woodpeckers (*Picoides pubescens*) at one locality, James obtained an r^2 of only 0.062 when body mass was regressed on wing length). Mass and tarsus length, or a principal component factor which combines the two, are clearly better indices of overall size in the taxa we investigated. The handful of studies to date also seem to indicate that tarsus length and mass are somewhat more heritable than wing length (Hailman 1986). Murphy (1986) and Rising (1987) have published data which also suggest that wing length is a poor measure of overall size in small birds. The proposal (F. Götmark pers. comm.) that wing and tail length serve as better indices of age and wear (activity) than of overall size deserves investigation.

One of the central messages of our findings is an empirical demonstration of the need for caution in interpreting univariate metrics of body size. Although the temptation is strong to draw sweeping conclusions about the effect of body size on contest outcomes or energetics based on patterns (or lack of patterns) observed with univariate body-size metrics, such conclusions might often be in error. We urge field-workers to obtain repeated measures of body mass, or a large series of external measurements of different body components, before making strong conclusions about selection on body size.

ACKNOWLEDGMENTS

We thank Nancy Langston for helping to measure female Red-winged Blackbirds, and David Gori, Frank

Götmark, R. F. Johnston, J. C. Rice, and Sievert Rohwer for commenting on the manuscript.

LITERATURE CITED

BOAG, P. T. 1983. The heritability of external morphology in Darwin's ground finches (*Geospiza*) on Isla Daphne Major, Galápagos. Evolution 37: 877-894.

HAILMAN, J. P. 1986. The heritability concept applied to wild birds. Pp. 71-95 *in* Current ornithology, vol. 4 (R. F. Johnston, Ed.). New York, Plenum.

JAMES, F. C. 1970. Geographic size variation in birds and its relationship to climate. Ecology 51: 365-390.

MURPHY, M. T. 1986. Body size and condition, timing of breeding, and aspects of egg production in Eastern Kingbirds. Auk 103: 465-476.

NILES, D. M. 1973. Adaptive variation in body size and skeletal proportions of Horned Larks of the southwestern United States. Evolution 27: 405-426.

RISING, J. D. 1987. Geographic variation of sexual dimorphism in size of Savannah Sparrows (*Passerculus sandwichensis*): a test of hypotheses. Evolution 41: 514-524.

ROBINS, J. D., & G. D. SCHNELL. 1971. Skeletal analysis of the *Ammodramus-Ammospiza* grassland sparrow complex: a numerical taxonomic study. Auk 88: 567-590.

ROHWER, S. 1972. A multivariate assessment of interbreeding between the meadowlark, *Sturnella*. Syst. Zool. 21: 313-338.

———, P. W. EWALD, & F. C. ROHWER. 1981. Variation in size, appearance, and dominance within and among sex and age classes of Harris' Sparrows. J. Field Ornithol. 52: 291-303.

SCHLUTER, D. 1984. Morphological and phylogenetic relations among the Darwin's finches. Evolution 38: 921-930.

ZINK, R. M. 1982. Patterns of genic and morphologic variation among sparrows in the genera *Zonotrichia, Melospiza, Junco,* and *Passerella*. Auk 99: 632-649.