ORNITOLOGIA NEOTROPICAL 9: 23–30, 1998


# SUGGESTIONS FOR MEASURING EXTERNAL CHARACTERS OF BIRDS

**Kevin Winker**

University of Alaska Museum, 907 Yukon Drive, Fairbanks, Alaska, 99775, U.S.A.
E-mail:ffksw@uaf.edu.

*Abstract.* It is noted that there are no modern references thoroughly describing methods for measuring the external characteristics of birds. Perhaps this explains an all too frequent use of methods that vary from suboptimal to inappropriate. Several common problems are addressed and functional solutions and methods are suggested for measuring external body components and body mass in birds. Some studies from a diverse array of avian investigations are cited to serve as model examples of effectively measuring external characters and of using measurement data.
*Accepted 18 December 1997.*

*Key words: Measurement, morphometrics, techniques, methodology, statistics, comparative method.*

## INTRODUCTION

Recently, I have seen a number of methods used or suggested for the measurement of external body characters and body mass in birds that were not appropriate. Professional ornithologists are usually aware of proper methodologies, or, if not, they are usually aware of examples in the literature where sound methods can be found. However, there are no modern sources where one can find avian measurement thoroughly discussed, making it all too easy for beginning researchers to adopt less suitable or wholly inappropriate methods.

A complete review of the measure and mismeasure of birds is needed, but such a review is not the goal of this paper. Instead, I discuss some seemingly common problems in the measurement of external characters and offer suggestions for functional solutions. Intended as a brief introduction to the subject, this paper will be most useful to researchers who have little experience with the measurement of birds, whether in the field or museum. An array of internal body components, particularly skeletal, are regularly measured in museum studies, but these will not be treated here. Some examples of these measurements and their use can be found in Schnell (1970), Robins & Schnell (1971), Zink (1983), and Rising (1987).

In this paper I cite studies that can serve as model examples of appropriate methods. The guidelines offered are not exhaustive, but have proven satisfactory for my own work in the measurement of well over ten thousand birds in the field and museum. I hope these suggestions will prove useful to others who desire increased value and acceptability of their data.

## LENGTH OF EXTERNAL BODY COMPONENTS

It is inappropriate to measure the lengths of the commonly measured external body components to accuracies of ± 0.01 mm. Although many modern calipers can report such precise values, these measurements are simply not replicable when measuring external body components. With practice, precision calipers can be used to obtain replicable measurements of external characters to accuracies of ± 0.1 mm.

To reduce observer error, one should practice measuring the components to be examined on a small series of birds (e.g., 10 individuals) until accurate replication is achieved. Occasional remeasurement of this reference series is also a good practice. Remeasurement allows one to be sure of consistency over time (e.g., Boag 1983, Prescott 1994), and may show that some characters are not suitable for morphometric analysis (e.g., Zink 1983). If using dial calipers, be aware that the jaws of these instruments must often be completely closed to verify that the closed distance is indeed still zero. Slippage and skipping are not uncommon with these instruments, particularly with increased instrument age, corrosion, and the accumulation of dust and dirt. For this reason, vernier calipers are preferred by some workers.

Electronic calipers linked to a computer for directly recording data into a computer database have become popular for studies involving large numbers of specimens or characters (e.g., Marcus 1982). This can be an improvement for unequivocal characters (e.g., skeletal studies). However, I remain unconvinced of their utility for measuring external characters of birds in museum studies because of the ease of overlooking specimen-related artifacts during measurement. Manually recording each measured length constitutes immediate datum review, and through such review I often discover that outliers are due to aberrancies of the specimen (e.g., hidden shot damage, poorly positioned or everted wings or tarsometatarsi, or previously unobserved molt). Biological outliers are a reality, but must be checked while the specimen is at hand. Electronic calipers can make it all too easy to include artifacts as real data. When using these tools, one should review recorded measurements immediately.

For components that are difficult or impossible to measure directly with calipers (e.g., tarsometatarsus, or "tarsus," on some museum specimens), it is possible to first establish this distance with dividers and measure the distance between divider tips (see illustrations in Baldwin *et al.* 1931). This is not preferred, because it introduces another source of error (taking secondary measurements), but it can enable one to obtain measurements that are not possible with calipers alone. For consistency, however, once dividers have been adopted for a particular measure, they should be applied to all homologous measurements in the study.

It is important to take standard measurements so that results are comparable among studies and data can be replicated by other workers. The seminal work by Baldwin *et al.* (1931) explained and depicted the most widely used standards for external measurements in ornithology. Unfortunately, many of these measurements are now either obsolete or only very rarely used. Because they are not very replicable, measurements of the total length and wingspan of birds have generally been abandoned nowadays (except in special cases, such as studies of the energetics of flight). The most widely used measures today include the lengths of wing chord, tail, tarsometatarsus (usually called "tarsus"), and bill. Many other external measurements can be taken, however, and details will vary with the species and the nature of the research (e.g., see Winkler 1988).

*Bill.* Several types of bill measurements can be found in the modern literature. The most common and replicable is the bill length measured from the anterior edge of the nostril to the tip (Baldwin *et al.* 1931: 16). This measurement is popular because in most taxa both ends of the measurement are readily definable points where calipers can be placed. Two other methods, the lengths of total culmen and exposed culmen, attempt to measure the entire length of the bill. Total culmen measurements extend to the bill tip from the notch on the forehead where the base of the culmen meets the skull (usually just inside the feathers on the forehead; Baldwin *et al.* 1931: 11). Exposed culmen measurements extend to the bill tip from the point where the tips of the forehead feathers begin to hide the culmen (Baldwin *et al.* 1931: 11). This is not an easily defined point, and hence this is the most variable of these bill measurements (pers. observ.).

In taxa without clearly defined anterior nostril edges, total culmen should be preferred; it is usually easy to find the notch where the base of the culmen meets the skull. Hummingbirds present a special problem because of nasal opercula covering the nostrils. While it is usually easy to measure bill from anterior edge of nostril in the field in hummingbirds by inserting caliper tips, this is not possible in museum specimens, where the opercula are dried and hard. In this case, measuring from the base of the nasal opercula is useful in some taxa (e.g., Hinkelmann 1996), whereas total culmen should be preferred in taxa with a less well defined opercular edge. Among those working with passerines, bill from anterior edge of nostril is often referred to as just "bill." But there are multiple ways to measure the bill, and detailed methods should be provided for every study.

*Live birds versus museum specimens.* Because museum specimens shrink upon drying, direct comparison between live birds and prepared specimens must be made with extreme caution (see Winker 1993, 1996). The lengths of individual primaries have been shown to be immune from shrinkage problems (Jenni & Winkler 1989). Measurements of primaries have the added advantage of being obtainable from various museum preparations where wing chord values are of dubious value (e.g., open wings, or skins with "wrists" incompletely everted). Further, primary lengths can add information (such as shape) that complements wing chord values (unpubl. data).

The care and practice required to achieve highly accurate (e.g., ± 0.1 mm) caliper measurements of wing and tail lengths in museum specimens is often unwarranted for measuring birds in the field, given the nature of the research and the ever-present need for speed in handling live birds (although with practice one can achieve equal speed with calipers). For these reasons, a flat ruler is often used: one with a vertical stop at zero for measuring wing length and one with a narrow edge for insertion between the central rectrices for measuring tail length. With these instruments, precision to the nearest ± 0.5 mm can generally be attained. When measuring wings, the chord is the more commonly used measurement (rather than flattened length), and in my experience is the most replicable (although see Parkes 1988). Workers should be aware of the subconscious tendency to read lengths to the nearest line on the ruler: if your ruler is marked only to the single mm, you may subconsciously record fewer values to the half mm than actually exist (e.g., see Mueller 1990). A ruler marked to the half mm and careful examination of data to determine whether approximately half of the measurements taken end in "x.5" should enable one to report accuracies to ± 0.5 mm (see Mueller 1990 for an example of data analysis).

*Wing length.* Flat rulers are generally inappropriate for the direct measurement of wing lengths on most museum specimens (excepting large birds), because one risks damaging the specimen in moving the wings to place them on the ruler properly. Calipers should be used with smaller specimens, and are preferable for measuring tail lengths as well. Flat rulers with vertical stops at zero can be used successfully for measurements of both tail and wing (with care) in large specimens (e.g., hawks and larger birds).

*Size.* A single measurement taken to represent individual size (e.g., wing chord) is not sufficient for careful studies. Single characters rarely correlate well with other size-related characters (e.g., mass) among individuals of a single species or among closely related species. What Arnold & Wade (1984) called "single-trait myopia" is widespread among studies including data on avian size and form. Multivariate approaches are preferable (see also Rising & Somers 1989, Freeman & Jackson 1990). Tarsometatarsus ("tarsus") and tibiotarsus lengths have been suggested as reasonable univariate size estimates based on analysis of four passerine species (Rising & Somers 1989, Freeman & Jackson 1990). However, tarsus length alone would have been misleading in a study of sexual size dimorphism in several neotropical passerines (Winker *et al.* 1994). I have also found tarsus length to be much less useful than wing or tail lengths in predicting the fat content of individual Tennessee Warblers (*Vermivora peregrina*; unpubl. data). Rather than begin using a new set of univariate characters to estimate size, it is better to avoid using univariate measures as size estimates and instead use multivariate approaches. Rising & Somers (1989) gave an excellent discussion of this problem and offered several solutions.

Some generally good examples of making and analyzing external avian measurements, and of the overall utility of these data in avian studies, can be found in Abbott *et al.* (1977), Smith & Zach (1979), Boag (1983, 1984), Blondel *et al.* (1984), Price (1984), Schluter (1984), Schluter & Smith (1986), Aldrich & James (1991), Grant & Grant (1994), and Graves (1996). Useful treatments of morphometrics can be found in Bookstein *et al.* (1985) and Rohlf & Bookstein (1990).

## BODY MASS

Workers should remember that when we "weigh" birds we do so to obtain a measure of their mass (Chardine 1986). Because of its portability, the spring scale has generally replaced the formerly more common triple-beam balance for measuring body mass of birds in the field. Many spring scales are very precise instruments. Using a 30 g Pesola scale, for example, one can theoretically attain an accuracy of ± 0.1 g in the field. At best, these scales are only marked at 0.5 g intervals, however. Larger capacity scales (e.g., 500 g) are marked at intervals of 5 g. Workers must be aware of the subconscious tendency to read these instruments to the nearest mark, and of the need to examine their data to determine whether the level of precision thought possible was indeed obtained (see comments on rulers above, and Mueller 1990). Small portable electronic balances are also becoming common in laboratories and banding stations; calibration of these instruments must be checked. Finally, any scale must be frequently tared.

When making comparisons of body mass between species, or between sexes of the same species, one must critically evaluate published values to determine how and when these values were obtained. Published mass values are useful for some purposes, but can be misleading for others. It must not be forgotten that the body mass of an individual often varies dramatically through the year, and that this variation can be quite different between species and sexes. For example, small migrant landbirds can double their body mass in autumn when depositing fat for migration (Odum *et al.* 1961). Nonmigrant birds are not known to show such dramatic seasonal changes, but will often show lesser seasonal fluctuations (unpubl. data). During the breeding season, females may show substantial mass gains, whereas males often lose mass during the period of territorial defense. Short-term fluctuations in mass also occur, such as during the course of a day due to feeding or egg laying, but these changes are generally ignored except in very fine-scale studies.

When making broad comparisons among species, the primary literature (e.g., Sanft 1970, 1973, Clench & Leberman 1978, Salvador 1988) and sources such as Dunning (1993) are invaluable when used with caution. Care must be used when including these data in a study. In addition, all workers should recognize the futility of separating data by sex in monochromatic species when individuals have not been sexed either by direct gonadal examination or by demonstrably foolproof characters (e.g., sex-specific incubation patches). Also, it should be remembered that even direct gonadal examination can be subject to error (Clench 1976, unpubl. data).

Another example shows that reliance on published mass values can be misleading. Walkinshaw's (1953) data on body mass during the breeding season in the Prothonotary Warbler (*Protonotaria citrea*; repeated in Dunning 1993) suggest that females are 1.16 times larger than males. This should be immediately suspicious to the careful reader, because reverse sexual size dimorphism, while common groups such as Accipitridae, is highly unusual in nine-primaried oscines. Indeed, Walkinshaw's (1953: 166) data on wing lengths suggest that males are larger than females in wing chord, and perhaps larger in "culmen" and "tarsus" (=tarsometatarsus) lengths, as well. Removing the highly variable fat component gives a more realistic view of the mass relationship between males and females in this (and other) species. Mass data from a study using birds from which fat was extracted (Odum, in Dunning 1993) suggests a more credible relationship between male and female *Protonotaria citrea*: males have a slightly higher fat-free mass than females of the same age class (male:female ratio of 1.04; 47 males, 20 females; Dunning 1993: 325). A multivariate examination of ten adults of each sex confirms this relationship: males show significantly higher values than females on the first principal component in a principal components analysis of the log-transformed measures of wing chord, tail, tarsometatarsus, and bill lengths (unpubl. data). In my experience, the data assembled by Dunning (1993), while valuable for many reasons, have too many weaknesses to be used in studies of sexual size dimorphism (at least at the level of family and below). There are many other inappropriate uses of these data, and researchers should justify their use of such sources when reporting their results.

## ACKNOWLEDGMENTS

I thank Gary Graves and Kenneth Parkes for stimulating discussion and comments, and James Wiley, François Vuilleumier, and two anonymous reviewers for suggested improvements.

## REFERENCES

Abbott, I., L. K. Abbott, & P. R. Grant. 1977. Comparative ecology of Galápagos ground finches (*Geospiza* Gould): Evaluation of the importance of floristic diversity and interspecific competition. Ecol. Monogr. 47: 151–184.

Aldrich, J. W., & F. C. James. 1991. Ecogeographic variation in the American Robin (*Turdus* *migratorius*). Auk 108: 230–249.

Arnold, S. J., & M. J. Wade. 1984. On the measurement of natural and sexual selection: applications. Evolution 38: 720–734.

Baldwin, S. P., Oberholser, H. C., & L. G. Worley. 1931. Measurements of birds. Sci. Publ. Clev. Mus. Nat. Hist. 2: 1–165.

Blondel, J., F. Vuilleumier, L. F. Marcus, & E. Terouanne. 1984. Is there morphological convergence among Mediterranean bird communities of Chile, California, and France? Evol. Biol. 18: 141–213.

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

Boag, P. T. 1984. Growth and allometry of external morphology in Darwin's finches (*Geospiza*) on Isla Daphne Major, Galapagos. J. Zool. London 204: 413–441.

Bookstein, F., B. Chernoff, R. Elder, J. Humphries, G. Smith, & R. Strauss. 1985. Morphometrics in evolutionary biology. Acad. Nat. Sci. Phil. Spec. Publ. 15: 1–277.

Chardine, 1986. Mass or weight: what is measured and what should be reported? Auk 103: 832.

Clench, M. H. 1976. Possible pitfalls in museum specimen data. North Am. Bird Bander 1: 20–21.

Clench, M. H. & R. C. Leberman. 1978. Weights of 151 species of Pennsylvania birds analyzed by month, age, and sex. Carnegie Mus. Nat. Hist. Bull. 5: 1–87.

Dunning, J. B., Jr. 1993. CRC handbook of avian body masses. CRC Press, Boca Raton, Florida.

Freeman, S., & W. M. Jackson. 1990. Univariate metrics are not adequate to measure avian body size. Auk 107: 69–74.

Grant, P. R., & B. R. Grant. 1994. Phenotypic and genetic effects of hybridization in Darwin's finches. Evolution 48: 297–316.

Graves, G. R. 1996. Hybrid wood warblers, *Dendroica striata* x *Dendroica castanea* (Aves: Fringillidae: Tribe Parulini) and the diagnostic predictability of avian hybrid phenotypes. Proc. Biol. Soc. Wash. 109: 373–390.

Hinkelmann, C. 1996. Systematics and geographic variation in long-tailed hermit hummingbirds, the *Phaethornis superciliosus-malaris-longirostris* species group (Trochilidae), with notes on their biogeography. Ornitol. Neotrop. 7: 119–148.

Jenni, L., & R. Winkler. 1989. The feather-length of small passerines: a measurement for wing-length in live birds and museum skins. Bird Study 36: 1–15.

Marcus, L. 1982. A portable, digital, printing caliper. Curator 25: 223–226.

Mueller, H. C. 1990. Can Saw-whet Owls be sexed by external measurements? J. Field Ornithol. 61: 339–346.

Odum, E. P., C. E. Connell, & H. L. Stoddard. 1961. Flight energy and estimated flight ranges of some migratory birds. Auk 78: 515–527.

Parkes, K. C. 1988. [Review of] Identification guide to North American passerines. Auk 105: 599.

Prescott, D. R. C. 1994. Intraspecific and geographical trends in body size of a differential migrant, the Evening Grosbeak. Auk 111: 693–702.

Price, T. D. 1984. The evolution of sexual size dimorphism in Darwin's finches. Am. Nat. 123: 500–518.

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

Rising, J. D., & K. M. Somers. 1989. The measurement of overall body size in birds. Auk 106: 666–674.

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

Rohlf, F. J., & F. L. Bookstein, eds. 1990. Proceedings of the Michigan morphometrics workshop. Univ. Mich. Mus. Zool. Spec. Publ. 2: 1–380.

Salvador, S. A. 1988. Datos de peso de aves Argentinas. Hornero 13: 78–83.

Sanft, K. 1970. Gewichte südamerikanischer Vögel-*Nonpasseres*. Beitr. Vogelkd. 16: 344–354.

Sanft, K. 1973. Gewichte südamerikanischer Vögel-*Passeres*. Beitr. Vogelkd. 19: 406–423.

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

Schluter, D., & J. N. M. Smith. 1986. Natural selection on beak and body size in the Song Sparrow. Evolution 40: 221–231.

Schnell, G. D. 1970. A phenetic study of the suborder Lari (Aves) I. Methods and results of principal components analyses. Syst. Zool. 19: 35–57.

Smith, J. N. M., & R. Zach. 1979. Heritability of some morphological characters in a Song Sparrow population. Evolution 33: 460–467.

Walkinshaw, L. H. 1953. Life-history of the Prothonotary Warbler. Wilson Bull. 65: 152–168.

Winker, K. 1993. Specimen shrinkage in Tennessee Warblers and "Traill's" Flycatchers. J. Field Ornithol. 64: 331–336.

Winker, K. 1996. Specimen shrinkage versus evolution: I'iwi morphology. Cons. Biol. 10: 657–658.

Winker, K., G. A. Voelker, & J. T. Klicka. 1994. A morphometric examination of sexual dimorphism in the *Hylophilus, Xenops,* and an *Automolus* from southern Veracruz, México. J. Field Ornithol. 65: 307–323.

Winkler, H. 1988. An examination of concepts and methods in ecomorphology. Acta Congr. Int. Ornithol. 19: 2246–2253.

Zink, R. M. 1983. Evolutionary and systematic significance of temporal variation in the Fox Sparrow. Syst. Zool. 32: 223–238.