

DIAGNOSABILITY VERSUS MEAN DIFFERENCES OF SAGE SPARROW SUBSPECIES

MICHAEL A. PATTEN^{1,2,3} AND PHILIP UNITT²

¹Department of Biology, University of California, Riverside, California 92521, USA; and
²San Diego Natural History Museum, P.O. Box 121390, San Diego, California 92112, USA

ABSTRACT.—A subspecies is a collection of populations within a biological species that are diagnosably distinct from other such collections of populations. That infraspecific designation has motivated a litany of spirited debates over the past half-century, from impassioned pleas for its retention to heated outcries for its abolition. We believe that the vast majority of attacks on the subspecies concept have resulted from displeasure with its improper application, not from serious flaws in the concept itself. The recognition of diagnosable subspecies allows one to address many questions not easily answered otherwise, ranging from dispersal and migration to local selection and adaptation and biogeographic affinities, yet that goal was lost for many years. Many taxonomists in the late nineteenth century and first half of the twentieth century named subspecies on the basis of average differences between populations under study, a procedure at odds with identification of diagnosable populations. To resolve that dilemma, we make explicit the established 75% rule for subspecies recognition, including formalizing the rule and developing a simple statistic to test whether diagnosability is met. The equations can be adapted readily to any level of diagnosability. We apply the concept and the statistic to a revision of the subspecies of the Sage Sparrow (*Amphispiza belli*). Rather than the seven named subspecies or the five that are generally considered valid, we show that only three aggregates of populations are diagnosable, and thus only three subspecies should be recognized: (1) *A. b. belli* in chaparral and sage scrub of coastal California, northwestern Baja California, and San Clemente Island; (2) *A. b. cinerea* in desert scrub of west-central Baja California; and (3) *A. b. nevadensis* in sagebrush and saltbush of the Great Basin and interior California. Consistent application of the 75% rule will result in fewer trinomials and a more biologically meaningful and taxonomically useful subspecies concept. Received 19 September 2000, accepted 11 June 2001.

RESUMEN.—Una subespecie es un grupo de poblaciones que forman parte de una especie biológica y que tienen características diagnósticas que las separan de otros grupos de poblaciones de ese tipo. Esta designación infraespecífica ha motivado una letanía de debates acalorados durante el último medio siglo que van desde súplicas apasionadas por mantenerla hasta llamados a gritos por abolirla. Nosotros creemos que la gran mayoría de ataques al concepto de subespecie han sido el resultado de inconformidad causada por su aplicación inadecuada y no de problemas serios del concepto mismo. El reconocimiento de subespecies diagnosticables permite abordar muchas preguntas que no serían contestables fácilmente de otra manera. Éstas van desde cuestiones relacionadas con dispersión y migración hasta selección local, adaptación y afinidades biogeográficas. Sin embargo, esto no fue posible por muchos años pues muchos taxónomos de finales del siglo diecinueve y principios del siglo veinte nombraron subespecies con base en diferencias promedio de las poblaciones de estudio, un procedimiento que es incompatible con la identificación de poblaciones diagnosticables. Para resolver ese dilema, hacemos explícita la regla del 75% para el reconocimiento de subespecies, incluyendo su formalización y el desarrollo de un estadístico simple para determinar si hay o no diagnosticabilidad. Las ecuaciones se pueden adaptar fácilmente a cualquier nivel de diagnosticabilidad. Nosotros aplicamos el concepto y el estadístico al hacer una revisión de las subespecies de *Amphispiza belli*. En lugar de las siete subespecies nombradas o de las cinco que generalmente se consideran válidas, demostramos que sólo tres grupos de poblaciones son diagnosticables. Por tanto, sólo deberían reconocerse tres

³ Present address: Environmental Studies Program, Dartmouth College, Hanover, New Hampshire 03755, USA. E-mail: michael.a.patten@dartmouth.edu

subspecies: (1) *A. b. belli* del chaparral y matorral de *Artemisia* de la costa de California, el noroeste de Baja California y la Isla San Clemente; (2) *A. b. cinerea* del matorral desértico del oeste-centro de Baja California; y (3) *A. b. nevadensis* de los matorrales de *Artemisia* y de *Atriplex* del Great Basin y el interior de California. La aplicación consistente de la regla del 75% conducirá a la existencia de menos trinomios y a un concepto de subespecie con un mayor significado biológico y una mayor utilidad taxonómica.

THE SUBSPECIES CONCEPT

THE CONCEPT OF subspecies has been widely criticized in the past half century, beginning with the well-known critiques by Wilson and Brown (1953) and Inger (1961). There is no doubt that subspecific taxonomy has been fraught with problems, but we wish to make an important distinction. Most criticism leveled at the subspecies concept is not properly directed at the concept at all but rather should have been directed at its frequent misapplication. The host of invalid subspecies names has resulted from a lack of appreciation of just what a subspecific taxon should be. Rather than quantifying the diagnosability of a population and thus the predictability of assignment of individual specimens to that population (Barrowclough 1982), many subspecies have been named on mean differences only. The latter practice led to naming of subspecies even along perfectly smooth clines, where in principle a near limitless number could be named provided sample sizes from each point were extremely large. Sadly, such careless naming was rife for decades, compounded by names based on artifacts like the effects of abrasion or post-mortem color changes.

Despite those strong attacks, various prominent systematists continued to champion the subspecies concept (e.g. Mayr 1954, Smith and White 1956, Simpson 1961, Parkes 1982). In essence, a subspecies is a collection of populations occupying a distinct breeding range and diagnosably distinct from other such populations (Mayr and Ashlock 1991). Proper description of a subspecies allows for a myriad of important studies in patterns of geographic variation, adaptation, biogeography, migration, and speciation. Indeed, the standard paradigm of speciation (Mayr 1942) states that all species must go through a subspecies stage, making the concept one of evolutionary import. It is therefore imperative that taxonomists and systematists employ a quantifiable definition of subspecies.

DIAGNOSABILITY VERSUS MEAN DIFFERENCES

Since the criticisms of a half-century ago, a good deal of effort has been expended on developing statistical procedures for describing geographic variation and drawing taxonomic boundaries (e.g. Pimentel 1958, 1959; Gabriel and Sokal 1969, Sneath and Sokal 1973, Thorpe 1976). Unfortunately, these procedures focused nearly without fail on detecting significant mean differences between populations, not on determining extent of overlap.

The standard level for defining a subspecies is based on the "75% rule" (Amadon 1949, Mayr 1969). Stated simply, to be a valid subspecies 75% of a population effectively must lie outside 99% of the range of other populations for a given defining character or set of characters. More stringent criteria of 90% separation (Amadon and Short 1992) and even 100% separation (Marshall 1967) have been advocated. The latter criterion is tenable only if a biological concept of species is used; otherwise, the diagnosis of a subspecies is equivalent to the diagnosis of a phylogenetic species (*sensu* Cracraft 1983). Although phylogenetic species are postulated to correspond to monophyletic units, that correspondence is seldom if ever demonstrated. Such correspondence does not necessarily follow from evolutionary theory (Avice 2000), so the definition of species on the basis of diagnosability alone is tenuous. Supposed phylogenetic species are often, in our view, good subspecies, not good species. The subspecies concept we advocate—diagnosable clusters of populations of biological species occupying distinct geographic ranges—is inconsistent with a typological, "phylogenetic" species concept. Concealed by the nesting of the two categories in the Linnaean hierarchy is that they address qualitatively different aspects of biology: the species addresses reproductive and behavioral criteria, the subspecies morphological diagnosability. We advocate use of a higher level of subspecies diagnosability, with 95% being an obvious choice if only for paral-

lelism with the standard of $\alpha = 0.05$ in statistical tests, but we acknowledge that the 75% rule is the most firmly established in the literature, and we find it serves well for explaining our technique and presenting our example. The technique we develop below, however, is amenable to any level of separation desired.

Despite claims that the 75% rule was strictly followed, most taxonomists nevertheless employed a looser rule of mean differences (Rand and Traylor 1950). A valid subspecies should be diagnosably different from all other populations, not merely exhibit mean differences. Otherwise, individuals cannot be identified, predictability is lost, and the category is deprived of its most useful applications. Admittedly, the lower boundary for defining a valid diagnosable subspecies is arbitrary. That arbitrariness has been a focal point of criticism for the subspecies concept (e.g. Wilson and Brown 1953, Inger 1961, Selander 1971). One must bear in mind, however, that definitions for placement into all taxonomic groups other than a species are arbitrary (presuming some sort of biology-based species concept). Just how much difference is needed to define a genus? What level of distinction is necessary to define a new phylum? Do all genera or phyla differ from each other equally? Even if all such higher categories could be defined cladistically or on the basis of some degree of genetic difference, arbitrary definitions are unavoidable. Also consider that our beloved $\alpha = 0.05$, the underpinning of inferential statistics, is an arbitrary level. We accept a Type I error rate of 5% not because it is superior to any other; it is merely accepted convention.

Even today, otherwise important contributions to geographic variation and subspecific taxonomy place too much importance on mean differences rather than diagnosability. Statistically significant differences from *t*-tests, ANOVA, or various multiple comparison tests do not a subspecies make (*contra* Selander 1971, Bilgin 1999). Likewise, powerful multivariate tests such as discriminant function analysis can tell us how well populations are differentiated on the basis of mean differences of multiple characters considered simultaneously, but interpretation of results from such tests needs to be more rigorous than the simple reporting of percentage correct classification (e.g. Unitt et al. 1996).

Numerous authors, failing to embrace fully the notion that named subspecies should be diagnosable even at this arbitrary 75% (or higher) level, is the equivalent of numerous workers publishing results lacking significance at $\alpha = 0.05$. The latter problem was corrected decades ago through strict adherence to the "5% Type I error rate rule." It is time a strict rule of subspecies diagnosability was enforced with equal vigor. Subspecies diagnoses need to be quantitative and statistically justified to the fullest extent possible. After all, as astutely noted by Michener (1963), owing to the rule of priority in nomenclature, taxonomy is the only scientific endeavor in which careless work is preserved in perpetuity.

FORMALIZING THE QUANTIFICATION OF SUBSPECIES

With the 75% rule definition, it is possible to quantify the diagnosis of a subspecies with statistical precision. The task is simple for categorical characters, such as the presence or absence of a color or marking. Classifying specimens accordingly, with the character treated as either present or absent, quickly yields a 2×2 table that can be tested with chi-square statistics or, more simply, counted to determine if the desired level of diagnosability has been achieved. For continuously varying characters, we assume (as do virtually all statistics) that the resultant phenotype is normally distributed. Under this assumption, the 75% rule reduces to meaning that at least 75% of the distribution of the population under study lies outside the distribution of the population of the reference population (Fig. 1). Assuming a standard normal distribution,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-[(x-\mu)^2/2\sigma^2]}$$

the distribution of a given trait, x is

$$f(x) = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-[(x-\mu)^2/2\sigma^2]} dx$$

where μ is the population mean and σ is the population standard deviation. Thus, in terms of parameters for population 2, the 75% rule can be expressed as

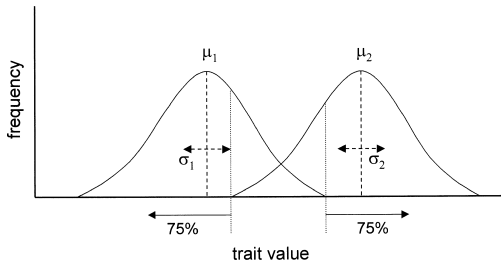


FIG. 1. A graphical representation of the 75% rule showing the distributions of a measured trait in two populations (assuming normality). If a population is valid subspecies then at least 75% of its distribution will lie outside the range of other populations.

$$0.75 \leq \frac{\int_b^d \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-[(x-\mu_2)^2/2\sigma_2^2]} dx}{\int_c^d \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-[(x-\mu_2)^2/2\sigma_2^2]} dx}$$

which reduces to

$$0.75 \leq \frac{\int_b^d e^{-[(x-\mu_2)^2/2\sigma_2^2]} dx}{\int_c^d e^{-[(x-\mu_2)^2/2\sigma_2^2]} dx}$$

where μ_1 and μ_2 are the means for populations 1 and 2, respectively, and σ_1 and σ_2 are the associated standard deviations. Limits to the integrals are from Figure 1, with $a \approx \mu_1 - 2.33\sigma_1$, $b \approx \mu_1 + 2.33\sigma_1$, $c \approx \mu_2 - 2.33\sigma_2$, and $d \approx \mu_2 + 2.33\sigma_2$ (the constant is from a standard Z table, for 99.9% of the area under the curve).

The integral can be solved using standard u substitution, with $u = (x - \mu)/\sigma\sqrt{2}$ and $du = 1/\sigma\sqrt{2} dx$ Unfortunately, the solution to

$$\frac{1}{\sigma\sqrt{2}} \int e^{-u^2} du$$

is complex. From Gröbner and Hofreiter (1961),

$$\int e^{-u^2} du = \frac{\sqrt{\pi}}{2} \Phi + C$$

where

$$\Phi = \frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt$$

which in turn has a complex, iterative solution of

$$\frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt = \frac{2u}{\sqrt{\pi}} \left[1 - \sum_{k=1}^{\infty} \frac{(-1)^k u^{2k}}{k!(2k+1)} \right]$$

(from Dwight 1961). Substituting the appropriate u (in this case for population 2), the solution becomes

$$\frac{x - \mu_2}{2\sigma_2^2} \left(1 - \frac{\left(\frac{x - \mu_2}{\sigma_2\sqrt{2}}\right)^2}{1!(3)} + \frac{\left(\frac{x - \mu_2}{\sigma_2\sqrt{2}}\right)^4}{2!(5)} - \frac{\left(\frac{x - \mu_2}{\sigma_2\sqrt{2}}\right)^6}{3!(7)} + \frac{\left(\frac{x - \mu_2}{\sigma_2\sqrt{2}}\right)^8}{4!(9)} - \frac{\left(\frac{x - \mu_2}{\sigma_2\sqrt{2}}\right)^{10}}{5!(11)} \dots \right)$$

which becomes more complex once evaluated at the appropriate limits, with a similar term in the numerator and denominator. Thus, an exact solution to the percentage of population overlap exists, but the equations are rather unwieldy and are not amenable to analyses of samples rather than entire populations.

A more satisfactory solution can be derived from statistical sampling theory. Previous statistical efforts at determining fit to the 75% rule (e.g. Amadon 1949, Unitt 1987) have used a normal distribution with associated Z scores. Such a solution is accurate only when large sample sizes ($n \geq 100$) are used; the t distribution should be used instead. Purely on the basis of the distributions of the populations (Fig. 1), the 75% rule reduces to a boundary problem in which one can use critical values of t at the appropriate one-tailed α levels to determine if the rule is met. For example, to satisfy the rule at least 75% of the distribution of population 2 must lie to the right of 99% of the distribution of population 1 (for the character in question). Thus, if the trait is normally distributed, the rule is

$$\bar{x}_1 + S_1(t_{0.01,df_1}) \leq \bar{x}_2 - S_2(t_{0.25,df_2})$$

where population 2 has the larger sample mean, \bar{x} is the sample's mean, and S is the sample's standard deviation. Note that this equation uses sample means and standard deviations for each population and the appropriate critical values of t . The t values are determined, respectively, for $\alpha = 0.01$ and the degrees of

freedom of population 1 ($df_1 = n_1 - 1$, where n_1 is the sample size for population 1) and $\alpha = 0.25$ and $df_2 (= n_2 - 1)$. With terms rearranged this equation can be expressed as

$$0 \leq \bar{x}_2 - S_2(t_{0.25,df_2}) - \bar{x}_1 - S_1(t_{0.01,df_1})$$

providing, then, a pairwise diagnosability index (D_{ij}) where $D_{ij} \geq 0$ means that population i is diagnosable from population j and $D_{ij} < 0$ means that it is not. This test must be performed in both directions. In our example, population 1 must be diagnosable from population 2 ($D_{12} \geq 0$) and population 2 must be diagnosable from population 1 ($D_{21} \geq 0$). Swapping critical t values in the equation above, for D_{21} , yields the complementary equation for D_{12} . Subspecies based on combinations of several characters can be compared and defined by our technique either by testing variables separately or by combining them, by means of a multivariate procedure such as principal components analysis, into a single variable.

SUBSPECIFIC TAXONOMY OF THE SAGE SPARROW

Seven subspecies of the Sage Sparrow (*Amphispiza belli*) have been named. Two, *A. b. campicola* and *A. b. xerophilus*, were properly synonymized in the literature many years ago. The former is a synonym of *A. b. nevadensis* (Paynter 1970) from which it is essentially indistinguishable (*contra* Behle 1985). The latter is based on one point along the steep cline between the dark *A. b. belli* and pale *A. b. cinerea* (cf. Jehl 1968, Miller et al. 1957). Four or all five of the remaining subspecies are generally recognized (e.g. American Ornithologists' Union 1957, Miller et al. 1957, Rising 1996). Three are resident in cismontane California and Baja California, the nominate in coastal California and northwestern Baja California, *A. b. clementeae* on San Clemente Island, and *A. b. cinerea* in west-central Baja California (Fig. 2). Two are pale interior subspecies, the large *A. b. nevadensis* of the Great Basin and small *A. b. canescens* of central interior of California (Fig. 2). Although it makes little biological sense, it has been widely reported that both pale interior subspecies migrate into the Colorado Desert and southern Arizona in winter (Grinnell and Miller 1944, Phillips et al. 1964, Rea 1983). *A. b. canescens* was defined on the basis of significant

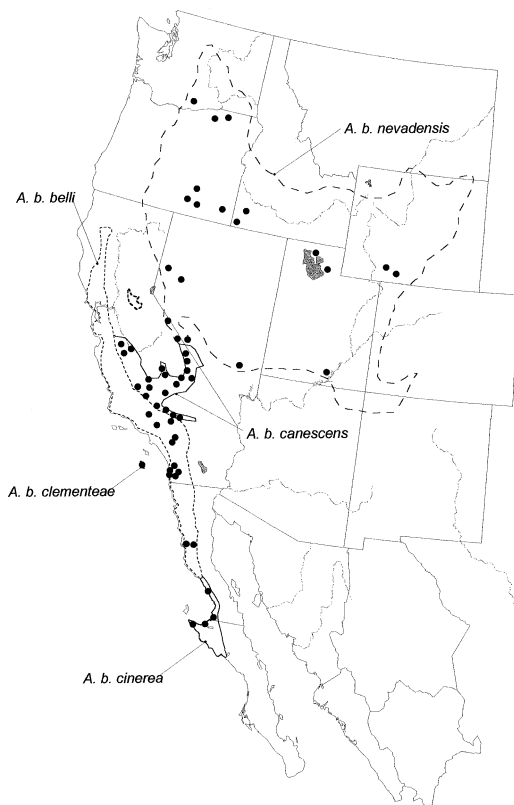


FIG. 2. Ranges of described subspecies of the Sage Sparrow (*Amphispiza belli*) from interior western North America. Dots signify collection locations of specimens examined.

differences in mean mensural characters (Grinnell 1905, Johnson and Marten 1992), particularly wing chord. Males average larger than females throughout the species' range.

METHODS

We assessed validity of each of those five subspecies on the basis of the characters used in their original descriptions. There are three principal sources of variation: (1) size, especially wing length; (2) darkness of the mantle; and (3) degree of streaking on the mantle. We determined size by measuring wing chord of specimens at the Field Museum of Natural History, Chicago (FMNH); Natural History Museum of Los Angeles County, Los Angeles (LACM); Museum of Vertebrate Zoology, University of California, Berkeley (MVZ); San Diego Natural History Museum (SDNHM); and Western Foundation of Vertebrate Zoology, Camarillo, California (WVZ). Although wing chord is partly related to overall size and partly to migratory distance (Zink and Remsen 1986), sub-

TABLE 1. Mantle darkness of described subspecies of the Sage Sparrow. Values are the L metric from a Minolta CR-300 Chroma Meter, with higher values indicating paler coloration. Data are from specimens at SDNHM and LACM.

Named subspecies (range)	Mean (L)	SD	Range	n
<i>A. b. belli</i> (Southwestern California/Northwestern Baja California)	29.8	1.68	27.8–33.2	18
<i>A. b. clementeae</i> (San Clemente Island)	31.0	1.14	28.7–33.7	14
<i>A. b. cinerea</i> (Central Baja California)	37.9	1.35	35.6–40.0	9
<i>A. b. canescens</i> (Central California)	36.6	0.74	35.8–37.2	3
<i>A. b. nevadensis</i> (Great Basin)	37.5	1.88	35.0–41.3	16

species of the Sage Sparrow were named mainly on that metric. *A. b. canescens* was also described in part on tail length, so we included tail measurements in our comparison of it and *A. b. nevadensis*. We measured mantle darkness with a Minolta CR-300 Chroma Meter using fresh specimens (collected September–January), from the breeding range, at SDNHM. The L metric, a unitless measure of darkness, is higher for pale colors and lower for dark colors. We tested for foxing, postmortem reddening, and fading of feathers, by regressing L against year of collection for specimens of *A. b. nevadensis* (for which we had specimens from 1899 to 1996). We assessed mantle streaking as either present or absent. We analyzed measured characters by means of t -tests and, in the case of the data on wing and tail length for *A. b. canescens* versus *A. b. nevadensis*, by means of a discriminant function analysis so that both characters could be considered simultaneously. We assessed diagnosability using the index (D_{ij}) developed above. Other standard museum abbreviations are Academy of Natural Sciences, Philadelphia (ANSP) and National Museum of Natural History, Smithsonian Institution, Washington, D.C. (USNM).

RESULTS

Sage Sparrows show no evidence of foxing ($r = 0.03$, $df = 18$, $P > 0.05$) so correction for specimen age was unnecessary. There is no overlap between mantle darkness of *A. b. belli* and the pale subspecies *A. b. cinerea*, *A. b. nevadensis*, and *A. b. canescens* (Table 1). In that respect alone the subspecies is 100% diagnosable ($D_{ij} > 1.8$ for all pairwise comparisons). There is slight clinal variation in mantle darkness of *A. b. belli*, the birds becoming a bit paler in the southern part of its range (e.g. the vicinity of San Quintin in northwestern Baja California). Even so, individuals as pale or paler can be found in San Diego County, California (e.g. SDNHM 15678, Kearney Mesa, 28 January 1932).

A. b. clementeae is also a dark-mantled subspecies, although it is reportedly paler and shorter billed than *A. b. belli* (van Rossem 1932). We found, however, that mean bill lengths of *A. b. belli* (7.14 mm, $SD \pm 0.26$) and *A. b. clementeae* (7.19 mm, $SD \pm 0.26$) do not differ ($t = 0.45$, $df = 29$, $P > 0.05$). With no significant difference in a character according to a t -test, evaluating it for diagnosability is pointless. By contrast, average mantle darkness of birds from San Clemente Island and from the adjacent mainland of San Diego County does differ significantly ($t = 2.34$, $df = 29$, $P < 0.03$). However, those subspecies are not diagnosable on mantle darkness ($D_{cb} = -3.91$, $D_{bc} = -2.98$), an unsurprising finding given the broad overlap in this character (Table 1).

None of the pale subspecies differ from each other in mantle darkness (Table 1; t -tests, $P > 0.05$ for all pairwise comparisons). *A. b. cinerea* is readily diagnosed by the complete lack of streaking on the mantle feathers (formed by longitudinal dark marks around the shaft of the feathers on the back). No specimens of *A. b. cinerea* have even faint mantle streaks, whereas $>97\%$ (151 of 155) of *A. b. nevadensis* and *A. b. canescens* have bold mantle streaks. The four specimens of *A. b. canescens* from the western part of its range that had only faint mantle streaks approached some *A. b. belli*, which has a faintly streaked mantle. *A. b. cinerea* is further characterized by its smaller average wing chord (Table 2), but its diagnosability rests on absence of mantle streaking ($D_{ij} < -3.5$ for all pairwise within-sex wing chord comparisons to *A. b. canescens*).

The two pale subspecies from the Central Valley, Mojave Desert, and Great Basin do not differ in mantle darkness ($t = 1.48$, $df = 7$, $P > 0.05$) or in extent of mantle streaking ($>97\%$ of both have streaking). Those populations do,

TABLE 2. Wing chord statistics for pale subspecies of the Sage Sparrow (*Amphispiza belli*) described from xeric western North America. Data are from museum specimens housed at LACM, MVZ, SDNHM, and WFVZ. Within sex, mean wing-chord differs significantly between named subspecies (*t* tests, $P < 0.001$ for all comparisons).

Named subspecies (sex)	Mean	SD	Range	<i>n</i>
<i>A. b. cinerea</i> (male)	65.4 mm	3.10	59.1–68.7 mm	13
<i>A. b. canescens</i> (male)	70.9 mm	2.88	65.0–76.5 mm	45
<i>A. b. nevadensis</i> (male)	78.7 mm	2.79	73.0–85.5 mm	38
<i>A. b. cinerea</i> (female)	63.0 mm	2.77	58.6–68.6 mm	12
<i>A. b. canescens</i> (female)	67.2 mm	2.77	63.0–75.0 mm	42
<i>A. b. nevadensis</i> (female)	73.4 mm	2.30	65.0–76.5 mm	30

however, differ significantly in average size, with *A. b. nevadensis* being larger, especially in wing chord (Table 2; Grinnell 1905, Johnson and Marten 1992). Even so, the populations do not meet the 75% rule definition of a subspecies (Fig. 3). Among males, only 66.7% of *A. b. canescens* are smaller than the smallest *A. b. nevadensis*, whereas 73.7% of *A. b. nevadensis* are larger than the largest *A. b. canescens* ($D_{nc} = -1.05$, $D_{cn} = -0.94$). Females fare even worse, with 66.7% of *A. b. canescens* smaller than the

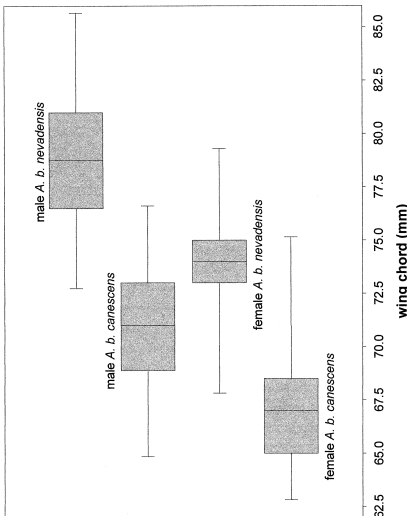


FIG. 3. Wing-chord distributions for subspecies of the Sage Sparrow (*Amphispiza belli*) of interior western North America. Boxes show the median line and the 25th and 75th percentiles; bars show the full range of variation. Despite significant differences in mean wing chord, the subspecies are not diagnosable under the 75% rule (i.e. the 75th percentile does not lie outside the range of the other subspecies). We present this figure as a graphical aid only; see text for the quantitative demonstration of lack of diagnosability.

smallest *A. b. nevadensis* and a mere 10.0% of *A. b. nevadensis* larger than the largest *A. b. canescens* ($D_{nc} = -2.05$, $D_{cn} = -1.35$). Despite reportedly strong genetic differentiation between those populations (Johnson and Marten 1992), they cannot be separated at the minimal acceptable level. The size variation may be clinal. The original description of *A. b. canescens* differentiated the subspecies on both its shorter wing chord and tail (Grinnell 1905). Combining those two measures does not improve diagnosability. There is broad overlap between subspecies for both males and females (Fig. 4); in none of the four sex–subspecies classes is >75% of individuals classified in the correct subspecies with $P \geq 0.9$ confidence.

Notably, some summer specimens from within the range of either subspecies demonstrate the broad overlap in size. For example, two breeding males of *A. b. canescens* from Tulare County, California, 28 July 1911 (MVZ 20461, 20462), have wing chords ≥ 76 mm. Likewise, a breeding female *A. b. "canescens"* from Palmdale, Los Angeles County, California, 28 March 1934 (WFVZ 10080), has a wing chord of 75 mm. Conversely, a female *A. b. nevadensis* from central Mono County, California, 9 September 1917 (MVZ 28408), has a wing chord of only 68 mm; and a male from adjacent Nevada 21 June 1972 (MVZ 163098), measures only 73 mm.

Because tracing migration and dispersal is a key application of subspecies, we tested applicability of the traditional subspecific classification on a set of specimens collected on the species' wintering grounds. As noted above, both *A. b. nevadensis* and *A. b. canescens* allegedly winter together in the Colorado Desert, with the latter reportedly more numerous in the Salton Sink of southeastern California (van

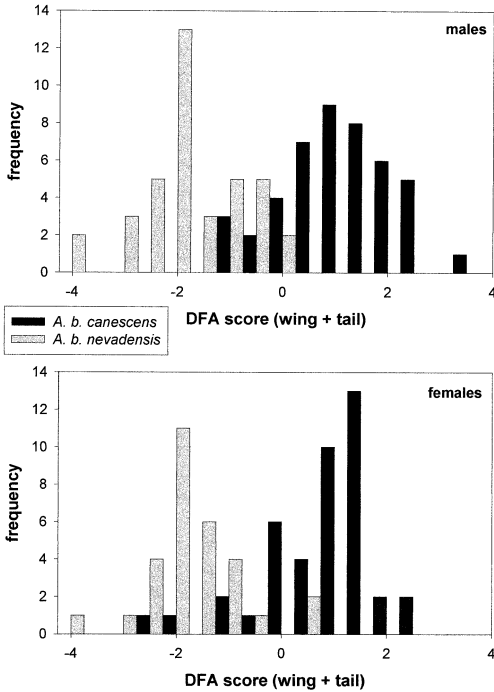


FIG. 4. Scores from discriminant function analyses (DFA) using wing chord and tail length in concert for separating male and female Sage Sparrow specimens. Note the broad overlap in *A. b. canescens* and *A. b. nevadensis*, despite near lack of specimens from the reported intergradation zone.

Rossem 1911). We measured wing chord of 22 specimens from the Salton Sink. Of those, half would be classified as *A. b. canescens*, four as *A. b. nevadensis*, and seven are equivocal. A fine example of the last category is an unsexed bird from Mecca 2 February 1934 (FMNH 127943). Its wing chord is 74 mm, within the size range of females and males of either subspecies! One-third of Sage Sparrow specimens from the Salton Sink not being identifiable implies that the subspecies are too poorly differentiated to be useful.

CONCLUSIONS

Discontinuities of variation in the Sage Sparrow allow recognition of only three diagnosable subspecies:

Amphispiza belli belli (Cassin, 1850)

Emberiza belli Cassin, 1850, Proc. Acad. Nat. Sci. Philadelphia 5:104

Amphispiza belli clementae Ridgway, 1898, Auk 15:230
Amphispiza belli xerophilus Huey, 1930, Trans. San Diego Soc. Nat. Hist. 6:229

Holotype.—ANSP 24036; age and sex?; near Sonoma, Sonoma County, California; date?; J. G. Bell.

Diagnosis.—Mantle dark; mantle streaks distinct, though muted; size averages small relative to that of populations in the Mojave Desert and the Great Basin.

Range.—The coastal slope of central California south through northwestern Baja California and on San Clemente Island off southern California; sedentary. Intergrades with *A. b. cinerea* at Santa Catarina Landing, Baja California.

Amphispiza belli nevadensis (Ridgway, 1873)

Poospiza belli nevadensis Ridgway, 1873, Bull. Essex Inst. 5:191

Aimophila belli canescens Grinnell, 1905, Condor 7:18
Aimophila belli campicola Oberholser, 1946, J. Washington Acad. Sci. 36:388

Holotype.—USNM 53516; adult male; near Wright's Canyon, west slope of Humboldt Mountains, Pershing County, Nevada; 11 September 1867; Robert Ridgway.

Diagnosis.—Mantle pale; mantle streaks distinct; size averages large relative to that of populations in coastal California and Baja California.

Range.—The Great Basin and Rocky Mountain region of western North America west through the Mojave Desert to the San Joaquin Valley of California; extends in winter south to northeastern Baja California, northern Sonora, northwestern Chihuahua, and western Texas.

Amphispiza belli cinerea C. H. Townsend, 1890

Amphispiza belli cinerea C. H. Townsend, 1890, Proc. U.S. Natl. Mus. 13:136

Holotype.—USNM 117575; adult male; Bahía de Ballenas, Baja California, México; 3 or 4 May 1888; Charles H. Townsend.

Diagnosis.—Mantle pale; mantle unstreaked; size averages small relative to that of populations in transmontane California and the Great Basin.

Range.—West-central Baja California in and near the Vizcaino Desert; sedentary.

Had we focused on mean differences we obviously would have recognized more subspecies. On the other hand, had we employed a stringent criterion of 95% separation, as we advocate above, we would still be left with three subspecies of Sage Sparrow, and an even broader overlap of *A. b. nevadensis* and *A. b. canescens*. In either case, the problem we highlighted regarding determination of the population source of Sage Sparrows wintering in the Salton Sink underscores the fallacy of naming populations on mean differences. If mean differences were all that mattered, a sufficiently large sample at each of x evenly spaced points along the cline of increasing wing chord of *A. b. nevadensis* from the western to the northern part of its range would lead to the naming of exactly x subspecies. Such a situation is clearly unacceptable and, worse, biologically meaningless. Yet three subspecies were named along this postulated cline.

Johnson and Marten (1992) reported sizable genetic differences between *A. b. nevadensis* and *A. b. canescens*. Their finding does not affect criteria for recognition of subspecies on the basis of morphology. Genetic studies frequently reveal variation not reflected in morphology. Conversely and importantly, morphological studies frequently reveal variation not reflected in the tiny fraction of the genome studied by genetic techniques (e.g. Rohwer et al. 2001). Furthermore, statistics such as the sum of squares simultaneous test procedure that Johnson and Marten (1992) employed successfully, detect significant mean differences but do not assess diagnosability.

One might wonder how it came to be that two interior subspecies of Sage Sparrow were named and how their validity became so widely accepted (American Ornithologists' Union 1957, Phillips et al. 1964, Rea 1983, Johnson and Marten 1992). Nonetheless, this example is far from unique. A host of subspecies names persists in the literature even though the original description was based on mean differences rather than on diagnosability. *A. b. canescens* is but one of many subspecies we could have presented to make that point. Taxonomy, like any good science, requires careful reanalyses of evidence before changes should be adopted (Patten and Erickson 1996). Many invalid trinomials will persist until someone takes the time to quantify them and reassess their validity.

We believe the subspecies concept to be valid and important. It has the potential to enhance our understanding of speciation, migration, dispersal, geographic variation, and local selection. From our own studies of faunistics it is clear that we would know far less about population movements and biogeographic affinities were it not for careful analyses of subspecific variation. Like any other concept, however, the subspecies concept must be applied properly for it to be of any value (Mayr and Ashlock 1991). Subspecies must be defined on diagnosability, not on mean differences. If misused, the utility of the trinomen is lost and with it goes a vast array of information not easily obtained by other means.

ACKNOWLEDGMENTS

We thank Carla Cicero and Ned K. Johnson (Museum of Vertebrate Zoology), René Corado (Western Foundation of Vertebrate Zoology), Kimball L. Garrett (Natural History Museum of Los Angeles County), and David Willard (Field Museum of Natural History) for access to specimens in their care or loan of material. We thank Heath Wakelee of Minolta Corporation for his guidance with the colorimeter, and we thank the California Department of Fish and Game for making it available to us. This paper benefited from careful reviews by J. V. Remsen, Jr., Frederick H. Sheldon, and an anonymous referee.

LITERATURE CITED

- AMADON, D. 1949. The seventy-five percent rule for subspecies. *Condor* 51:250–258.
- AMADON, D., AND L. L. SHORT. 1976. Treatment of subspecies approaching species status. *Systematic Zoology* 25:161–167.
- AMADON, D., AND L. L. SHORT. 1992. Taxonomy of lower categories—Suggested guidelines. *Bulletin of the British Ornithologists' Club* 112A:11–38.
- AMERICAN ORNITHOLOGISTS' UNION. 1957. Checklist of North American Birds, 5th ed. American Ornithologists' Union, Washington, D.C.
- AVISE, J. C. 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.
- BARROWCLOUGH, G. F. 1982. Geographic variation, predictiveness, and subspecies. *Auk* 99:601–603.
- BEHLE, W. H. 1985. Utah birds: Geographic distribution and systematics. Utah Museum of Natural History Occasional Publication, no. 5.
- BILGIN, C. C. 1999. The use of coefficient of difference (CD) in systematic zoology and the subspecies

- concept. *Turkish Journal of Zoology* 23(Supplement 3):769–773.
- CRACRAFT, J. 1983. Species concepts and speciation analysis. *Current Ornithology* 1:159–187.
- DWIGHT, H. B. 1961. *Tables of Integrals and Other Mathematical Data*, 4th ed. Macmillan Publishing, New York.
- GABRIEL, K. R., AND R. R. SOKAL. 1969. A new statistical approach to geographic variation analysis. *Systematic Zoology* 18:259–278.
- GRINNELL, J. 1905. The California Sage Sparrow. *Condor* 7:18–19.
- GRINNELL, J., AND A. H. MILLER. 1944. The Distribution of the Birds of California. *Pacific Coast Avifauna*, no. 27.
- GROBNER, W., AND N. HOFREITER. 1961. *Intergral-tafel: Erster Teil Unbestimmte Integrale*. Springer-Verlag, Wien, Germany.
- INGER, R. F. 1961. Problems in the application of the subspecies concept in vertebrate taxonomy. Pages 262–285 in *Vertebrate Speciation* (W. F. Blair, Ed.). University of Texas Press, Austin.
- JEHL, J. R., JR. 1968. Type specimens of birds in the San Diego Natural History Museum. *Transactions of the San Diego Society of Natural History* 15:133–139.
- JOHNSON, N. K., AND J. A. MARTEN. 1992. Macrogeographic patterns of morphometric and genetic variation in the Sage Sparrow complex. *Condor* 94:1–19.
- MARSHALL, J. T. 1967. Parallel variation in North and Middle American screech-owls. *Monographs of the Western Foundation of Vertebrate Zoology*, no. 1.
- MAYR, E. 1942. *Systematics and the Origin of Species*. Columbia University Press, New York.
- MAYR, E. 1954. Notes on nomenclature and classification. *Systematic Zoology* 3:86–89.
- MAYR, E. 1969. *Principles of Systematic Zoology*. McGraw-Hill, New York.
- MAYR, E., AND P. D. ASHLOCK. 1991. *Principles of Systematic Zoology*, 2nd ed. McGraw-Hill, New York.
- MICHENER, C. O. 1963. Some future developments in taxonomy. *Systematic Zoology* 12:151–172.
- MILLER, A. H., H. FRIEDMANN, L. GRISCOM, AND R. T. MOORE, EDs. 1957. Distributional check-list of the birds of Mexico. Part 2. *Pacific Coast Avifauna*, no. 33.
- PARKES, K. C. 1982. Subspecific taxonomy: Unfashionable does not mean irrelevant. *Auk* 99:596–598.
- PATTEN, M. A., AND R. A. ERICKSON. 1996. Subspecies of the Least Tern in Mexico. *Condor* 98:888–890.
- PAYNTER, R. A., JR. 1970. Family Emberizidae. Pages 3–443 in *Check-list of Birds of the World*, vol. 13 (R. A. Paynter, Jr. and R. W. Storer, Eds.). Harvard University Press, Cambridge, Massachusetts.
- PHILLIPS, A., J. MARSHALL, AND G. MONSON. 1964. *The Birds of Arizona*. University of Arizona Press, Tucson.
- PIMENTEL, R. A. 1958. Taxonomic methods: Their bearing on speciation. *Systematic Zoology* 7:139–156.
- PIMENTEL, R. A. 1959. Mendelian infraspecific divergence levels and their analysis. *Systematic Zoology* 8:139–159.
- RAND, A. L. 1948. Probability in subspecific identification of single specimens. *Auk* 65:416–432.
- RAND, A. L., AND M. A. TRAYLOR. 1950. The amount of overlap allowable for subspecies. *Auk* 67:169–183.
- REA, A. M. 1983. *Once a River*. University of Arizona Press, Tucson.
- RISING, J. D. 1996. *A Guide to the Identification and Natural History of North American Sparrows*. Academic Press, San Diego, California.
- ROHWER, S., E. BERMINGHAM, AND C. WOOD. 2001. Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. *Evolution* 55:405–422.
- SELANDER, R. K. 1971. Systematics and speciation in birds. Pages 57–147 in *Avian Biology*, vol. 1 (D. S. Farner and J. R. King, Eds). Academic Press, New York.
- SIMPSON, G. G. 1961. *Principles of Animal Taxonomy*. Columbia University Press, New York.
- SMITH, H. M., AND F. N. WHITE. 1956. A case for the trinomen. *Systematic Zoology* 5:183–190.
- SNEATH, P. H. A., AND R. R. SOKAL. 1973. *Numerical Taxonomy*. W. H. Freeman, San Francisco, California.
- THORPE, R. S. 1976. Biometric analysis of geographic variation and racial affinities. *Biological Reviews* 51:407–452.
- UNITT, P. 1987. *Empidonax traillii extimus*: An endangered subspecies. *Western Birds* 18:137–162.
- UNITT, P., K. MESSER, AND M. THÉRY. 1996. Taxonomy of the Marsh Wren in Southern California. *Proceedings of the San Diego Society of Natural History*, no. 31.
- VAN ROSSEM, A. 1911. Winter birds of the Salton Sea region. *Condor* 13:129–137.
- VAN ROSSEM, A. J. 1932. On the validity of the San Clemente Island Bell's Sparrow. *Auk* 49:490–491.
- WILSON, E. O., AND W. L. BROWN, JR. 1953. The subspecies concept and its taxonomic application. *Systematic Zoology* 2:97–111.
- ZINK, R. M., AND J. V. REMSEN, JR. 1986. Evolutionary processes and patterns of geographic variation in birds. *Current Ornithology* 4:1–69.

Associate Editor: F. Sheldon