Null Expectations in Subspecies Diagnosis

Author(s): Michael A. Patten
Source: Ornithological Monographs No. 67:35-41. 2010.
Published By: The American Ornithologists' Union
URL: http://www.bioone.org/doi/full/10.1525/om.2010.67.1.35

BioOne (www.bioone.org) is an electronic aggregator of bioscience research content, and the online home to over 160 journals and books published by not-for-profit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Web site, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/page/terms_of_use.

Usage of BioOne content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Ornithological Monographs
Volume (2010), No. 67, 35–41
© The American Ornithologists' Union, 2010.
Printed in USA.

CHAPTER 3

NULL EXPECTATIONS IN SUBSPECIES DIAGNOSIS

Michael A. Patten[1]
Oklahoma Biological Survey and Department of Zoology, University of Oklahoma,
Norman, Oklahoma 73019, USA

Abstract.—The utility of subspecies in studies of evolution and migration and in conservation planning has been debated hotly for a half-century. Inconsistent and sometimes sloppy application of the subspecies concept has led some to deem it a failure, but recent quantitative definitions of subspecies have put the concept on more rigorous footing. Nonetheless, the molecular revolution has added fuel to the fire as researchers attempt to test subspecies by genetic means. Until a sound and defensible null expectation is developed for genetic differentiation of subspecies, genetic approaches will be fraught with problems. A test for monophyly is insufficient, because parapatric subspecies interbreed by definition. Moreover, because much geographic variation may arise via natural selection, tests restricted to selectively neutral genetic data are likewise problematic. Moreover, long-standing charges of subjectivity in the naming and diagnosis of subspecies must be addressed if subspecies are to continue to be accepted as valid taxonomic entities. Statistical advances, including pairwise tests, spline regression, module identification in neural networks, Monmonier's algorithm, and unsupervised, fuzzy *k*-means cluster analysis offer considerable promise as means of identifying and quantifying geographic variation in an objective yet statistically rigorous manner.

Key words: algorithms, diagnosability, null models, statistics, subspecies.

Expectativas Nulas en la Diagnosis de Subespecies

Resumen.—La utilidad de las subespecies en los estudios de evolución y migración, así como en la planeación de la conservación, ha sido debatida fuertemente por medio siglo. La aplicación inconsistente y a veces descuidada del concepto de subepecie ha llevado a que algunos lo consideren un fracaso, pero el desarrollo reciente de definiciones cuantitativas de las subespecies ha puesto al concepto sobre unas bases más rigurosas. Sin embargo, la revolución molecular le ha agregado combustible al fuego en la medida en que los investigadores han intentado poner a prueba la validez de las subespecies usando herramientas genéticas. Mientras no se desarrolle una expectativa nula sobre la diferenciación genética de las subespecies que sea razonable y defensable, los enfoques genéticos estarán rodeados de problemas. Una prueba de monofilia es insuficiente, porque las subespecies con distribuciones parapátricas, por definición, se entrecruzan. Además, debido a que buena parte de la variación geográfica puede surgir como consecuencia de la selección natural, las pruebas basadas en datos genéticos que son selectivamente neutros también son problemáticas. Más aún, las críticas en cuanto a que existe subjetividad sobre la nomenclatura y la diagnosis de las subespecies deben ser abordadas para que las subespecies puedan seguir siendo aceptadas como entidades taxonómicas válidas. Los avances estadísticos, incluyendo las pruebas por pares, las regresiones de "splines", la identificación de módulos en redes neurales, el algoritmo de Monmonier y los análisis de conglomerados difusos y no supervisados de *k* medias, son altamente promisorios como medios para identificar y cuantificar la variación geográfica de manera objetiva y estadísticamente rigurosa.

---

[1] E-mail: mpatten@ou.edu

Ornithological Monographs, Number 67, pages 35–41. ISBN: 978-0-943610-86-3. © 2010 by The American Ornithologists' Union. All rights reserved. Please direct all requests for permission to photocopy or reproduce article content through the University of California Press's Rights and Permissions website, http://www.ucpressjournals.com/reprintInfo.asp. DOI: 10.1525/om.2010.67.1.35.

The subspecies has had a long and contentious history as a taxonomic unit. After its first use—in 1844, by an ornithologist no less (Sibley 1954, Borgmeier 1957)—the subspecies became a staple of museum taxonomy during natural history's great age of discovery (Goetzmann 1986). An overwhelming number of the world's avian subspecies were named during that time (see Sabrosky 1955), and certainly the zeal of that era contributed mightily to the surfeit of trinomials with which any present-day taxonomist must contend. Criticism of the very concept of subspecies was inevitable given the plethora of subspecific names, many based on trivial perceived differences or tiny sample sizes. Wilson and Brown (1953) fired the first serious salvo, spawning a surfeit of their own: some 20 responses and meta-responses were published in *Systematic Zoology* alone over the next 5 years (Starrett 1958), and others weighed in through the pages of influential books (e.g., Inger 1961, Simpson 1961, Mayr 1969). Calls for balance included Smith and White's (1956) observation that constructive criticism of how subspecies are defined or diagnosed could be, and was being, misconstrued to imply that the subspecies concept itself lacked merit. The end result of the fervent debate was not a resolution but an entrenchment of opposing positions. One need look no further than the well-known forum in *The Auk* (Wiens 1982) to see how little the debate had progressed: 11 prominent avian taxonomists and systematists devoted 23 pages to covering the same arguments heard out, but apparently not exhausted, three decades earlier.

Another round of salvos began recently (Zink 2004, Remsen 2005, Haig et al. 2006, Phillimore and Owens 2006), but by this time the tone of the argument had shifted. Zink (2004) concluded that subspecies obscure biological diversity, specifically in birds, and that this intraspecific category of biodiversity misleads conservation efforts—conclusions reached despite an exclusive reliance on genetic surveys, even though it is not clear what a subspecies should look like genetically. Subspecies names have been applied almost exclusively to populations that differ phenotypically. Because of a lack of widely accepted standards, numerous—perhaps as many as half of all (see Patten et al. 2003:71)—avian trinomials will not stand up to scrutiny, but before we can bring modern statistical and genetic techniques to bear on the problem, it is imperative that we have a clear understanding of what a subspecies is and a firm grasp of the null expectations of the underlying genetics.

The Nature of the Subspecies

Fundamentally, "subspecies are . . . the place at which we stop lumping populations" (Groves 1986). Rensch (*fide* Mayr 1942:106) provided one of the first clear definitions of a subspecies, also known as a "geographical race" or, most commonly, a "race," a loaded term that should be avoided (Patten 2009). Rensch's definition emphasized interfertile individuals that differed in both geographic range and morphology, with morphological differences assumed to have a genetic basis. Mayr himself shortened the definition but again emphasized geography, genetics, and taxonomic distinctness. Four important features about the nature of subspecies can be inferred from these definitions and from others proffered over the years: (1) a subspecies is not reproductively isolated from other subspecies of that species, (2) its defining features have a genetic or developmental basis, (3) it has a unique breeding range separate from that of other subspecies, and (4) it is diagnosably distinct from other subspecies. This fourth feature is what Mayr (1942) meant by subdivisions differing taxonomically.

Patten and Unitt (2002) combined these features into a clear definition of a subspecies: "a collection of populations occupying a distinct breeding range and diagnosably distinct from other such populations." That is, a subspecies represents a level of biological organization below the species that has phenotypic properties that are sufficiently distinct (i.e., separable statistically, as defined below) from other populations. That subspecies are a biological entity whose limits are determined statistically should not deter anyone from making use of this level of biological organization. Subspecies reflect what we can readily observe and have served as the firm basis for decades of literature that has advanced knowledge of distribution, migration, biogeography, ecology, and systematics.

Note that there is nothing in the Rensch-Mayr definition that requires subspecies to be monophyletic. Presence or absence of monophyly is not a relevant criterion for defining or assessing the validity of subspecies, and such a requirement cannot be added. Indeed, this problem exists in defining species, in that monophyly is almost always tested by means of gene trees, particularly using mtDNA (e.g., Zink and Barrowclough 2008), yet as Edwards (2009) observed, monophyly of a gene tree is not a suitable criterion for delimiting a species because natural selection, founder effects, and sampling error affect monophyly but have no bearing on species limits. In the case of subspecies, by the time monophyly could be corroborated, a population may or may not have passed the threshold to become a biological species. Yet reproductive isolation and monophyly evolve independently or are correlated only loosely: reproductive isolation may exist well before monophyly is achieved (e.g., following strong natural or sexual selection) or vice versa (e.g., isolated populations that do not differ phenotypically or behaviorally).

Subspecies are frequently treated as incipient species (e.g., Zink 2004), which is incorrect. Subspecies are, by definition, a stage in the process of allopatric speciation (Mayr 1942b), but even though every species that has evolved in this way must have passed through a subspecies stage, it does not follow that any subspecies that reaches this stage will become a species. Consider a simple abstract model wherein a lineage progresses through different stages of differentiation through time:

$$A \Leftrightarrow B \Leftrightarrow C$$

in which stage A represents an undifferentiated population, stage B a subspecies, and stage C a species. Given allopatric divergence, to reach C, population A must go through B, but once at B the subspecies might continue to C (i.e., speciate) or revert to A (i.e., differentiation swamped by gene flow in secondary contact). A subspecies thus represents an early diagnosable stage through which a biological species must pass (Mayr 1942), and this stage may or may not be an incipient species.

Phenotypic divergence in allopatry may be the result of natural selection, sexual selection, or drift. The environment certainly plays a role in geographic variation (James 1983), but the suggestion that all geographic variation is solely the result of environmental forces is not supported by scientific evidence. For example, Gloger's rule—the tendency for a species to be paler in warm, arid climates and darker in cool, humid ones—is well documented in birds (Zink and Remsen 1986; cf. Chui and Doucet 2009). Underlying mechanisms of Gloger's Rule are open to debate and may include background matching, thermoregulation (Zink and Remsen 1986), adaptation to feather-degrading bacteria (Burtt and Ichida 2004), and other factors still to be identified. Regardless of the cause, the result is geographic variation in plumage with a putative basis in the nuclear genome. Let us not forget that Darwin's (1859) successful and extensive artificial selection of domestic pigeons underscores the genetic basis of wide phenotypic variation within a bird species.

Philosophical Underpinnings

Numerous studies have attempted to apply modern genetic analyses to determine whether one or several subspecies are valid. Many of these studies, across a wide range of organisms, have concluded that a subspecies is not valid because either no genetic difference was detected in some putatively neutral genetic marker or reciprocal monophyly was not found (e.g., Dijkstra and Jellyman 1999, Burbrink et al. 2000, Zink et al. 2000, Daniels et al. 2005, Ramey et al. 2005, Marthinsen et al. 2007). A few authors have drawn more tempered conclusions (e.g., Mock et al. 2002, Bulgin et al. 2003, Swei et al. 2003), and some have even reported new subspecies on the basis of genetic data (e.g., Päckert et al. 2006).

Although a particular species concept may seem to be purely an operational choice that has little bearing on research but says a great deal about one's philosophy, it may be more appropriate to think of species concepts as competing Kuhnian paradigms. Sites and Crandall (1997) intimated that the choice of a species concept should be stated as a hypothesis to be tested because the choice affects interpretation of the results; that is, the concept used affects taxonomic conclusions (Gamauf et al. 2005). On the basis of its definition, the subspecies lies squarely in the purview of the biological species concept. Any diagnosably distinct, geographically circumscribed population may qualify as a subspecies if it is not reproductively isolated from other such populations (and assuming that variation is not merely smoothly clinal). Put simply: if reproductively isolated, the populations are biological species; if not, they are subspecies of a polytypic biological species. It follows, then, that any researcher making use of the subspecies concept in its proper context is making use of the biological species concept to designate phenotypically diagnosable units within the species.

The phylogenetic species concept classifies a population as a distinct species if it is diagnosable by at least one heritable character and if reciprocal monophyly is evident (Cracraft 1983). This definition bears more than a passing similarity to that of a subspecies, but note that it requires 100% diagnosability. Any statistical rule for diagnosing a subspecies at <100%, such as the widely used "75% rule" (see below), renders phylogenetic species and subspecies nonequivalent. Under the phylogenetic species concept, any 100% diagnosable population is elevated to the rank of species, but any population that falls short of 100% diagnosability has no taxonomic status—for example, McKitrick and Zink (1988) advocated not naming subspecies. That subspecies are not recognized in the phylogenetic species paradigm has profound philosophical ramifications for testing a subspecies' validity scientifically. In a hypothetico-deductive framework, the null hypothesis must be "not diagnosable," regardless of paradigm. In the phylogenetic species paradigm, however, the sole alternative hypothesis is that diagnosable taxa are not subspecies but species. These two alternatives exclude the possibility of a subspecies being valid: a population is either undifferentiated or represents a phylogenetic species; there is no middle ground. On purely philosophical grounds, this situation is untenable.

Null Expectations

Most of the papers that I have cited above failed to detect reciprocal monophyly of mtDNA trees and thus concluded that subspecies designations were incorrect. Using monophyly as a criterion meant that these authors were, perhaps unwittingly, equating species and subspecies, in that both taxonomic levels were being held to that same expectation. Recognizing that subspecies are, by definition, a stage in the process of allopatric speciation, the proper null expectation for reciprocal monophyly of a set of subspecies is that it will not exist. Failure to reject this null expectation tells us nothing about a subspecies' validity. Let us further recall that a properly defined subspecies is a diagnosable population not isolated reproductively from other such populations, a definition that allows for the possibility of persistent gene flow between parapatric populations, although not enough to swamp phenotypic diagnosability. Accordingly, a broader question surfaces: What is the null expectation for the genetic differentiation of subspecies, particularly at a neutral locus?

A null expectation of no difference among populations under study is both obvious and unhelpful, yet it is fundamentally difficult to construct a test for H_0 = no genetic differentiation. By definition, a neutral marker is not associated with local adaptation, so this locus may be relatively uniform throughout two distinct populations that hybridize across a shared boundary. Taxa that hybridize even modestly are likely to have at least some introgressed mtDNA, and in some cases mtDNA of one taxon replaces mtDNA in another taxon despite no evidence of introgressed nuclear genes or of mixed morphology (Ballard and Whitlock 2004). Yet ongoing (potential) hybridization characterizes a subspecies—the lack of reproductive isolation is why subspecies, although distinct morphologically, are not classified as biological species.

A further complication arises because subspecies are often posited to be the products of local adaptation, and such divergence may have been geologically recent (see Reznick et al. 2004). Rapid, disruptive selection may yield distinct phenotypes with an underlying genetic basis that is invisible to a neutral marker because differentiation is evident only in non-neutral portions of the nuclear genome. A selective sweep may alter gene frequencies to a point that they no longer reflect phylogeny accurately. Balancing selection (e.g., through frequency-dependent selection) can decouple mtDNA phylogeny and taxon pedigrees (Rand 1996). And strong stabilizing selection on morphology and behavior may produce two allopatric populations that are uniform phenotypically and behaviorally but, if they have been isolated long enough, distinct genetically.

The point about genetic divergence raises another philosophical issue: use of neutral markers assumes that we need only concern ourselves with time. Not to belittle mutation rates, population sizes, sex ratios, or dispersal, but time and natural selection constitute the two pillars of divergence. Emphasizing only time is easier operationally, but doing so ignores our growing understanding of how speciation works. If all speciation occurred via random accumulation of genetic differences when populations are allopatric, the extent to which a neutral genetic marker differs will be in proportion to the degree of divergence between the taxa under study. Any other mechanism that promotes speciation, such as disruptive or directional selection in the face of gene flow (Nosil 2008), would create discordance between the neutral marker and the signal of diverging loci and phenotypic characters under selection (for a review, see Winker 2009).

Any attempt to determine the validity of a subspecies needs to take these caveats into account and use them to construct a meaningful null. This task will not be easy. Whether divergence was recent or the result of strong selection or occurred despite persistent gene flow, we might well expect "no difference" in gene frequencies or haplotypes and expect no reciprocal monophyly. If no difference is the expectation, how can we reject H_0?

Statistical Approaches

A working null hypothesis of no genetic differentiation when phenotypic differentiation is evident must, at a minimum, examine a larger portion of the genome. Perhaps the day will come when enough genomes have been characterized to pinpoint specific genes associated with observed phenotypic differences. Another option would be to develop a model-comparison approach (*sensu* Anderson et al. 2000), wherein the weight of evidence for competing alternative phylogenetic schemes—assuming that they could be characterized—could be judged objectively. Until either of these lofty goals is reached, the best position for advocates of the subspecies concept is to opt for statistical rigor in defining what they mean.

It is no secret that numerous named subspecies are ill defined, in that they do not conform to a rigorous, formal definition of subspecies based on diagnosability (here meaning an emphasis on effect size, not on mean differences). One solution would be to discard the notion of subspecies (e.g., McKitrick and Zink 1988); if they do not exist, they cannot be ill defined. That other levels of taxonomic classification suffer from similar problems typically passes unremarked by those who favor discarding subspecies. A rank-free system (e.g., de Queiroz and Gauthier 1992) would eliminate all such problems, but rank-free systems are not without drawbacks (Benton 2000). Accordingly, rather than discard subspecies, we ought to ensure that they are well defined and defensible. If a researcher then chooses to ignore subspecies, at least it will be for philosophical reasons rather than discomfort that a particular taxonomic scheme may be arbitrary or trivial.

Statistical assessment of subspecies dates back to the advent of the 75% rule (Amadon 1949). At its core, this rule was an attempt to emphasize that a large effect size, not whether means differ significantly, is the statistic of interest as well as what should matter biologically. (Effect size is a statistical term for a measure, in terms of standard deviation, of distance between the sampling distributions of H_0 and H_A; i.e., it is a measure of the degree of overlap between the null distribution and that of the alternative.) Mayr's (1969:189) related coefficient of difference (CD),

$$\mathrm{CD} = \frac{\bar{x}_b - \bar{x}_a}{\mathrm{SD}_b + \mathrm{SD}_a}$$

stressed this point further, as we can see in comparing it with a simple formula for effect size (**d**; Cohen 1988:20):s

$$\mathbf{d} = \frac{\bar{x}_b - \bar{x}_a}{\sigma}$$

where σ is the (pooled) standard deviation or the SD for either population (they are assumed to be equal). Yet despite numerous efforts to solidify this foundation (Rand and Traylor 1950, Simpson 1961, Mayr 1969), the emphasis on a large effect size still goes largely unheeded. And from a practical standpoint, the vast majority of subspecies were named before this framework was built (see Remsen, this volume), so the vast majority of subspecies have not been evaluated with modern statistical methods (Patten and Unitt 2002, Remsen 2005).

Two approaches have been proposed in the past decade. The first is a pairwise test of diagnosability (Patten and Unitt 2002), a technique based on the *t*-distribution that can be applied either to a single morphometric or other quantitative (e.g., scores from a colorimeter) character or to a composite value derived from a multivariate statistic such as a discriminant analysis (see Marantz and Patten, this volume). The pairwise test is limited to testing for differences among predefined groups. The other recent method involves application of spline- and step-regression techniques to identify breaks among characters that vary clinally (Skalski et al. 2008). It is unclear how sensitive this method is for detecting small but potentially biologically meaningful steps, because the simulation results presented by Skalski et al. (2008) have abrupt breaks in character values.

Three additional techniques would be worth developing for use in diagnosing morphologically distinct groups that have separate geographic ranges (i.e., subspecies). The first is a simulated annealing algorithm that detects modules (analogous to clusters) in a network (Guimerà and Amaral 2005). The algorithm is unsupervised, which means that there is no need to define in advance the number of clusters desired. Instead, a measure of degree of modularity signifies how distinct the modules are. Input matrices are a set of links, which could be between spatial location (Carstensen and Olesen 2009) and morphological traits or between individuals and those traits. If the latter were used, then short links would indicate that individuals were close to one another and, thus, potentially in the same module. Software for this algorithm is available (see Olesen et al. 2007).

Monmonier's maximum-difference algorithm (Monmonier 1973, Manni et al. 2004) is a spatially explicit technique that can use any distance matrix to determine where barriers (natural breaks) exist in the data. The algorithm begins with a map of sites using specific coordinates (e.g., latitude and longitude). A distance (dissimilarity) matrix is mapped onto a triangulation created among sites such that each pairwise line between sites has an associated distance. The algorithm builds barriers, beginning with the maximum pairwise distance and continuing until a loop forms or the map's edge or another computed barrier is hit. Distances can be from any semimetric index, such as F_{ST} (Nicholls et al. 2006) or Jaccard's index (Patten and Smith-Patten 2008), and are thus amenable to either genetic or morphometric data. Although Monmonier's algorithm is supervised, there are ways to determine *a priori* how many barriers to select (Patten and Smith-Patten 2008). A key advantage to this technique stems from a recent software implementation that allows input of multiple matrices (Manni and Guérard 2004). As such, either bootstrap matrices can be generated to determine support for specific barriers (Manni et al. 2004, Patten and Smith-Patten 2008) or matrices from multiple data sets can be entered to determine levels of correspondence among them.

On purely heuristic grounds, perhaps the most appealing statistical option is an unsupervised fuzzy cluster analysis, a technique virtually unknown in biology but used widely in pattern recognition and artificial intelligence. In classical set theory, group membership is discrete if sets are exclusive. Object x is an element either of set **A** or of set **B**. In fuzzy set theory, group membership is a probability, which means that x might have $P = 0.73$ of being an element of **A** and $P = 0.27$ of being an element of **B**. Such a probabilistic partitioning seems more naturally in accord with the fuzzy nature of subspecies—they are defined by real or potential gene flow and, therefore, every individual may not be classifiable. Operationally, membership probabilities can be assigned by scaling linearly between minimum and maximum values for the variable in question (Roberts 2008). These assigned P values are then used in a clustering algorithm to build k clusters on the basis of minimizing some predefined objective function (Gath and Geva 1989, Liu and George 2005). Software applications can be found on the Internet (e.g., Kenesei et al. 2006).

Although the foregoing discussion is geared toward morphological data, closely related techniques are already used widely for genetic data. For example, assignment tests (e.g., Piry et al. 2004) classify individuals probabilistically—using a Bayesian approach with priors set by group of origin—among a set of populations, and genetic assignment can correspond well to subspecies (e.g., Pruett et al. 2008a). Algorithms such as STRUCTURE (Pritchard et al. 2000) work like an unsupervised k-means cluster analysis in that genetic data are used, again within a Bayesian framework and this time coupled with Monte Carlo randomization, to determine the number of groups most likely given the data. STRUCTURE further allows genetic data to be constrained by phenotype, and individuals can be assigned a genetic score that can be correlated against a morphological score (e.g., Patten et al. 2004).

Concluding Remarks

Conservation policies and wildlife management decisions often are based on subspecific taxonomy (Stanford 2001, Haig et al. 2006). Practitioners must be able to defend subspecies diagnoses critically and objectively, a move that will doubtless lead to a more cautious and conservative intraspecific taxonomy, which is warranted given the high stakes and limited funding available for biodiversity conservation. There is certainly no point in complaining that others ignore subspecific designations if advocates cannot defend them quantitatively. Conversely, advocates of the phylogenetic species concept ought to be forthcoming about their philosophical inability to test the validity of a subspecies—thus far, no one has explained how such a test can be done within that paradigm.

The utility of subspecies rests on our ability to assign, with a high degree of confidence, individual specimens to a particular geographic population. A properly named subspecies is a surrogate for geographically isolated pools of phenotypic variation whose basis is putatively genetic. Diagnosability of subspecies has been defined statistically for qualitative and quantitative characters (e.g., Patten and Unitt 2002). A roughly ordered expectation of which types of characters are best suited to proper diagnosis of subspecies is pattern > color ≈ shape > size. This ordered expectation may not hold in all cases; it only suggests that size, for example, is more likely than pattern to vary clinally. Smoothly clinal variation has no place in the proper description of subspecies. We instead need to emphasize only those characters that exhibit "breaks" between geographic sites (see Skalski et al. 2008).

To achieve rigor in subspecies diagnosis will require strict adherence to the now 60-year-old recognition that extent of overlap is what matters. Between the statistical and genetic techniques summarized above and recent advances in the quantification and analysis of shape (González-José et al. 2008) and color (Endler and Mielke 2005), there are few excuses to avoid the careful work needed to put subspecies on a firm quantitative and objective footing. Yet there are recent examples of subspecific diagnoses that relied on tests of mean differences (e.g., Engelmoer and Roselaar 1998, Jiguet 2002, Cabot and Urdiales 2005) rather than focusing on effect size. Editors and peer reviewers cannot allow this practice to continue, and the time to integrate approaches is now.

Acknowledgments

I thank L. R. Bevier, J. V. Remsen, Jr., P. Unitt, and K. Winker for input over the years; each has been instrumental in shaping my thinking on this subject. I further thank K. Winker and S. M. Haig for inviting me to present a version of this paper at the symposium on subspecies. F. C. James, J. V. Remsen, Jr., B. D. Smith-Patten, and K. Winker provided excellent comments on various drafts of this paper.