

Neither Individually Distinctive Songs Nor “Lek Signatures” are Demonstrated in Suboscine Screaming Pihas

Author: Kroodsma, Donald

Source: The Auk, 128(4) : 789-790

Published By: American Ornithological Society

URL: <https://doi.org/10.1525/auk.2011.128.4.789>

The BioOne Digital Library (<https://bioone.org/>) provides worldwide distribution for more than 580 journals and eBooks from BioOne's community of over 150 nonprofit societies, research institutions, and university presses in the biological, ecological, and environmental sciences. The BioOne Digital Library encompasses the flagship aggregation BioOne Complete (<https://bioone.org/subscribe>), the BioOne Complete Archive (<https://bioone.org/archive>), and the BioOne eBooks program offerings ESA eBook Collection (<https://bioone.org/esa-ebooks>) and CSIRO Publishing BioSelect Collection (<https://bioone.org/csiro-ebooks>).

Your use of this PDF, the BioOne Digital Library, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Digital Library 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 is an innovative nonprofit that 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.



The Auk 128(4):789–790, 2011
 © The American Ornithologists' Union, 2011.
 Printed in USA.

Neither individually distinctive songs nor “lek signatures” are demonstrated in suboscine Screaming Pihas.—Why vocal learning has evolved in some taxa (most notably, songbirds, parrots, and hummingbirds) but not others remains an unanswered evolutionary question. Recent evidence from suboscine bellbirds (*Procnias* spp.; Kroodsma 2005, Saranathan et al. 2007), however, reveals that song learning may be more common than previously thought, and so with considerable excitement I began to read Fitzsimmons et al.'s (2008) article about “lek signatures” and the possibility of song learning in the suboscine Screaming Piha (*Lipaugus vociferans*; hereafter “piha”). By the end of the article, however, I was sadly disappointed.

In short, the data used in the study consist of 256 songs recorded from 26 male pihas that were distributed among four “exploded leks” (with 3, 4, 6, and 13 males in the four leks). Birds were recorded sequentially between 0730 and 1300 hours over 5 days, with birds in the same lek usually recorded on the same day, and 13 frequency or time measurements were made or calculated from the 8 or 10 best recorded songs from each male.

The first major problem is with sampling. Because the reported song differences among males and among leks are so subtle and are not obvious in sonagrams (see Fitzsimmons et al. 2008: figure 3) but are instead measured by computer, to the nearest 22 Hz and 2.9 ms, considerable care must be taken to ensure that the results of the analyses are not artifacts of the sampling methods. If each male is sampled during a single half-hour period, for example, but different males are sampled over a significant portion of the day (from 0730 to 1300 hours), in a variety of unknown contexts (female present or absent in lek, interaction with neighboring males current or not), over a 5-day span, then it is imperative that each male also be sampled at a variety of times of day, in a variety of motivational contexts, and on different days. If, like Fitzsimmons et al., one chooses not to sample each male in multiple sessions, one cannot attribute the between-session variation to individuals, only to the different “recording sessions.” Nor can one convincingly demonstrate that songs vary among leks if one has sampled the leks in succession and each lek primarily on only one day.

The biological reason for these statistical limitations is simple: songs of birds are highly expressive, and they can vary in subtle to striking ways with time of day and motivational context. Examples of such variation in song abound in the literature. Individual song phrases of a Scarlet Tanager (*Piranga olivacea*) can differ in duration from early morning to midmorning, for example, and some birds (e.g., the non-songbird Belted Kingfisher [*Ceryle alcyon*]) change frequency characteristics of their vocalizations from one context to another (Kroodsma 2009). For the pihas, as motivation

to sing waned and males slowed their song rates from early morning to midday, or if a female appeared in the lek, even the most subtle changes in songs due to different contexts could have a marked effect on the outcome of the analyses. Given the non-obvious differences among the sonagrams, I would not be surprised if the reported results were even dictated, in part, by the relatively poor recordings used in the study: with three different shotgun microphones with different polar responses being used to record singing birds at various positions in the rainforest canopy, the resulting large and varying amounts of degradation in the recordings could have a significant effect on fine-scale frequency and time measurements (see reverberation lasting >100 ms in the sonagrams).

The second major problem is in analysis. One set of analysis problems arises directly from the sampling failures. For example, although it is reported that “all 13 song features were more variable *among males* than *within males*” (Fitzsimmons et al. 2008:911; my italics) and that canonical discriminant function analysis “assigned 93.2% of songs to the correct *male*” (p. 912; my italics)—the authors thereby claiming that *males* have individually distinctive songs—all that can really be claimed is that all 13 song features were more variable *among recording sessions* than *within recording sessions* and that the canonical discriminant function analysis assigned 93.2% of songs to the correct *recording session* (but see below for other problems with the discriminant analysis). Without sampling individual males in different recording sessions, thereby controlling for known and unknown sources of song variation, the observed song variation cannot be attributed to male distinctiveness, only to recording-session distinctiveness.

Most troublesome among the analysis problems are those that plague the discriminant function analysis that the authors rely on for their “lek signature” conclusion; I provide three examples. (1) Most glaring is the issue of “simple pseudoreplication” (Hurlbert 1984). The 10 (or 8) songs from each male are incorrectly used as if they are independent samples, as in the following statements: “canonical discriminant analysis assigned 76.4% of *songs* to the correct lek...[and the] discriminant analysis capably differentiated *songs* from the four different leks...suggesting that variation between *males* at different leks is small, but present” (Fitzsimmons et al. 2008:912–913; my italics). The sampling problem cripples this analysis from the outset, but, additionally, the desired test is whether males differ among leks (as stated in the conclusion of the above quotation), and to use the 256 songs as if they represented 256 different males is false replication (McGregor et al. 1992); this error is serious and not a mere technicality, because this particular kind of “pseudoreplication...tends to produce (sometimes grossly) incorrect results...and discriminability of...groups...can be drastically overestimated” (Mundry and

Sommer 2007:965). (2) The authors appear to have used the discriminant function to classify the same songs that they used to compute the function; this kind of circular analysis vastly increases the likelihood that the function will seem to correctly classify the songs, and is simply bad practice (Tabachnick and Fidell 2007). (3) The function “assigned 76.4% of songs to the correct lek, well above the 25% level of correct assignment expected by chance”; statistical significance for the function is implied, yet no such test was done, and it seems incorrect to simply assert a 25% chance level of assignment when, to complicate matters, 50% of the songs come from one lek. Preferably, one determines *a priori* the chance classification probability for each category and then determines how close the classification comes to those probabilities (Tabachnick and Fidell 2007).

The third problem is in interpretation. I provide one example: “Our finding that Screaming Pihas sing individually distinctive songs adds to growing evidence that there may be a learned component to song in some suboscines” (Fitzsimmons et al. 2008:913). This statement is in the final sentence of the paper, the place where an author wants to leave the reader with a lasting impression about the significance of a study, yet the statement is nonsensical and, even worse, misleading, because songs in a wide range of species (most likely all species) are individually distinctive whether the songs are learned or not. In nonlearning flycatchers (*Empidonax* spp.), for example, songs are individually distinctive, perhaps best documented by the two papers the authors cite about the Alder Flycatcher (*E. alnorum*; Lovell and Lein 2004a, b); the birds even use the variation to discriminate among individuals. Even if individually distinctive songs had been demonstrated for the pihas, such a finding would have no bearing on whether the songs were learned or not.

Given the paper’s problems in sampling, analysis, and interpretation, Fitzsimmons et al. (2008) cannot reach any valid conclusions about whether songs are individually distinctive. Nor do they present valid evidence of songs differing from lek to lek. Nor are the findings relevant to the question of vocal learning in suboscines.

When papers like this appear in print, authors rightly share blame with others who facilitate the publication process, including reviewers and editors. How this extended responsibility can fail is illustrated not only by the initial publication of Fitzsimmons et al. (2008) but also by the reluctance of those involved in the review process to share my desire that a severe, but fair, review be published. This cavalier attitude toward the design of research and the collection and analysis of numbers is unacceptable, because such permissiveness undermines the very science we claim to be doing. The present case is not unique, and such flawed papers can do considerable damage if they go unchallenged. If the research model is emulated by others and if the conclusions and logic are accepted as reported, progress in understanding birds is confused and stymied (for additional discussion, see Byers and Kroodsma 2009). We deserve better from each other, and we should hold each other to a higher standard.

Acknowledgments.—I thank K. Yasukawa and editors S. Sealy and M. Murphy for their advice on the content of this letter, and two reviewers whose advice I sought but who chose to remain anonymous.—DONALD KROODSMA, 36 Kettle Pond Road, Amherst, Massachusetts 01002, USA. E-mail: dekroodsma@yahoo.com

LITERATURE CITED

- BYERS, B. E., AND D. E. KROODSMA. 2009. Female mate choice and songbird song repertoires. *Animal Behaviour* 77:13–22.
- FITZSIMMONS, L. P., N. K. BARKER, AND D. J. MENNILL. 2008. Individual variation and lek-based vocal distinctiveness in songs of the Screaming Piha (*Lipaugus vociferans*), a suboscine songbird. *Auk* 125:908–914.
- HURLBERT, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211.
- KROODSMA, D. E. 2005. *The Singing Life of Birds: The Art and Science of Listening to Birdsong*. Houghton Mifflin, Boston.
- KROODSMA, D. E. 2009. *Birdsong by the Seasons: A Year of Listening to Birds*. Houghton Mifflin Harcourt, Boston.
- LOVELL, S. F., AND M. R. LEIN. 2004a. Neighbor-stranger discrimination by song in a suboscine bird, the Alder Flycatcher, *Empidonax alnorum*. *Behavioral Ecology* 15:799–804.
- LOVELL, S. F., AND M. R. LEIN. 2004b. Song variation in a population of Alder Flycatchers. *Journal of Field Ornithology* 75:146–151.
- MCGREGOR, P. K., C. K. CATCHPOLE, T. DABELSTEEN, J. B. FALLS, L. FUSANI, H. C. GERHARDT, F. GILBERT, A. G. HORN, G. M. KLUMP, D. E. KROODSMA, AND OTHERS. 1992. Design of playback experiments: The Thornbridge Hall NATO ARW Consensus. Pages 1–9 in *Playback and Studies of Animal Communication* (P. K. McGregor, Ed.). Plenum Press, New York.
- MUNDRY, R., AND C. SOMMER. 2007. Discriminant function analysis with nonindependent data: Consequences and an alternative. *Animal Behaviour* 74:965–976.
- SARANATHAN, V., D. HAMILTON, G. V. N. POWELL, D. E. KROODSMA, AND R. O. PRUM. 2007. Genetic evidence supports song learning in the Three-wattled Bellbird *Procnias tricarunculata* (Cotingidae). *Molecular Ecology* 16:3689–3702.
- TABACHNICK, B. G., AND L. S. FIDELL. 2007. *Using Multivariate Statistics*, 5th ed. Allyn & Bacon, Boston.

Received 2 July 2009, accepted 30 July 2010.

The Auk 128(4):790–792, 2011

© The American Ornithologists' Union, 2011.

Printed in USA.

Further analysis supports the conclusion that the songs of Screaming Pihas are individually distinctive and bear a lek signature.—The conclusion of our field study (Fitzsimmons et al. 2008) was that three complementary methods of analysis demonstrated significant differences in song features between individual male Screaming Pihas (*Lipaugus vociferans*) and, to a lesser degree, distinctiveness based on the lek at which they were recorded. Kroodsma (2011) presents many criticisms of our paper, related to our sampling approach, our acoustic measurements, our analytical approach, and our interpretation. Here, we address these criticisms and provide additional data and analyses in support of our conclusions. We argue that, despite some deficiencies, our investigation provides an interesting contribution to the literature on suboscine songbirds.

Sampling approach.—Kroodsma’s (2011) first main criticism is related to our sampling approach. He points out that each male’s songs were recorded during only one recording session, and he suggests that our observed differences are due to differences between recording sessions rather than differences between individuals. As