Statistical Inference from Color-banding Data

GEOFFREY E. HILL[1] AND DAVID E. CARR[2]

The use of colored leg bands to identify birds at a distance has become a standard technique in field ornithology. Typically, the researcher chooses colors that are most visible and distinguishable in the field; little attention is paid to similarity between leg band colors and the coloration of the species under study. Thus, the natural coloration of an individual is often enhanced inadvertently.

Studies of the effects of leg-band color on mate choice and reproductive success in the Zebra Finch (*Poephila guttata*) demonstrated the impact that colored leg bands can have on a bird's reproductive and social interactions (Burley 1981, 1985, 1986; Burley et al. 1982). Burley's results justifiably caused concern among biologists whose field studies were founded on the assumption that leg-band colors have no significant effect on the biology of birds. Despite widespread concern, however, only a few published studies have considered the effects of colored leg bands on wild or wild-caught birds (Watt 1982, Brodsky 1988).

Recently, Hagan and Reed (1988) claimed to demonstrate a significant effect of red leg bands on the reproductive biology of Red-cockaded Woodpeckers (*Picoides borealis*). There are, however, several fundamental problems with the analysis of Hagan and Reed that cause us to question the validity of their conclusions. Because of the serious implications of Hagan and Reed's paper to field ornithology, we felt it was important to point out several flaws in their analysis.

Hagan and Reed used data from a long-term study of Red-cockaded Woodpeckers. Birds were banded with 10 different colors, but only 9 were used in the analysis because 1 of the colors was used only briefly. Each bird received 3 or 4 colored bands and one aluminum band. To look for effects of band color on life-history parameters (including success at fledging young, the sex ratio of nestlings, and the survival of fledged young), colors were considered one at a time. For each color, birds were divided into those wearing the color and those lacking it. Thus, 9 separate ANOVAs were run to analyze the fledgling production data, and 9 separate Chi-square tests of independence were used to analyze both the sex ratio of offspring and nestling survivorship. Only the color red was found to have an effect on fledgling production and offspring sex ratio; red, light green, and white were found to affect nestling survivorship.

[1] Museum of Zoology and Department of Biology, University of Michigan, Ann Arbor, Michigan 48109 USA.

[2] Department of Zoology, University of Maryland, College Park, Maryland 20742 USA.

Unfortunately, Hagan and Reed did not correct for the increased probability of a Type I error (rejecting a true H_0) when >1 comparison involving a single dependent variable is made (Steel and Torrie 1980: 182). The probability of committing at least one Type I error when making 9 comparisons with an $\alpha = 0.05$ is $P = 1 - (1 - \alpha)^9 = 0.37$. In other words, when testing 9 different colors for an effect on reproductive success, one would have a 37% chance of finding a significant effect when none existed. Even at $\alpha = 0.01$, there is still a 9% probability of committing at least one Type I error. To get an acceptable Type I error, Hagan and Reed should have used $\alpha = 0.005$ (Sokal and Rohlf 1981: 242). With an appropriate α, there is no statistically significant effect of red color bands on the number of fledglings, nestling survival, or fledgling sex ratio and no basis for their conclusions.

Another problem with Hagan and Reed's analysis is that the data they used were not independent. First, the same observation can appear in >1 "treatment." For example, if a bird wore neither a red nor a white band, that individual was entered into the red "treatment" data set for red vs. non-red comparisons as well as the white "treatment" for white vs. non-white. Similarly, if a bird wore both red and white bands, that bird's data were used in both the red and the white "treatments." In addition, observations of a single bird made in different years were treated as independent samples. It is not immediately clear what effect this lack of independence would have on their analysis (in terms of making it more or less conservative), but it obviously violates a basic assumption of the statistical tests that they employed.

Finally, the most disturbing problem with Hagan and Reed's data is that colors were not assigned to the birds in a random or even a haphazard fashion. Instead, colors were used to define groups so that all members of a group received the same two to three colors on one leg. This marking scheme confounds the effect of "group" with all of its associated genetic and environmental (territory quality, group size, number and quality of helpers) correlations, and the effect of "band color." For example, if a group occupied a territory of low quality and was defined by red color, there could be a spurious correlation between red color and poor breeding success. In their discussion, Hagan and Reed ignore the problem of systematic application of the same color combination to individuals within a group and deal only with the nonrandom assignment of color combinations to groups and colors to individuals. They conclude that there are no significant differences between red-banded and non-red-banded males in number of helpers, group size, or spatial biases. However, Hagan and Reed did not and could not have addressed all potential group effects or combinations of group effects that may have resulted in the patterns that they observed. Indeed, they indicate that non-red-banded males tended to have more helpers than red-banded males. The difference was not significant ($P = 0.12$), but fewer helpers could be symptomatic of general group features that could have decreased the reproductive output of red-banded males.

Because of the widespread use of colored leg bands in field ornithology and the likelihood that in some cases colored bands have an undesired effect on the birds under study, there is a serious need for field studies to test the effects of leg bands on wild birds. At the same time, because of the potential implications for many studies that are founded on the assumption that colored leg bands have no effect on birds, investigations of the effects of color bands must be carefully designed and data must be analyzed and interpreted conservatively. We believe that the study by Hagan and Reed is sufficiently flawed in both the design of the experiment and the interpretation of results that it should be discounted as evidence for an effect of colored leg bands.

We thank Robert Payne for his comments on an earlier draft of this manuscript.

Literature Cited

Brodsky, L. M. 1988. Ornament size influences mating success in male Rock Ptarmigan. Anim. Behav. 36: 662-667.

Burley, N. 1981. Sex-ratio manipulation and selection for attractiveness. Science 211: 721-722.

———. 1985. Leg band color and mortality patterns of captive breeding populations of Zebra Finches. Auk 102: 647-651.

———. 1986. Sex-ratio manipulation in color banded populations of Zebra Finches. Evolution 40: 1191-1206.

———, G. Krantzberg, & P. Radman. 1982. Influences of color banding on the conspecific preferences of Zebra Finches. Anim. Behav. 30: 444-455.

Hagan, J. M., & J. M. Reed. 1988. Red color bands reduce fledgling success in Red-cockaded Woodpeckers. Auk 105: 498-503.

Sokal, R. R., & F. J. Rohlf. 1981. Biometry. New York, W. H. Freeman and Co.

Steel, R. G. D., & J. H. Torrie. 1980. Principles and procedures of statistics: a biometrical approach. New York, McGraw-Hill Book Co., Inc.

Watt, D. J. 1982. Do birds use color bands in recognition of individuals? J. Field Ornithol. 53: 177-179.

Received 21 December 1988, accepted 31 January 1989.

Response to Hill and Carr

John M. Hagan[1] and J. Michael Reed[2]

Hill and Carr (1989) present three criticisms of our analyses of color-band effects on Red-cockaded Woodpeckers (*Picoides borealis*; Hagan and Reed 1988): (1) that, by failing to adjust the experimentwise error rate as a result of performing multiple tests, our conclusions concerning effects of red bands are invalid, (2) that we have two problems with independence of observations (the sample observations of reproductive success appeared in >1 treatment, and observations of a single individual bird made in different years were treated as independent samples), and (3) that color bands were not assigned randomly to the individuals in our study.

The first criticism, that the error rate was not properly adjusted, is based on a fallacious interpretation of our analytical approach. In fact, the error rate needs no adjustment for the analysis of the color red. We clearly stated in the introduction that our a priori hypotheses concerned only red bands. These hypotheses were constructed in advance of analysis, and based on biological reality and conclusions reached by Burley in research on Zebra Finches (*Poephila guttata*) (1981; 1985a, b; 1986a, b, c; Burley et al. 1982). Plumage of the Red-cockaded Woodpecker is entirely black and white, with the interesting exception that males have a small tuft of red feathers above and behind the eye that is revealed during intense agonistic encounters. Our a priori hypotheses were based on plumage color of this particular species. Thus, for each population parameter, we had only a single hypothesis, regarding a single color, red. The Type I error rate for these a priori hypotheses is 0.05, for $\alpha = 0.05$.

We analyzed the effects of other band colors subsequent to our tests of "red hypotheses" as a result

[1] Manomet Bird Observatory, P.O. Box 936, Manomet, Massachusetts 02345 USA.

[2] Department of Zoology, Campus Box 7617, North Carolina State University, Raleigh, North Carolina 27695 USA.