Studies in Avian Biology No. 6:450–451, 1981.

SUMMARIZING REMARKS: SAMPLING DESIGN

JAKE RICE[1]

Several major points addressed by the papers in this session on experimental design have already been discussed by previous summarizers. The fact that previous speakers make redundant parts of my summary of this session only serves to emphasize that these design considerations are essential points, demanding the attention of the biometrical and the ecological communities. Rather than repeat many of those points, I will reiterate just a few general comments here and then expand on some of the specific points raised by speakers in this session.

The frequent references to the basic points of experimental design emphasize both how fundamental and how straightforward these basics are. At their simplest, the purposes of experimental design are to maximize accuracy and precision through the minimization of bias and within-group variation. As numerous papers at this conference have amply documented, much work remains for biometricians interested in methods of counting bird populations, for there are many sources of bias and variance. Rather than list them yet again, I simply refer readers to Verner's introductory comments to this session, and the summary comments of the sessions on "Estimating Birds per Unit Area," "Comparison of Methods," and "Observer Variability."

Next I wish to focus on an additional caveat in design considerations: The variance and bias minimization must be achieved under constraints of fixed, and usually limiting funds and resources. Gates presented a brief introduction to optimization of resource use in allocating available time and personnel to counting efforts. Surprisingly little information is required to use those optimality calculations: (1) we must know the cost of laying out transects, mapping grids or stations; (2) we must know the cost of sampling the transect, grid or point (in practice, both of the above pieces of information, or at least rough estimates of them, are readily available); and (3) finally, we need to specify the desired power of our study. Although "power" is a statistical term, the basic point here is a biological one: how small a difference does one want to be able to detect, either for the management implications of the differences, or for the importance of the difference in evaluating theoretical predictions. It may be difficult for a researcher to specify precisely the magnitude of such differences, but usually all that is needed is a general estimate, and such estimates are usually possible.

The true difficulty in the optimal allocation problem is that one also needs an estimate of the inherent variance of the system under study. The importance of pilot studies in good experimental design was brought out by the papers of Gates (1981), Dawson (1981b), and Pollock (1981). Blondel et al. 1981 used a somewhat different approach, but their stress of confidence intervals for their population estimates stems from the importance of these same within-group variance sources. For some reason, field ecologists are loathe to conduct pilot studies, and even more reluctant to use in optimization models the knowledge of system variance provided by their own prior work or literature sources. It strikes me as strange that ecologists complain frequently about inadequate resources, and yet are reluctant to determine the best use of whatever resources they do have.

This reluctance cannot be due either to a lack of access to statistical consultants or to hesitancy about the use of computers and sophisticated mathematical algorithms. The same ecologists readily find statisticians to consult once they have data in hand, and they are then willing to conduct all sorts of complex statistical routines, such as principal components analysis, multiple regressions and the like. I feel that the reluctance of ecologists to involve design consultants in the initial planning of a study is due to their fear of getting advice akin to that one would derive from Dawson's Figure 5. This figure shows the intimidatingly large number of counts that are necessary to detect even a modest difference between numbers of uncommon species. We are generally afraid to know the magnitude of the Type II errors in our statistics. Other papers in this session also illustrate aspects of this propensity toward high Type II error rates in count studies, for example, the size of the confidence interval around a measure as rough as the mean number of species per I.P.A. count shown in the paper by Blondel et al. 1981. In this case, after more than 30 counts, the confidence interval still seems to be around four units (species) wide, whereas the mean is only slightly greater than 12 units (species). The paper by Morrison et al. (1981) illustrates the other side of the problem, in that their Table 2 shows that with few counts even "replicate" plots can dif-

[1] Center for Environmental Studies, Arizona State University, Tempe, Arizona 85281.

fer significantly (high Type I errors due to inadequate sampling effort, as they conclude) and their Table 1 shows that even after 10 stations, the coefficients of variation in their density estimates are greater than 10% in 4 of the 6 study areas and greater than 30% in two of them. Such variation in total avian density again indicates that probably many plots would be necessary to have much power for between plot comparisons of uncommon species.

In this light, two points in the paper by Pollock (1981) deserve special note. The first is that studies should be designed specifically to meet the assumptions of the models to be tested. If one begins a field project unaware of what uses will be made of the data, the study is premature; clearer research objectives need to be established before field work commences. Second, Pollock points out that often a small study is no better than no study at all. In many cases I would go even further and say that a small study, with an inherently high Type II error rate, is *worse* than no study at all. An inadequately supported study, doomed from the onset to produce low accuracy and a high likelihood of not detecting differences of a biologically significant magnitude, can be misused readily. Researchers rarely have any control over the uses made of their data. Regardless of how many qualifications may be put on the reports of studies known to be inadequate, we are all aware that such findings can easily be taken out of context. Until we are sure of the professional training (and unfortunately, in at least a few cases, the ethical standards) of the possible users of our studies, we must face seriously the ethics of conducting inadequately supported studies. The unfortunate part of the problem, however, is that withholding ecological findings known to be inadequate merely leads public and private agencies to make and implement environmental decisions in a vacuum. The wheels do not stop turning without our input.

An unrelated point worthy of additional comment is the issue of variance estimation. All papers in this session address this issue in one way or another. Everyone agrees that replication is essential for a good study, but the biologists and statisticians at this conference do not agree on the best means for conducting these replicates. The question of the relationship of statistical independence to repeated censuses of the same area especially needs further consideration, although both Dawson (1981b) and Gates (1981) argue that such replication is legitimate and Morrison et al. (1981) and Blondel et al. (1981) demonstrate its usefulness.

We also need to clarify exactly what the variance is. A tendency exists among the current generation of ecologists to assume that each data point and each difference we observe is the result of deterministic processes, including both biological factors and the sources of bias that Verner itemized in his introduction to this session. Some long-term studies, especially recent works of Wiens and Rotenberry in shrub-steppe habitats (Wiens 1977, Wiens and Rotenberry 1979), and our work on the lower Colorado River riparian systems (Rice et al. In press, Ohmart et al. MS), demonstrate some substantial stochastic components to many of the population parameters we are trying to estimate with our censuses. It is essential for design considerations that biologists quantify the true magnitudes of these stochastic processes, and help to separate them from the tangle of sources of bias in our studies.