

SKUTCH, A. F. 1952. On the hour of laying of birds' eggs. *Ibis* 94: 49-61.
 ———. 1976. Parent birds and their young. London, Univ. Texas Press.
 WEATHERHEAD, P. J., & S. B. McRAE. 1990. Brood care

in American Robins: implications for mixed reproductive strategies by females. *Anim. Behav.* 39: 1179-1188.

Received 18 October 1990, accepted 22 March 1991.

Comments on the Sample Sizes Used to Test the Effect of Experimental Brood Enlargement on Adult Survival

JEFF GRAVES

*Department of Biology and Preclinical Medicine, University of St. Andrews,
 St. Andrews, Fife KY16 9TS, United Kingdom*

Demonstrating that an experimental manipulation does not have an effect on adult survival can frequently be difficult. In particular, a small sample size may be inadequate to have a reasonable chance of finding a real difference between experimental and control groups (type II error). In recent reviews of the literature on experimental studies of the cost of reproduction, Linden and Møller (1989), Dijkstra et al. (1990), and Nur (1990) cited studies of increased reproductive effort on adult mortality in 11 species that was caused by experimental enlargement of brood size. They reported that in only four of these species (Askenmo 1979; Nur 1984, 1988a; Reid 1987; Dijkstra et al. 1990) was there a significant increase in adult mortality for at least one sex subsequent to experimental brood enlargement. Because more experimental brood enlargement manipulations reduced adult fecundity in the next breeding attempt rather than increased adult mortality, it is tempting to conclude that the trade-offs that reduce adult fecundity are more likely than those that reduce adult survival (Linden and Møller 1989). Unfortunately, many of the studies that found no significant effect on adult mortality have sample sizes below that required to have even a reasonable probability of detecting a significant difference between the experimentals and the controls (i.e. inadequate statistical power).

Although several authors commented that small sample size results in a low probability of rejecting the null hypothesis (Nur 1988b, 1990; DeSteven 1980; Linden and Møller 1989), none of the studies that reported failure to find any significant increase in adult mortality after experimental brood mortality also calculated the statistical power of their results. There exist discussions of the problem of determining the sample size required to detect a significant difference between two populations (Fleiss 1981, Cohen 1988). Fleiss (1981) provided a simple formula that gives an approximation of the sample sizes required, and he also provided an appendix that lists the sample sizes necessary to achieve different levels of power with different significance levels and different pairs of proportions for the experimental and control groups. Co-

hen (1988) has the most extensive treatment, with tables for sample sizes and for calculating the power of any particular test.

Significance levels are now rigidly set by convention at 0.05 or less, but there is as yet no general agreement on the critical level of power. Toft and Shea (1983) recommend that β , the probability of type II error, be set at 0.05, so that the power is $1 - \beta$ (i.e. 0.95 or greater). This would require very large sample sizes; far in excess of what is usually possible in field studies, unless the difference between the two groups was very marked indeed. Fleiss (1981) and Cohen (1988) suggest as a reasonable compromise the criterion that type II error should be set at four times the level set for type I error. Thus, when the significance level is 0.05 the criterion for type II error would be set at $1 - 0.20$ (i.e. 0.80). The acceptance of such a criterion for the statistical power of a test would presuppose that type I errors are in general four times as serious as type II errors (Cohen 1988).

In the case of changes in adult mortality when brood size has been experimentally enlarged or reduced, there is usually some information beforehand on the expected mortality of the controls. Setting the significance level at the usual 0.05 and the type II at 0.80, a curve for the sample sizes required can be generated for any level of mortality for the experimental subjects. There is a very rapid increase in sample size required as the mortality of the experimentals approaches that of the controls (Fig. 1). If the control mortality is 40% and the experimental is 45%, then to have an 80% chance of detecting it at the 0.05 level of significance, 1,573 birds in both the groups will be required.

These results greatly affect the interpretation of most of the studies on brood enlargement and its effect on adult mortality. Of the eight studies that found no effect on adult mortality, six used sample sizes for experimental or control groups of 30 or fewer nests (DeSteven 1980, Harris 1970, Hegner and Wingfield 1987, Korpimäki 1988, Orell and Koivula 1988, Røskaft 1985). With sample sizes as small as this, the experimental birds would have to suffer a massive

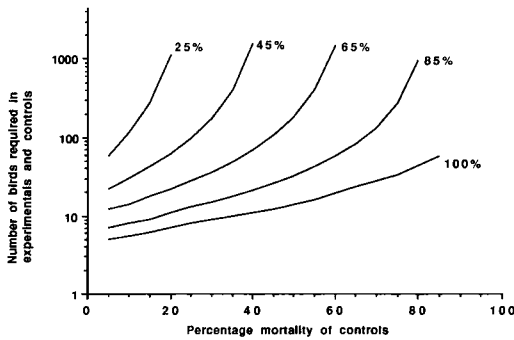


Fig. 1. Estimates of the sample sizes of birds required in each group to test the difference between percentage mortality of each group when α , the probability of type I error, is set at 0.05 and β , the probability of type II error, is set at 0.20. The percentage of mortality of the experimental group is given at the top of the curve.

increase in mortality before there would be an 80% chance of detection at the 0.05 level. If we assume that the observed differences in the proportions of mortality in the adults with normal and enlarged broods represents the true difference, it is possible to calculate the statistical power for these experiments from the tables in Cohen (1988) or the formula in Fleiss (1981). In all six cases the probability of a type II error is 0.9 or more, the statistical power is 0.1. Although several of the studies also provided data on reduced broods as well, I have treated only the two classes of control and enlarged. Finding sample sizes for more than two classes is difficult because of the great variety of possible deviations away from the expected (Sokal and Rohlf 1981).

Two of the studies that reported no effect of increased brood size on adult mortality used sample sizes in the hundreds, large enough to have an 80% chance of detecting a significant difference provided that the difference in the mortality of the two groups had been ca. 0.15 or greater (Gustafsson and Sutherland 1988, Pettifor et al. 1988). Because neither of these papers list the levels of mortality observed for any groups, it is not possible to calculate the power of the test. However, even these studies have sample sizes inadequate for detecting a difference of 0.10 or less at the 0.05 significance level with a probability of 0.8 or more. Both of these studies also collected data over several years, which lessens the possibility that any selection pressure against large broods might be relaxed in any one year (Nur 1988b). However, lumping data over different years might obscure a cost of reproduction that was present in some years but not in others.

Two of the three studies that found evidence of an effect on parental survival had very high levels of mortality in at least one sex. Askenmo (1979), in a

4-yr study, found that mortality in male Pied Flycatchers (*Ficedula hypoleuca*) increased from 62% for controls to 82% for males with enlarged broods. Nur (1984, 1988a) found in a 3-yr study that female Blue Tits (*Parus caeruleus*) had a significant increase in mortality with increasing brood size. Mortality was approximately 70% for very small clutches and 85% for large clutches. In both studies the clutch size was increased within the range of normal clutches. In the third study, Reid (1987) found that male and female Glaucous-winged Gull (*Larus glaucescens*) showed reduced survival if given supernormal broods. He combined the adult survival of all the supernormal-sized broods (4–7 chicks), and compared this with the adult survival rate of adults with normal-sized broods (1–3 chicks). This gave an adequate sample size of 319 adults with broods enlarged to supernormal size. It would also be of interest for life history studies to measure the effect the normal maximum brood size of three has on adult survival (Linden and Møller 1989). Reid found that mortality increased from 16.6% for adults with broods of one or two chicks to 21% for adults with broods of three chicks, but this difference was not significant. With a sample size of 105 adults with 3-chick broods, the probability of finding a significant difference at the 0.05 level for this difference in mortality of the two groups is only 20%. If this 4.4% increase was the true difference in mortality between adults with broods of one or two and those with three, then to have an 80% chance of finding it at the 0.05 level would require a sample size of >1,200 birds in both the experimental and control groups. There is also a sharp decrease in adult survival between parents with broods of 3 and those with broods of 4. Nur (pers. comm.) used linear logistic regression to reanalyze Reid's (1987) published data and found that there is a significant effect of brood size for broods of 1–4 chicks, excluding the supernormal broods of 5–7 chicks.

Because relatively small increases in mortality with increased brood size can be important in the analysis of life history traits, especially in long-lived species with a low level of annual mortality, it is important to try to use a sample size large enough to give a reasonably high probability of actually finding this difference if it exists. Because adequate sample sizes can run very high, this is obviously not possible in many field experiments. The alternative is to publish the power of the test that was possible as well as the exact probability value. Certainly reporting power when positive assertions are made from negative evidence is essential.

I thank A. Whiten for discussions about problems of sample size in fieldwork, and P.J.B. Slater, R. H. McCleery, A. Gordon, and an anonymous referee for comments. I also thank N. Nur for the very extensive comments and suggestions as a referee, and for bringing both his then unpublished chapter and Fleiss's book to my attention. This work was supported by

Natural Environment Research Council grant GR3/7249.

LITERATURE CITED

- ASKENMO, C. 1979. Reproductive rate and the return rate of male Pied Flycatchers. *Am. Nat.* 114: 748-753.
- COHEN, J. 1988. *Statistical power analysis for the behavioral sciences*, second ed. New Jersey, Lawrence Erlbaum Assoc.
- DESTEVEN, D. 1980. Clutch size, breeding success and parental survival in the Tree Sparrow (*Iridoprocne bicolor*). *Evolution* 34: 278-291.
- DIJKSTRA, C., A. BULT, S. BIJLSMA, S. DAAN, T. MEIJER, & M. ZIJLSTRA. 1990. Brood size manipulations in the Kestrel (*Falco tinnunculus*): effects on offspring and parent survival. *J. Anim. Ecol.* 59: 269-285.
- FLEISS, J. L. 1981. *Statistical methods for rates and proportions*, second ed. New York, J. Wiley and Sons.
- FORBES, L. S. 1990. A note on statistical power. *Auk* 107: 438-439.
- GUSTAFSSON, L., & W. J. SUTHERLAND. 1988. The costs of reproduction in the Collared Flycatcher *Ficedula albicollis*. *Nature* 335: 813-815.
- HARRIS, M. P. 1970. Breeding ecology of the Swallow-tailed Gull *Creagrurus furcatus*. *Auk* 104: 470-480.
- HEGNER, R. E., & J. C. WINGFIELD. 1987. Effects of brood-size manipulation on parental investment, breeding success, and reproductive endocrinology of House Sparrows. *Auk* 104: 470-480.
- KORPIMÄKI, E. 1988. Costs of reproduction and success of manipulated broods under varying food conditions in Tengmalm's Owl. *J. Anim. Ecol.* 57: 1027-1039.
- LINDEN, M., & A. P. MØLLER. 1989. Cost of reproduction and covariation of life history traits in birds. *Trends Ecol. Evol.* 4: 367-371.
- NUR, N. 1984. The consequences of brood size for breeding Blue Tits I. Adult survival, weight change and the cost of reproduction. *J. Anim. Ecol.* 53: 479-496.
- . 1988a. The consequences of brood size for breeding Blue Tits III. Measuring the cost of reproduction: survival, future fecundity and differential dispersal. *Evolution* 42: 351-362.
- . 1988b. The cost of reproduction in birds: an examination of the evidence. *Ardea* 76: 155-168.
- . 1990. The cost of reproduction in birds, evaluating the evidence from manipulative and non-manipulative studies. In *Population biology of passerine birds, an integrated approach* (J. Blondel, A. Gosler, J. D. Lebreton, and R. McCleery, Eds.). Berlin, Springer-Verlag Press.
- ORELL, M., & K. KOIVULA. 1988. Cost of reproduction: parental survival and production of recruits in the Willow Tit *Parus montanus*. *Oecologia* 77: 423-432.
- PETTIFOR, R. A., C. M. PERRINS, & R. H. MCCLEERY. 1988. Individual optimization of clutch size in Great Tits. *Nature* 336: 160-162.
- REID, W. V. 1987. The cost of reproduction in the Glaucous-winged Gull. *Oecologia* 74: 458-467.
- RØSKAFT, E. 1985. The effect of enlarged brood size on the future reproductive potential of the Rook. *J. Anim. Ecol.* 54: 255-260.
- SOKAL, R. R., & F. J. ROHLF. 1981. *Biometry*, second ed. San Francisco, W. H. Freeman and Co.
- TOFT, C. A., & P. J. SHEA. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. *Am. Nat.* 122: 618-625.

Received 19 September 1990, accepted 26 May 1991.

Courtship of Ducklings by Adult Male Chioe Wigeon (*Anas sibilatrix*)

GWEN BREWER¹

James Ford Bell Museum of Natural History, University of Minnesota, 10 Church Street SE, Minneapolis, Minnesota 55455 USA

Unlike most birds with seasonal pair bonds, many waterfowl pair much in advance of the next breeding season (Rohwer and Anderson 1988). For example, in Holarctic dabbling ducks (Anatini), courtship and pair formation occur up to 8 months before breeding begins (Hepp and Hair 1983). During the early period of bond formation, males court females and try to

lead them away from other birds. When females show their preference for one male and rejection of others by giving inciting displays, temporary associations between males and females are formed. As bonds strengthen, the members of a pair maintain close proximity, synchronize their activities, repel rivals, and perform displays that reinforce the bond (McKinney in press). Pair bonds continue to be tested during spring migration and after arrival on the breeding grounds, although most birds are paired by this time.

¹ Present address: Department of Biology, Frostburg State University, Frostburg, Maryland 21532 USA.