

DESIGN AND DURATION OF PERTURBATION EXPERIMENTS: IMPLICATIONS FOR DATA INTERPRETATION

KENNETH L. PETERSEN AND LOUIS B. BEST

Abstract. Studies of effects of habitat perturbations on birds sometimes lack adequate controls, pre-treatment data, or long-term postdisturbance data. We studied the effects of a prescribed fire on a bird community in shrubsteppe habitat in southeastern Idaho. The study comprised two years preburn and seven years postburn data collection on experimental and control plots. Interpretation of our results varies, depending on inclusion or exclusion of data from control plots, inclusion or exclusion of preburn data, and the number of years of postburn data incorporated in the analysis. For example, the Brewer's Sparrow (*Spizella breweri*) population declined the first two years after burning but subsequently showed a neutral response. Without control data, we would have concluded that fire had a negative impact on Sage Sparrow (*Amphispiza belli*) numbers when in fact the population also declined on unburned controls. Several fire-induced changes in nest-site selection by Brewer's and Sage sparrows would not have been detected without preburn data. Inadequate study design and duration can lead to inaccurate conclusions and misdirected conservation efforts.

EL DISEÑO Y LA DURACIÓN DE EXPERIMENTOS DE PERTURBACIÓN: EFECTOS EN LA INTERPRETACIÓN DE DATOS

Sinopsis. A veces los estudios de los efectos de las perturbaciones de hábitat en las aves carecen de controles adecuados, de datos previos, o de datos a largo plazo posteriores al cambio. Estudiamos los efectos de un fuego programado en una comunidad de aves en un hábitat de estepa arbustiva en el sureste de Idaho. El estudio comprendió dos años de recolección de datos antes del fuego, y siete años después del fuego en parcelas experimentales y de control. La interpretación de nuestros resultados varía según la inclusión o la exclusión de los datos de las parcelas de control, según la inclusión o la exclusión de datos previos al fuego, y según el número de años de datos posteriores al fuego que se incluyan en el análisis. Por ejemplo, la población del Gorrión de Brewer (*Spizella breweri*) disminuyó durante los primeros dos años después de la quema pero luego registró una respuesta neutra. Sin datos de control habríamos concluido que el fuego tenía un efecto negativo en los números del Gorrión de Artemisia (*Amphispiza belli*), cuando de hecho la población disminuyó también en los controles sin quema. El fuego produjo varios cambios en la selección de los Gorriónes de Brewer y de Artemisia de los sitios de nidos que no habríamos descubierto sin los datos previos al fuego. Un estudio con diseño y duración inadecuados puede traducirse en conclusiones erróneas y en tentativas de conservación mal encaminadas.

Key Words: *Amphispiza belli*; Brewer's Sparrow; data interpretation; experimental design; prescribed fire; Sage Sparrow; shrubsteppe; *Spizella breweri*.

Because so many grasslands have been destroyed or altered, studies of how various habitat perturbations affect grassland bird populations play an important role in bird conservation. The most powerful design for such studies is to collect pre- and postdisturbance data from control and disturbed areas (Stewart-Oaten et al. 1986) over a sufficiently long period to reveal the dynamics of the system (Wiens 1989). Historically, however, many studies have been deficient in one or more of these requirements. For example, we surveyed the 1994 and 1995 issues of the *Auk*, *Journal of Field Ornithology*, *Condor*, *Wilson Bulletin*, and *Journal of Wildlife Management*. We identified 35 papers that investigated effects of natural or artificial habitat alterations (including food additions or deletions) on some aspect of bird biology. Of these, 33 (94%) either did not sample pre-disturbance conditions, had no control sites, or were brief (< 5 yr) in duration. Twenty-five (71%) of the 35 studies were

deficient in at least two of these design aspects. The most common weakness was short-term data collection; 21 of the studies spanned 2 yr or less, and 12 were based on only 1 yr of data collection.

These shortcomings occur not necessarily because investigators are unaware of the elements of study design but because they often are faced with unavoidable limitations and therefore have few options in planning their studies. For example, a disturbance such as a fire, storm, or human-caused disaster may afford an unexpected opportunity for study, but unless the site is already being monitored, no pre-disturbance data are available (e.g., Bowman et al. 1995, Latta et al. 1995). Or a study may be of limited duration because of funding constraints. Although investigators should not forego the opportunity to study habitat perturbations in such situations, they must exercise caution in interpreting the results. This paper demonstrates potential errors

that may occur in data interpretation when any one of the aforementioned requirements of research design is not met.

STUDY AREA AND METHODS

We studied the effects of prescribed fire on a bird community in shrubsteppe habitat in southeastern Idaho (Petersen and Best 1987). The study area was about 11 km south of Howe, Butte County, Idaho, and the vegetation was dominated by big sagebrush (*Artemisia tridentata*) and bunchgrasses (e.g., *Agropyron*, *Oryzopsis*, *Sitanion*). The study comprised 2 yr (1980–1981) of preburn data and 7 yr (1982–1988) of postburn data collected on two experimental and two control plots. Each plot measured 250 × 250 m, was gridded at 25-m intervals to facilitate recording bird observations and nest locations, and was large enough to encompass 6–12 average-sized territories of most of the songbird species that inhabited the study area. Plots were positioned far enough apart (200–300 m) so that bird territories on one plot did not overlap those on another.

The fire substantially reduced mean coverage of sagebrush on experimental plots and moderately stimulated development of herbaceous vegetation (Petersen and Best 1987). Even 7 yr after burning, there was no evidence of sagebrush recovery on burned sites. Big sagebrush does not resprout after fire, and fire-induced changes in coverage may persist for more than 10 yr (Wright and Bailey 1982). In contrast, coverage of shrubs and herbaceous vegetation on control plots did not vary significantly over the course of the study. Thus, we were confident that the significant variations we documented in the bird community, even several years after the fire, could reasonably be attributed to burning.

We censused bird populations in June each year by delineating territories either by spot mapping (International Bird Census Committee 1970) or by using the "flush" technique (Wiens 1969). Both techniques involved recording positions and movements of individuals on maps of the grids. Each plot was visited at least five times in each year, typically for 3–4 hr each visit. In 1980–1984 we captured and color-banded Sage Sparrows (*Amphispiza belli*) and Brewer's Sparrows (*Spizella breweri*) to facilitate individual identification (Petersen and Best 1987); we banded 50–100% of the territorial Sage Sparrow males and 30–75% of the territorial Brewer's Sparrow males in each of these years. By the end of each season's census, we were confident that we had accounted for every territorial male. Population densities were determined by counting territories (and fractions thereof) lying within plot boundaries.

We also systematically searched all plots for nests during the first 5 yr (1980–1984) of the study (Petersen and Best 1985a, b, 1991), but only two species—Sage and Brewer's sparrows—were present in sufficient numbers to afford analysis of nest success and nest-site selection in response to fire. Active nests were visited at 1- to 2-d intervals, and nest success was calculated by Mayfield's (1975) method. Nest-site characteristics were measured after nests (either successful or failed) had been abandoned. These measurements included nest height, height of the shrub

supporting the nest, and coverage and average height of sagebrush within a 5-m radius of the nest.

We employed one-way analysis of variance (ANOVA) to ascertain whether, within each treatment, a parameter (e.g., population density or nest success) of a given species varied significantly among years of the study. In analyses of nest success, a significant ANOVA result was followed by Duncan's multiple range test to determine which years differed significantly from each other. We used t-tests to determine the significance of between-treatment differences in population parameters of a given species in each year of the study. Statistical significance was set at $P \leq 0.05$.

RESULTS AND DISCUSSION

IMPLICATIONS OF SHORT-TERM DATA COLLECTION

Drawing firm conclusions about bird responses to habitat perturbation from only one or a few years of postdisturbance data may be unwise. Analyses of our population data for Brewer's Sparrows in 2 yr pre- and 2 yr postburn (1980–1983) show that fire had an immediate negative impact on this species, reducing densities on experimental plots by about 50% (ANOVA: $F_{(3,4)} = 21.51$; Fig. 1). Brewer's Sparrow densities on control plots did not vary significantly over these years. Examination of all 9 yr of data, however, shows that the negative effect of fire was short-lived; densities increased substantially on experimental plots (ANOVA: $F_{(8,9)} = 21.66$) after 1983 but did not vary significantly on controls. Moreover, in 1985 and 1988 densities on experimental plots averaged significantly higher than on controls ($t_{1985} = 121.00$, $df = 2$; $t_{1988} = 6.00$, $df = 2$), a seeming reestablishment of the preburn pattern in which densities on experimental plots also exceeded those on controls. Thus, the long-term impact of fire seemed to be neutral for this species.

Similarly for the Sage Thrasher (*Oreoscoptes montanus*), short-term (1980–1982) data suggest that fire had no impact after two pre- and one postburn seasons (Fig. 2). There were no significant variations in densities on either treatment over these 3 yr, nor did densities differ significantly between treatments. When the long-term data are included in the analysis, however, densities were consistently higher on the burned plots (although the difference between treatments was significant only in 1985 [$t = 9.00$, $df = 2$]), suggesting that the effects of fire stimulated the population.

Vesper Sparrows (*Poocetes gramineus*) were not present on the study area in the first 4 yr of the study (two pre- and two postburn years) and then abruptly appeared on burned plots in the third postburn year and were present thereafter (Fig. 3). Thus, a study of as much as 4 yr duration would not have led to an accurate conclusion about the composition of this postburn bird

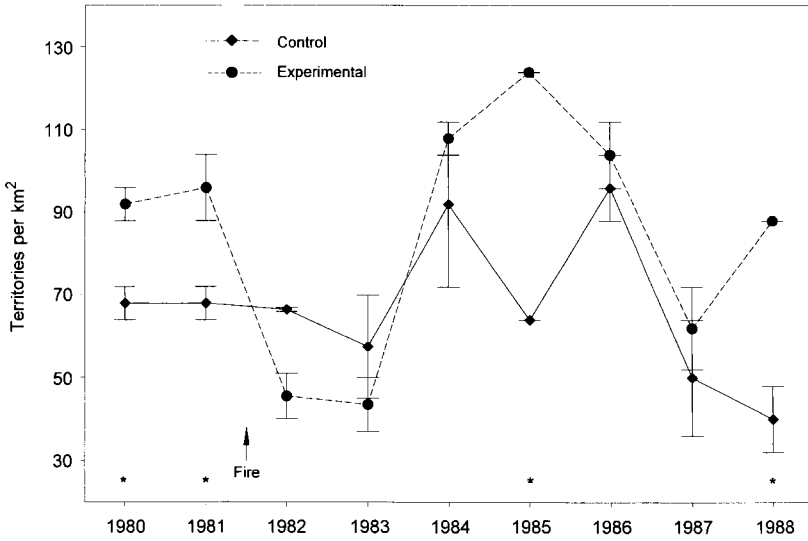


FIGURE 1. Population densities of Brewer's Sparrows on experimental and control plots in southeastern Idaho. Experimental plots were burned in the fall of 1981. Each data point and associated error bars represent values from two study plots ($\bar{X} \pm SE$). Asterisks signify that treatments differed significantly ($P \leq 0.05$, t-test) in particular years.

community. In our survey of 35 papers, only 7 (20%) examined data from 5 yr or more of study. We can be confident that the pattern observed in this instance reflected a perturbation effect because Vesper Sparrows consistently settled on burned plots after 1983 yet never did so on unburned controls. Furthermore, because the

two treatments were close enough to each other to be influenced by the same range of natural phenomena, it is difficult to imagine a factor other than the fire that would have produced this effect.

Our data on Western Meadowlarks (*Sturnella neglecta*) show that even 5 yr may not be suf-

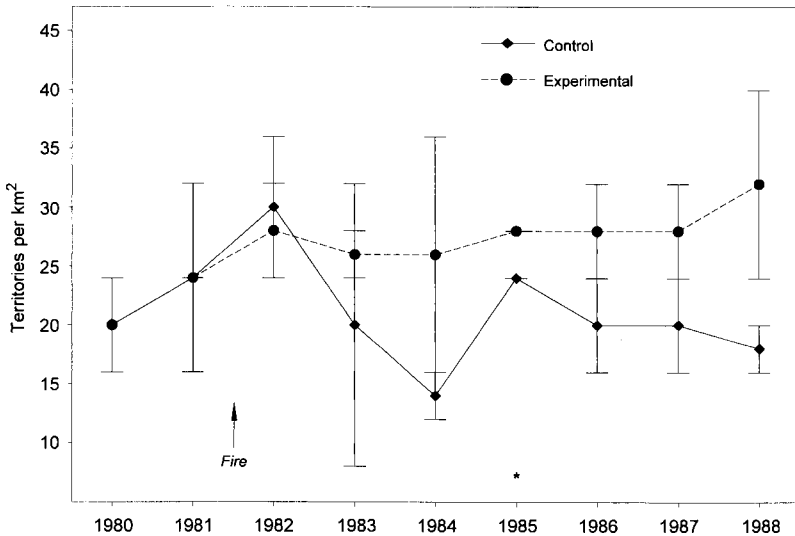


FIGURE 2. Population densities of Sage Thrashers on experimental and control plots in southeastern Idaho. Experimental plots were burned in the fall of 1981. Each data point and associated error bars represent values from two study plots ($\bar{X} \pm SE$). Control and experimental means were identical in 1980 and 1981. The asterisk signifies that treatments differed significantly ($P \leq 0.05$, t-test) in 1985.

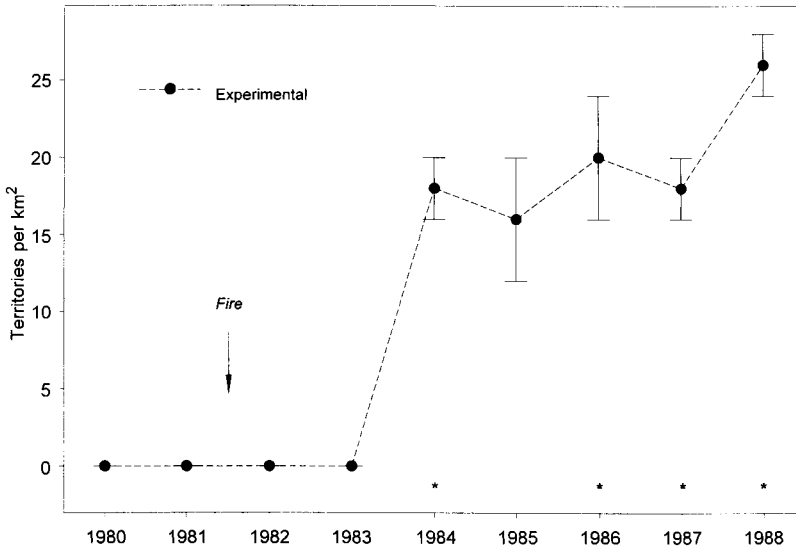


FIGURE 3. Population densities of Vesper Sparrows on experimental plots in southeastern Idaho, burned in the fall of 1981. Each data point and associated error bars represent values from two study plots ($\bar{X} \pm SE$). Asterisks signify that experimental means differed significantly ($P \leq 0.05$, t-test) from controls in particular years. Control plots supported no Vesper Sparrows during this time.

ficient to assess the long-term impact of disturbance. If our study had spanned only 5 yr (1980–1984), we likely would have concluded that fire had a mildly positive effect on Western Meadowlarks (Fig. 4). Population densities increased significantly (ANOVA: $F_{(4,5)} = 6.22$) on

experimental plots over this time but did not vary significantly on controls. And in 1984, mean density on burned plots was nearly significantly greater ($P = 0.09$) than on controls ($t = 3.00$, $df = 2$). Analysis of all 9 yr of data, however, does not seem to suggest a long-term in-

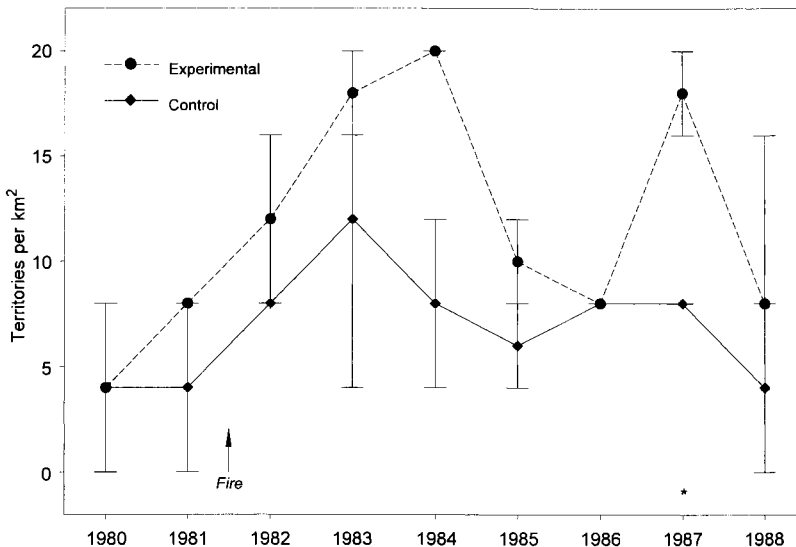


FIGURE 4. Population densities of Western Meadowlarks on experimental and control plots in southeastern Idaho. Experimental plots were burned in the fall of 1981. Each data point and associated error bars represent values from two study plots ($\bar{X} \pm SE$). Control and experimental means were identical in 1980 and 1986. The asterisk signifies that treatments differed significantly ($P \leq 0.05$, t-test) in 1987.

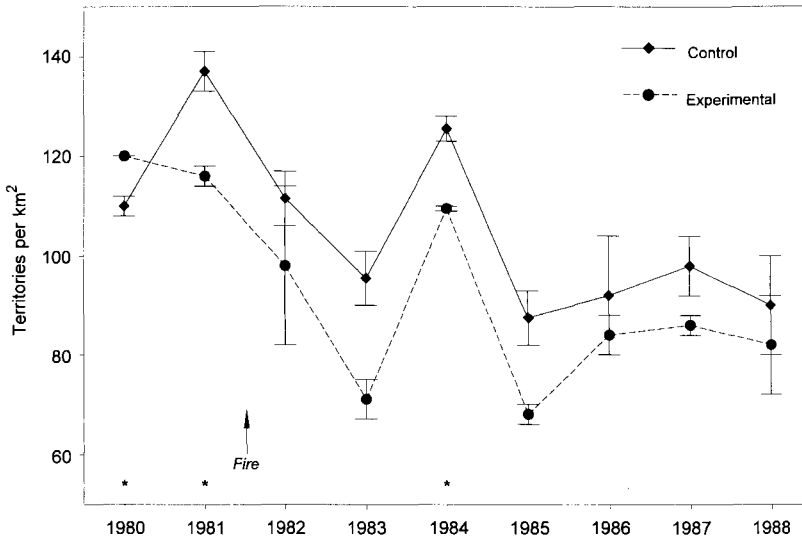


FIGURE 5. Population densities of Sage Sparrows on experimental and control plots in southeastern Idaho. Experimental plots were burned in the fall of 1981. Each data point and associated error bars represent values from two study plots ($\bar{X} \pm SE$). Asterisks signify that treatments differed significantly ($P \leq 0.05$, t-test) in particular years.

fluence of fire. Although mean density on experimental plots was greater than on controls in 1987 ($t = 5.00$, $df = 2$), density did not vary significantly on plots of either treatment when all 9 yr were incorporated in the analysis.

IMPLICATIONS OF NO CONTROLS

Although most (26 of 35) studies we surveyed included data from control areas, control data may occasionally be absent. For example, a

TABLE 1. DAILY SURVIVAL PROBABILITIES ($\bar{X} \pm SE$, $N = 2$ PLOTS)^a FOR BREWER'S SPARROW NESTS ON CONTROL AND EXPERIMENTAL (BURNED) PLOTS IN SOUTHEASTERN IDAHO

Year	Plots	
	Control	Experimental
1980	0.958 \pm 0.042 (41%) ^c	0.992 \pm 0.008 A ^b (84%)
1981	0.926 \pm 0.046 (20%)	0.970 \pm 0.018 AB (53%)
----- Fire -----		
1982	0.978 \pm 0.022 (63%)	0.932 \pm 0.009 C (23%)
1983	0.992 \pm 0.008 (84%)	1.000 \pm 0.000 A (100%)
1984	0.955 \pm 0.015 (38%)	0.956 \pm 0.002 BC (39%)

^a Number of nests sampled per year per treatment ranged from 6 to 18.
^b Means in the same column with no letters in common are different ($P \leq 0.05$, one-way ANOVA, Duncan's multiple range test).
^c Estimated nest success based on a nesting-cycle length of 21 d.

change or disturbance may occur unexpectedly on an area from which some data had already been collected (e.g., Miller et al. 1994, Hestbeck 1995, Wunderle 1995).

From an examination of Sage Sparrow densities on experimental plots alone, we would be strongly inclined to conclude that fire had a negative impact, at least in the short term (Fig. 5). Except for a spike in 1984, this conclusion could be drawn from the long-term data as well; densities varied significantly over 9 yr (ANOVA: $F_{(8,9)} = 8.19$). When control data are included, however, it is clear that, after the fire, changes in densities on control and experimental plots paralleled each other (9-yr ANOVA on controls: $F_{(8,9)} = 6.66$). Although mean density on controls exceeded that on burned plots in 1984 ($t = 6.28$, $df = 2$), this pattern also existed in 1981 before the fire ($t = 4.70$, $df = 2$). Therefore, the conclusion to be drawn from examination of both control and experimental data is that fire had no effect.

Similarly, patterns of change in Brewer's Sparrow nest success show that lack of control data might lead to an incorrect conclusion. Success declined on experimental plots immediately after burning, but there was also a decline on experimental plots between the two preburn years (Table 1). Nest success rebounded strongly in 1983 but then declined again. Without control data, one would be uncertain as to how much of this variation was attributable to fire. When data

TABLE 2. NEST-SITE CHARACTERISTICS ON CONTROL AND EXPERIMENTAL (BURNED) PLOTS IN SOUTHEASTERN IDAHO

	Sage Sparrow				Brewer's Sparrow			
	Control		Experimental		Control		Experimental	
	Preburn	Postburn	Preburn	Postburn	Preburn	Postburn	Preburn	Postburn
Nest height (cm)	32	29	38 ^a	28	34	34	41 [*]	32
Substrate height (cm)	61 ^{**b}	60	72 [*]	61	61	65	74	64
Sagebrush height (cm)	43	40	52 [*]	39	49	47	58	44
Sagebrush coverage (%)	24	24	32 [*]	18	31	26 ^{**}	32	15

Note: Each value is a mean from two plots. Standard errors have been omitted for clarity. The number of nests sampled preburn and postburn per treatment ranged from 23 to 60 for Sage Sparrows and from 11 to 28 for Brewer's Sparrows.

^a A single asterisk signifies that preburn differs significantly ($P \leq 0.05$, t-test) from postburn in either control or experimental.

^b Double asterisks signify that control differs significantly ($P \leq 0.05$, t-test) from experimental in either pre- or postburn.

from control plots are included, however, the pattern becomes clearer. Now, because of the immediate postburn increase (although statistically insignificant) on controls, it seems likely that the decline on experimental plots was fire induced. And it is evident that the pattern of the subsequent 2 yr (1983–1984) was independent of fire.

IMPLICATIONS OF NO PRE-ALTERATION DATA

Some habitat perturbations may occur unexpectedly on sites not previously studied. Although one would have no predisturbance data, the investigator still might wish to take advantage of the opportunity to collect postdisturbance data (e.g., Bowman et al. 1995). In our survey, 24 studies (69%) lacked pre-alteration data.

Analysis of Sage and Brewer's sparrow nest-site selection provides a good example of potential misinterpretation. In comparing controls against burned plots (postburn data only), nest height, substrate height, and height of sagebrush around the nest were virtually the same (Table 2). We might conclude from postburn data alone that, with the possible exception of sagebrush coverage, fire had essentially no effect on nest-site selection. When we include preburn data, however, we draw a different conclusion. On experimental plots, postburn means were significantly smaller than preburn means in several instances, but there were no significant postburn changes on control plots. Thus, fire did seem to influence nest-site selection.

CONCLUSIONS

We must design perturbation experiments as any good scientific experiment is designed: with adequate controls, with pre- as well as postdisturbance data, with replicated treatments, and with a time scale that is as long as we can reasonably make it. Then we may draw conclusions with some confidence, and we may be able to develop and apply management if it is needed. Deciding what "long-term" means may be difficult, but in general, the duration of a pertur-

bation study should be proportional to the life span of the birds being studied (Calder 1984, Wiens 1984). We showed that even 2–3 yr of postdisturbance data collection may be insufficient for short-lived passerines. The term of a perturbation study also will depend on the rate of habitat change after disturbance. If an altered site changes quickly toward the predisturbance condition or some other stable configuration (as may be the case in some eastern North American grasslands), comparatively few years will be required to assess the effects of the perturbation on the bird community. Where disturbance cycles and recovery times are long (as, for example, in sagebrush shrubsteppe), longer-term study is required.

Determining what constitutes an appropriate control also is not always straightforward. In a perturbation study aimed at generalizing average or "usual" effects of a disturbance to a large set of possible instances, treatments should be randomly assigned to study plots. Otherwise the experiment is controlled only in a subjective and approximate way (Hurlbert 1984). This was a limitation of our own study as well as of many others because, realistically, such randomization is often not possible. Random allocation of treatments to study plots may not be necessary, however, if the concern is with a particular impact in a particular place resulting from a particular disturbance (i.e., impact assessment; Stewart-Oaten et al. 1986).

Similarly, establishing replication in a perturbation experiment is not necessarily simple. First, care must be taken to avoid pseudoreplication, in which, for example, nests might be replicated but the actual experimental units (e.g., study plots) are not (Hurlbert 1984). Second, the number of replicates must be sufficient to enable investigators to infer treatment effects when such effects occur. The power of most inferential statistical tests to detect treatment effects increases as sample size increases, and determination of sample sizes adequate to ensure detec-

tion of treatment effects requires power analysis. Such analysis should occur before beginning a study because the role and utility of power analysis after data have been collected and analyzed are controversial (Thomas 1997). The consequences of failing to statistically detect a positive or negative perturbation effect may be as critical for bird conservation as the consequences of inferring perturbation effects when none exist. In our study, the necessity of accurately and comprehensively sampling all study sites, given available time and personnel, constrained us to two replicate plots per treatment. Because of this small sample size, we possibly failed to detect some fire effects that did in fact occur. Nevertheless, we were able to observe several significant trends and differences that we could not have documented without replication of study plots.

Our limited literature survey suggests that it is frequently difficult to incorporate all of the aforementioned aspects of study design. When this is true, however, it is not necessarily wise to abandon plans to collect data. Because opportunities to conduct well-designed perturbation experiments are scarce, we should take advantage of any situation (e.g., an unexpected wildfire) that can provide some information. Surely even studies that lacked predisturbance data, were short term, etc., have contributed to our understanding of bird biology in a disturbed landscape. As we have shown, however, such situations require caution in data interpretation. We cannot afford, for example, to conclude that a disturbance had a positive impact when in fact there was no effect or perhaps even a negative effect. This, of course, would lead to an inappropriate stewardship strategy that would waste time, energy, and money and might even be detrimental to the species or community we were trying to conserve. Unfortunately, regardless of the soundness of the design, there will always be uncertainty in data interpretation. This is so because, in the natural world, many factors cannot be controlled even by the most careful design. At best, we can design research in such a way as to minimize the probability of reaching false conclusions.

ACKNOWLEDGMENTS

P. Sievert and L. Erickson-Eastwood contributed substantially to field data collection. Field work was funded by the Office of Health and Environmental Research, U.S. Department of Energy. We thank O. D. Markham for the opportunity to use facilities and field

sites of the Idaho National Engineering Laboratory. This paper benefited greatly from the comments of two anonymous reviewers.

LITERATURE CITED

- BOWMAN, T. D., P. F. SCHEMPF, AND J. A. BERNATOWICZ. 1995. Bald Eagle survival and population dynamics in Alaska after the Exxon Valdez oil spill. *Journal of Wildlife Management* 59:317-324.
- CALDER, W. A., III. 1984. How long is a long-term study? *Auk* 101:893-894.
- HESTBECK, J. B. 1995. Response of Northern Pintail breeding populations to drought, 1961-92. *Journal of Wildlife Management* 59:9-15.
- HURLBERT, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187-211.
- INTERNATIONAL BIRD CENSUS COMMITTEE. 1970. Recommendations for an international standard for a mapping method in bird census work. *Audubon Field Notes* 24:723-726.
- LATTA, S. C., J. M. WUNDERLE, JR., E. TERRANOVA, AND M. PAGAN. 1995. An experimental study of nest predation in a subtropical wet forest following hurricane disturbance. *Wilson Bulletin* 107:590-602.
- MAYFIELD, H. F. 1975. Suggestions for calculating nesting success. *Wilson Bulletin* 87:456-466.
- MILLER, C. K., R. L. KNIGHT, L. C. McEWEN, AND T. L. GEORGE. 1994. Responses of nesting Savannah Sparrows to fluctuations in grasshopper densities in interior Alaska. *Auk* 111:962-969.
- PETERSEN, K. L., AND L. B. BEST. 1985a. Brewer's Sparrow nest-site characteristics in a sagebrush community. *Journal of Field Ornithology* 56:23-27.
- PETERSEN, K. L., AND L. B. BEST. 1985b. Nest-site selection by Sage Sparrows. *Condor* 87:217-221.
- PETERSEN, K. L., AND L. B. BEST. 1987. Effects of prescribed burning on nongame birds in a sagebrush community. *Wildlife Society Bulletin* 15:317-329.
- PETERSEN, K. L., AND L. B. BEST. 1991. Nest-site selection by Sage Thrashers in southeastern Idaho. *Great Basin Naturalist* 51:261-266.
- STEWART-OATEN, A., W. W. MURDOCH, AND K. R. PARKER. 1986. Environmental impact assessment: "pseudoreplication" in time? *Ecology* 67:929-940.
- THOMAS, L. 1997. Retrospective power analysis. *Conservation Biology* 11:276-280.
- WIENS, J. A. 1969. An approach to the study of ecological relationships among grassland birds. *Ornithological Monographs* 8:1-96.
- WIENS, J. A. 1984. Editorial. The place of long-term studies in ornithology. *Auk* 101:202-203.
- WIENS, J. A. 1989. *The ecology of bird communities*. Vol. 2. Processes and variations. Cambridge University Press, New York, NY.
- WRIGHT, H. A., AND A. W. BAILEY. 1982. *Fire ecology: United States and southern Canada*. J. Wiley and Sons, New York, NY.
- WUNDERLE, J. M., JR. 1995. Responses of bird populations in a Puerto Rican forest to Hurricane Hugo: the first 18 months. *Condor* 97:879-896.