University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Faculty Publications in the Biological Sciences | Papers in the Biological Sciences

10-1-2004

Spatial Heterogeneity, Not Visitation Bias, Dominates Variation in Herbivory: Comment & Reply

James F. Cahill
University of Alberta

Brenda B. Casper
University of Pennsylvania

David S. Hik
University of Alberta

Svata M. Louda
University of Nebraska - Lincoln, slouda1@unl.edu

Anne M. Parkhurst
University of Nebraska - Lincoln, aparkhurst1@unl.edu

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/bioscifacpub

Part of the Life Sciences Commons

Cahill, James F.; Casper, Brenda B.; Hik, David S.; Louda, Svata M.; Parkhurst, Anne M.; Bradley, Kate L.; Bakker, Elisabeth S.; Knops, Johannes M. N.; Damschen, Ellen I.; and Young, Lauren M., "Spatial Heterogeneity, Not Visitation Bias, Dominates Variation in Herbivory: Comment & Reply" (2004). *Faculty Publications in the Biological Sciences.* Paper 64.
http://digitalcommons.unl.edu/bioscifacpub/64

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications in the Biological Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
James F. Cahill, Brenda B. Casper, David S. Hik, Svata M. Louda, Anne M. Parkhurst, Kate L. Bradley, Elisabeth S. Bakker, Johannes M. N. Knops, Ellen I. Damschen, and Lauren M. Young

This article is available at DigitalCommons@University of Nebraska - Lincoln: http://digitalcommons.unl.edu/bioscifacpub/64

Comments

Ecology, 85(10), 2004, pp. 2901–2906
© 2004 by the Ecological Society of America

SPATIAL HETEROGENEITY, NOT VISITATION BIAS, DOMINATES VARIATION IN HERBIVORY: COMMENT

James F. Cahill, Jr.,[1,3] Brenda B. Casper,[2] and David S. Hik[1]

[1]*Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada*
[2]*Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

Making scientific measurements can alter the subject of the measurement (Heisenberg 1927), limiting one's ability to understand how systems operate when they are not being studied. The idea that such effects occur during ecological research is not new (e.g., Singer and Wedlake 1981, Major 1990, Duncan et al. 2002), but our recent studies have renewed awareness (Cahill et al. 2001, 2002, Hik et al. 2003). Visitation by researchers to focal plants can alter herbivory and plant growth (Cahill et al. 2001, 2002) due to both effects of handling focal plants and trampling neighbors (Cahill et al. 2002). ''Visitation effects'' occur in a variety of communities, appear widespread among species, and are generally of low intensity (Hik et al. 2003).

Presumably because these prior results have such important implications for the interpretation of ecological studies, other groups have tested the generality of our findings by replicating aspects of our initial experiment (Cahill et al. 2001) in other systems. In Cedar Creek, Minnesota, USA, Schnitzer et al. (2002) found that weekly visitation increased pathogen damage on some plants but had no effect on herbivory. Bradley et al. (2003) similarly conclude there is little evidence that weekly visitation alters herbivory or plant growth in three U.S. locations (Cedar Creek, Savannah River [South Carolina], and Reller Ranch/Madigan Prairie [Nebraska]).

Here we show that the experimental design and statistical analysis used by Bradley et al. (2003) do not adequately control for Type II experimental error (accepting the null hypothesis when it is in fact false). This is particularly important when responses have low effect sizes such as those previously demonstrated for visitation effects (Hik et al. 2003). Type II errors are of utmost concern with respect to confounding effects of experimental methodology, as their undetected effects greatly hinder understanding the ecology of the system under investigation.

We use the study of Bradley et al. (2003) to illustrate the impact that alternative interpretations of experimental design and statistical analysis have on the resulting conclusions. We describe specific shortcomings of the methodological choices made by Bradley et al. and show that a more appropriate statistical model indicates their data are consistent with our prior findings.

Bradley et al. (2003) experimental design

Methodological details are found within Bradley et al. (2003), and only a brief overview is presented here. Five plant species at each of the three U.S. locations Nebraska (NE), South Carolina (SC), and Minnesota (MN) were used, with species generally varying among locations (14 unique species). In NE and SC, all species occurred in two fields, while in MN each species occurred in 2–3 of six fields. Half of the 20–72 individuals of each species were assigned weekly visits and half were left unvisited as controls. The experiments were initiated over a three-week period with the most northern location (MN) started first. At each location, aboveground biomass was harvested after eight weeks, and a variety of other response variables were measured. Throughout this paper we focus only on two of their measures, shoot biomass and the proportion of leaves damaged by herbivory, as these variables are most similar to those examined in prior visitation studies (e.g., Hik et al. 2003), and both were measured for all species in all locations by Bradley et al. However, the issues presented here are applicable to all of the response variable analyzed by Bradley et al.

Bradley et al. (2003) statistical analysis

Bradley et al. analyzed each response variable separately for each species as a function of treatment (fixed effect) and field (random effect). Four response variables were measured for most species, resulting in 57 general linear model analyses. Bradley et al. (2003: 2217) viewed their study as a single experiment and used the Dunn-Sidák procedure (Sokal and Rohlf 1997) to lower the type I error rate (rejecting the null when it is in fact true) in order to ''limit the experimental error rate.'' As a result, individual results were viewed as significant if $P < 0.0009$.

Manuscript received 25 November 2003; revised and accepted 8 March 2004. Corresponding Editor: B. Shipley.
[3] E-mail: jc.cahill@ualberta.ca

Implications of this statistical model

A single experiment vs. multiple experiments.—Conducting 57 separate analyses is inconsistent with the idea of a single experiment performed in multiple locations. A more appropriate analysis would include treatment, location, species, and field as factors in the model, with field and species nested within location. This approach allows for testing for an overall treatment effect, as well as interaction terms (e.g., species × treatment). Post hoc tests could be conducted if a significant interaction is found, with *P* values adjusted as needed. Because the response to visitation varies among species, suggesting the initial title of ''The Herbivory Uncertainty Principle'' (Cahill et al. 2001:307), a direct test of species × treatment interactions seems warranted. Such a model is particularly appropriate with respect to visitation effects where low effect sizes can result in significant species × treatment interactions even without significant species-specific comparisons (Hik et al. 2003). This approach allows one to reduce the risk of inflating Type I error rate by reducing the number of tests performed. A disadvantage is that higher-order interactions may be difficult to interpret, although this is a consequence of the experimental design, rather than the analysis. If the experimental design allows for interactions to occur (e.g., field × treatment), the statistical model should allow for their detection.

The experimental methods used in the three locations differed enough that the study is justifiably viewed as three separate experiments, rather than a single study. Specifically, locations differed in (1) the species examined, (2) the number of fields used, (3) the response variables measured, and (4) the timing of the study relative to local phenology. These differences are important because visitation effects are species specific (Cahill et al. 2001, Hik et al. 2003) and because indirect effects of visitation, such as those caused by trampling neighboring plants, should depend upon the composition and phenology of the plant community. Trampling of emerging seedlings is likely of different ecological consequence than trampling of older and more established plants. Standardizing the timing of each experiment to the local growing season would minimize this latter problem (sensu Hik et al. 2003). Viewing this study as three experiments necessitates separate analyses for each location (= experiments), with field, treatment, species, and all possible interaction terms in each statistical model. Post-hoc tests could be conducted if interaction terms are significant. In MN, not every species occurred in every field, and thus empty cells exist in the ANOVA model, which could be dealt with by using Type IV, rather than Type III sums of squares.

This general approach has been used before with respect to visitation effects. Hik et al. (2003) treated each location as a separate experiment, although they used only one field in each location. Schnitzer et al. (2002) used site as a factor in their model, but treated species as a random, rather than a fixed effect. They, too, had missing cells in their analysis due to nonuniform species distributions among fields.

Adjusting P values.—The question of when and how to adjust alpha is a matter of debate (Cabin and Mitchell 2000, Moran 2003). The approach used by Bradley et al. (2003) maintains an experiment-wide Type I error rate at 5% by modifying the alpha value (Type I error rate) for individual tests using the Dunn-Sidák procedure (similar to the more conservative Bonferroni correction). While reasonable for a small number of comparisons, both procedures are extremely conservative in larger numbers of comparisons (Sokal and Rohlf 1997), as clearly occurred in Bradley et al.'s approach ($P < 0.0009$). Decreasing alpha also decreases statistical power, and thus a low Type I error rate on individual tests increases the likelihood of accepting null hypotheses when they are actually false (Type II error). Type II errors are particularly likely when effect sizes are small, such as in the case of visitation effects. Although ecologists have generally given more consideration to Type I than Type II errors, there is no a priori justification for this bias (DiStefano 2003). There are many situations in which Type II rather than Type I errors are of primary concern, but this is not widely acknowledged amongst ecologists (DiStefano 2003). We suggest that the failure to document any inadvertent confounding effects of experimental methodology, such as the effects of researcher visitation on plant growth and herbivory, decreases the intellectual value of a study. As a result, there is a great cost to making Type II errors in studies of research methodology such as that of Bradley et al. (2003).

Fortunately, there exist several alternatives to adjusting *P* values and increasing Type II error rates. The preferred method would be to limit the number of tests conducted to those addressing the specific hypotheses of interest. In the case of Bradley et al., this would be achieved through conducting analyses with species as a factor, as suggested above, rather than separate analyses for each species. If warranted, post hoc comparisons could adjust *P* values, though at the cost of an increased probability of Type II errors. Ranked from most prone to Type II error to most prone to Type I, adjustment options include: (1) Bonferroni or other correction for all post hoc tests conducted in all analyses, as in Bradley et al. (2003); (2) a sequential correction that increases the alpha values following each significant post hoc comparison; (3) correction for the number of post hoc tests involving a given response vari-

able, regardless of the number of tests on other response variables; (4) a sequential approach to number 3; and (5) no adjustment, regardless of the number of post hoc comparisons made. The last option recognizes that effect sizes in ecological studies are often low, and numerous tests near $P = 0.05$ may be more meaningful than a single post hoc test at a very low value of alpha (Moran 2003).

Which correction to use should depend upon the context of the study. In our initial paper on visitation effects (Cahill et al. 2001), we used a Bonferroni correction to determine species-level responses upon finding a species × treatment interaction. We did so because we felt our results were so unexpected that we wanted them to not be dismissed by an argument of inflated experiment-wise error rates (Type I). In contrast, Bradley et al. (2003) argue in support of accepting the null hypothesis, and thus Type II errors should have been of primary concern.

Appropriate levels of replication and F_{crit}.—Because Bradley et al. (2003) treat field as a random effect in the ANOVA model, the F ratio for the main effect of visitation is constructed using the field × treatment (f × t) interaction as the error MS ($MS_t/MS_{f\times t}$). Two treatment levels and two fields (for most species) result in few degrees of freedom (df = 1, 1) available for testing for the main effect of treatment and a correspondingly high F_{crit} value ($F_{crit} = 161$). The Dunn-Sidák adjustment of P ($P = 0.0009$) raises F_{crit} to approximately 500 000. Such a large critical value could have been avoided had the experimental design increased replication at the level of field, rather than plants within field, thereby increasing the denominator df for the main effect of visitation. For example, using four fields instead of two would have reduced critical values for the main treatment effect (df = 1, 3; $P = 0.05$, $F_{crit} = 10.1$; $P = 0.0009$, $F_{crit} = 179$).

Alternatively, some researchers may argue that since only two fields were used for most species, fields are a fixed, rather than a random effect. Such a change would cause the denominator of the F ratio to be the residual error, increasing the denominator df (df = 26–66) and reducing F_{crit} (~4) for the test of the main effect of treatment. Whether fields should be considered fixed or random in the Bradley et al. (2003) study is debatable, and a detailed discussion of fixed and random effects is beyond the scope of this paper (see Underwood 1997).

To demonstrate the potential impacts of very high critical values on Type II error rates, we created 20 simulated data sets, structured with two levels of field and treatment, and 15 cells of each field × treatment combination (60 cells total). This number is similar to that found for most species in the Bradley et al. (2003) data. We filled the cells with a normally distributed response variable (''biomass'') with a mean of 200 and 1 SD of 100, resulting in a coefficient of variation (CV) of 50%. The CV of the biomass data in the Bradley et al. study was around 100% for most species, and thus the simulated data were less variable than the actual data. In the initial construction of the data sets, we did not build in any ''true'' difference in means among fields or treatments. We then repeated the analyses three times, reducing the biomass values of the ''visited'' plants in the 20 data sets by 10, 50, and 99%. The 10% difference is similar to the effect size found in a meta-analysis of visitation effects on biomass (Hik et al. 2003), the 50% difference represents a large effect size, and the 99% difference represents such a large effect size that treatment effects would be visually obvious even without the use of statistics. Using $P = 0.0009$, a main effect of treatment was *never* found in ANOVA for any of the 20 data sets, even when one mean was 99% lower than the other. Using $P = 0.05$, a main effect of treatment was found in 2 of the 20 data sets in the control group (close to the expected 1/20), in 1 data set in the 10% group, in 5 data sets with a 50% difference, and in 14 data sets with a 99% difference. Even with $P < 0.05$, differences of the expected magnitude (10%) are not detected, and 99% differences are detected only about 75% of the time.

The consequences of the statistical model used by Bradley et al. (2003) are clear in their Table 1, where the treatment F ratios for several species are >20 but not significant. Significant terms were found only for field and the treatment × field interaction, due to the F ratios for these factors being constructed with the residual error MS as the denominator, with correspondingly higher df and F_{crit} values around 12 ($P = 0.0009$).

In total, the choices in design and analysis made by Bradley et al. (2003) meant that detecting visitation effects of the expected magnitude was highly unlikely. Their main conclusion that spatial heterogeneity is a more important effect than visitation could be due to it being the only main effect that could be reasonably detected with their statistical model.

Results from analyses of alternative statistical models

Bradley et al. (2003) kindly provided us with their data for reanalysis. We first reassessed their ANOVA results using $P = 0.05$ rather than $P = 0.0009$ as the indication of a significant effect. We then conducted a single ANOVA analysis for each location, including species as a fixed effect.

Uncorrected P values.—Of 15 species-level analyses, there were 0, 4, and 1 species with significant treatment, field, and treatment × field terms with respect to herbivory ($P < 0.0009$), and 0, 1, and 0 species with respect to biomass (Bradley et al. 2003: Table 1).

TABLE 1. Results of three-way ANOVA for leaf herbivory and shoot biomass at the end of the three experiments conducted by Bradley et al. (2003) in three locations in the United States.

Source of variation	*F* ratio over	Minnesota (MN) df	MS	*F*	*P*	Nebraska (NE) df	MS	*F*	*P*
Arcsine (proportion of leaves with herbivory at the end of the experiment)									
Visitation, V	V × F	1	0.329	0.78	0.428	1	0.100	8.59	0.209
Species, S	S × F	3	3.342	8.66	0.105	4	10.417	15.76	**0.010**
Field, F	Residual	4	1.030	7.20	<**0.001**	1	0.579	13.79	<**0.001**
V × S	V × S × F	3	0.791	1.55	0.415	4	0.033	0.35	0.832
V × F	Residual	4	0.423	2.96	**0.020**	1	0.012	0.29	0.593
S × F	Residual	2	0.386	2.70	0.069	4	0.661	15.74	<**0.001**
V × S × F	Residual	2	0.509	3.56	**0.030**	4	0.093	2.21	0.069
Residual		317	0.143			234	0.042		
Biomass at end of the experiment									
Visitation	V × F	1	17.949	3.28	0.145	1	427	0.04	0.881
Species	S × F	3	179.780	8.89	0.103	4	70 384	10.64	**0.021**
Field	Residual	4	4.143	0.90	0.463	1	14 382	4.18	**0.042**
V × S	V × S × F	3	5.060	43.62	**0.022**	4	4287	0.37	0.820
V × F	Residual	4	5.477	1.19	0.314	1	11 870	3.45	0.064
S × F	Residual	2	20.224	4.41	**0.013**	4	6612	1.92	0.107
V × S × F	Residual	2	0.116	0.03	0.975	4	11 575	3.36	**0.011**
Residual		336	4.591			258	3443		

Notes: In all analyses, visitation and species are fixed effects, with site a random effect. In Minnesota, different species occurred in different sites, resulting in missing cells in the ANOVA. There was no qualitative difference in any of the three states for either response variable as a function of using type III or type IV SS, and type III SS are presented. The appropriate MS error terms used for calculation of *F* ratios were determined as presented in Underwood (1997:372). Data are considered significant at $P < 0.05$ (boldfaced for ease of reading).

At $P = 0.05$ these values change to 0, 7, and 4 species for herbivory, and 0, 11, and 1 species for shoot biomass. These results are not indicative of widespread and strong visitation effects, but are consistent with our prior work in which about 1/3 of the species tested show some sort of visitation effect. The lack of any main effect of treatment is likely due to the high critical value ($F_{crit} = 161$) resulting from the structural issue of only having df = 1, 1. This problem cannot be corrected simply through alternative adjustments of *P* values, and instead requires a different experimental design.

A single ANOVA for each location.—We reanalyzed the data from each location by including species as a fixed factor in the model (SPSS 2002). The ability to detect a main treatment effect is still limited by low df, though more df are available for testing interaction terms that include treatment (Table 1). There were no qualitative differences in the results depending upon whether Type III or Type IV SS were used, and Type III SS are presented here. For both herbivory and biomass, field is a significant term in all three locations, supporting Bradley et al.'s (2003) conclusion regarding the importance of spatial heterogeneity. While treatment is not a significant main effect, there were significant treatment × species × field interactions in two of the three locations for herbivory (MN, SC), and in one location for shoot biomass (NE). There was also a significant treatment × species interaction for biomass in MN (Table 1). In total, there is evidence that visitation interacts with other factors to influence herbivory or biomass in all three locations (MN: herbivory and biomass; NE: biomass; SC: herbivory).

The prevalence of treatment × field × species interactions indicates that the effects of visitation varies spatially among fields and that the spatial pattern of visitation effects differs among species. The biological causes of these interactions are not clear, and are beyond the scope of this paper. However, there is little value in making statements about the relative importance of each factor separately (e.g., field or visitation), as interaction terms trump main effects. Instead, these data indicate visitation effects are one factor that influenced the growth and herbivory of plants in the Bradley et al. (2003) study. The interactions with field and species reinforce our initial determination that these effects are ''uncertain.''

Discussion

The results of our reanalysis show how choices in the experimental design and the statistical model are critical to determining the conclusions that can be drawn from data. They highlight the importance of understanding which error MS terms will be used to calculate the *F* ratios for effects of interest before the experiment is conducted (Underwood 1997), and thus replicating at the level that increases the denominator df (i.e., fields). The reanalyzed data from Bradley et

TABLE 1. Extended.

South Carolina (SC)			
df	MS	*F*	*P*
1	0.110	0.42	0.633
4	4.642	7.15	**0.041**
1	0.318	9.94	**0.002**
4	0.040	0.39	0.808
1	0.261	8.16	**0.005**
4	0.649	20.28	**<0.001**
4	0.103	3.25	**0.012**
276	0.032		
1	428.128	7.32	0.225
4	734.569	3.48	0.127
1	2328.264	29.35	**<0.001**
4	80.684	3.55	0.124
1	58.513	0.74	0.391
4	210.861	2.66	**0.033**
4	22.749	0.29	0.886
280	79.316		

al. (2003) also support our previous conclusions that visitation effects are real but unpredictable (Cahill et al. 2001, 2002, Hik et al. 2003). The results are not consistent with those of Schnitzer et al. (2002), who worked at one of the same locations as Bradley et al. (Cedar Creek, Minnesota, USA). It is particularly striking that different researchers using similar methods and working at the same site in the same year obtained different results regarding the impact of both visitation and spatial variation on herbivory (Schnitzer et al. 2002, Bradley et al. 2003).

Should ecologists worry about subtle effects? Yes, and we already do. There have been repeated calls to increase long-term monitoring of natural systems (e.g., Ehrlich 1996), partly because many ecological factors or treatments cause subtle effects that compound over time (e.g., Johnson et al. 2000). In the search for long-term and subtle effects, equally subtle but unintended effects associated with applying the treatments, or simply monitoring plots, become particularly troubling. It seems reasonable to employ a precautionary approach in dealing with visitation effects, reducing Type II, rather than Type I errors.

How do we do this? Ideally, methods other than repeated visitation of research plots may be available (e.g., remote censusing, imaging, etc.); however, there will certainly continue to be many questions for which it is essential that a researcher be physically present (e.g., to find tagged plants). We suggest that in such a case it would be prudent to conduct a study to determine the consequences of one's impact on the system (with low Type II error rates). Such a policy is standard for other methodologies (e.g., cage effects, carrier effects for insecticides, etc.), but inexplicably lacking in terms of human disturbance. In practice, this means establishing additional reference plots/plants that are monitored infrequently (e.g., once per year) and determining whether they differ from the main study plots in the variables of interest. The cost of this approach is that given limited time and money, increased numbers of ''visitation control'' plots will come at the expense of reduced treatment replicates. Under many circumstances it may be more important to increase the accuracy of the measured effects than to increase the precision of potentially inaccurate and uninformative results.

Where such a design is not possible, researchers should minimize their visits to research plots and collect only information essential to testing the specific questions of interest. Consideration should also be given to sampling less frequently, or even avoiding the use of longitudinal sampling designs when visitation effects are particularly likely to occur. Particular concerns are raised in experiments that impose a general treatment (e.g., CO_2 enrichment), and then employ teams of researchers to measure different response variables, thereby increasing total visits to individual plants or plots. To the plants, it doesn't matter how many times an individual visits, but may matter how many times anyone visits. The less preferred alternative is to assume that visitation effects do not occur or that nothing can be done about them.

We suggest that as ecologists we cannot escape the reality that by studying an individual, a community, or a process, we are potentially altering the biology of the research subjects. A search of the ISI database (22 October 2003) found 32 papers published since 1990 with the phrase ''handling effects'' in the title, abstract, or key words. These were reports in wildlife and conservation biology, agri/aqua-culture, physiology, and medicine, where researchers were concerned that handling study animals impacted experimental results. Plant biologists have also been interested in related processes, with a search for the term ''thigmomorphogenesis'' finding 94 papers in plant physiology, biomechanics, genetics, and anatomy. Demonstrating the ecological importance of these already-identified effects of handling requires investigation using appropriate experimental and statistical methods. We are not suggesting that the effects of visitation and touch overwhelm other processes. We simply maintain that they do exist, and that our ability to predict their magnitude and influence on ecological studies is very limited.

Acknowledgments

We thank K. Bradley and her colleagues for providing us with their data, enabling our reanalysis. We thank R. Rockwell, two anonymous reviewers, and the Cahill and Casper laboratory groups for comments that improved the quality of

this manuscript. We thank M. Dale for suggesting that we analyze simulated data. J. F. Cahill and D. S. Hik were supported by NSERC discovery grants during the preparation of this manuscript.

Literature Cited

Bradley, K. L., E. I. Damschen, L. M. Young, D. Kuefler, S. Went, G. Wray, N. M. Haddad, J. M. H. Knops, and S. M. Louda. 2003. Spatial heterogeneity, not visitation bias, dominates variation in herbivory. Ecology **84**:2214–2221.

Cabin, R. J., and R. J. Mitchell. 2000. To Bonferroni or not to Bonferroni: when and how are the questions. ESA Bulletin **81**:246–248.

Cahill, J. F., Jr., J. P. Castelli, and B. B. Casper. 2001. The herbivory uncertainty principle: visiting plants can alter herbivory. Ecology **82**:307–312.

Cahill, J. F., Jr., J. P. Castelli, and B. B. Casper. 2002. Separate effects of human visitation and touch on plant growth and herbivory in an old-field community. American Journal of Botany **89**:1401–1409.

DiStefano, J. 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. Functional Ecology **17**:707–709.

Duncan, R. S., D. G. Wenny, M. D. Spritzer, and C. J. Whelan. 2002. Does human scent bias seed removal studies? Ecology **83**:2630–2636.

Ehrlich, P. 1996. Conservation in temperate forests: what do we need to know and do? Forest Ecology and Management **85**:9–19.

Heisenberg, W. 1927. Über den anschaulichen inhalt der quantentheorestischen kinematik und mechanik. Zeitschrift Für Physik **43**:172–198.

Hik, D. S., M. Brown, A. Dabros, J. Weir, and J. F. Cahill,Jr. 2003. Prevalence and predictability of handling effects in field studies: results from field experiments and a meta-analysis. American Journal of Botany **90**:270–277.

Johnson, L. C., G. R. Shaver, D. H. Cades, E. Rastetter, K. Nadelhoffer, A. Giblin, J. Laundre, and A. Stanley. 2000. Plant carbon–nutrient interactions control CO_2 exchange in Alaskan wet sedge tundra ecosystems. Ecology **81**:453–469.

Major, R. E. 1990. The effect of human observers on the intensity of nest predation. Ibis **132**:608–612.

Moran, M. D. 2003. Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos **100**:403–405.

Schnitzer, S. A., P. Reich, B. Bergner, and W. P. Carson. 2002. Herbivore and pathogen damage on grassland and woodland plants: a test of the herbivore uncertainty principle. Ecology Letters **5**:531–539.

Singer, M. C., and P. Wedlake. 1981. Capture does affect probability of recapture in a butterfly species. Ecological Entomology **6**:215–216.

Sokal, R. R., and F. J. Rohlf. 1997. Biometry. Third edition. W. H. Freeman and Company, New York, New York, USA.

SPSS. 2002. SPSS for Windows. Version 11.5. SPSS, Chicago, Illinois, USA.

Underwood, A. J. 1997. Experiments in ecology: their logical design and interpretation using analysis of variance. Cambridge University Press, Cambridge, UK.

Ecology, 85(10), 2004, pp. 2906–2910
© 2004 by the Ecological Society of America

SPATIAL HETEROGENEITY, NOT VISITATION BIAS, DOMINATES VARIATION IN HERBIVORY: REPLY

Svata M. Louda,[1,4] Anne M. Parkhurst,[2] Kate L. Bradley,[1] Elisabeth S. Bakker,[1] Johannes Knops,[1] Ellen I. Damschen,[3] and Lauren M. Young[1,5]

[1]*School of Biological Sciences, University of Nebraska, Lincoln, Nebraska 68588-0118, USA*

[2]*Department of Statistics, University of Nebraska, Lincoln, Nebraska 68588-0712, USA*

[3]*Department of Zoology, North Carolina State University, North Carolina 27695, USA*

Manuscript received 12 April 2004; accepted 13 April 2004. Corresponding Editor: B. Shipley.

[4] E-mail: SLouda@UNL.edu

[5] Present address: Department of Botany and Agricultural Biochemistry; University of Vermont, Burlington, Vermont 05405 USA.

We are so prone to attribute a phenomenon to a single cause, that when we find an agency present, we are liable to rest satisfied therewith and fail to recognize that it is but one factor, and perchance a minor factor, in the accomplishment of the total result.

—Thomas C. Chamberlain (1890) [1965:756]

Cahill et al. (2004) criticize our recent study (Bradley et al. 2003) that contradicted the cause that they have adopted to explain an interesting phenomenon—high levels of variation in the intensity of herbivory among plants in nature. Their hypothesis is that visiting plants generally and significantly alters the intensity of herbivory (Cahill et al. 2001). If this visitation hypothesis were supported, then the interpretation of existing studies of the intensity and impact of herbivory on plant performance would need to be reexamined. However, few data support their hypothesis (Schnitzer et al. 2002, Bradley et al. 2003), including the evidence in their own studies (Cahill et al. 2001, 2002) and their reanalysis of our data (Cahill et al. 2004).

We found that spatial heterogeneity of environmental context (variation among sites) contributed more significantly to the variance in observed intensity of insect herbivory than did the fact that we ''visited and touched'' plants to measure them. Other, even earlier,

studies of the potential complex artifacts of experimental procedures in analysis of herbivory did not find evidence to support the hypothesis either (e.g., McGuinness 1997, Schnitzer et al. 2002). Thus, the continuing emphasis by Cahill and colleagues on the *potential* of visitation to affect intensity of herbivory makes their challenge to our research look like an attempt to hold onto an idea (''agency'') that is a small factor, with little general relevance, in the total explanation for variation in the intensity of herbivory in the field.

The focus of their challenge is not on the methods we used, which paralleled and in fact expanded upon their methods, nor on the data we collected. Instead, Cahill et al. (2004) suggest that our statistical approach was inappropriate ''. . . when responses have *low effect sizes* such as those previously *demonstrated*'' (italics ours). We argue that the real issue is not whether visitation might under some circumstances have low effect sizes on the intensity of herbivory, but rather whether such visitation generally alters herbivory sufficiently (with large effect sizes) to alter our understanding of the interaction and its implications for plant performance and dynamics.

They base their critique on a challenge of two aspects of our statistical analyses: (1) whether our levels of replication were sufficient to avoid making a Type II error (accepting the null when it is false), given our effort to avoid making a Type I error with multiple comparisons by using a Bonferroni procedure; and (2) whether our analyses by plant species rather than a more aggregated analysis by state would misrepresent the effect of visitation on intensity of herbivory, or perhaps secondarily on plant biomass. We address each of these points. We also argue that our basic conclusion, that spatial heterogeneity is larger than visitation effects as a factor in understanding variation in intensity of herbivory, is supported by both the incorrect reanalysis done by Cahill and colleagues and by our additional analyses presented here.

First, would our design, replication, and use of an admittedly conservative sequential Bonferroni correction for multiple comparisons have prevented us from us detecting small differences in herbivory between treatments? Multiple comparisons are possible only for species means, since the visitation treatment has only two levels. Thus, the Bonferroni correction applied only to species differences. We did a power analysis for visitation to determine whether our design would have allowed us to detect 10, 20, and 30% differences between visitation treatments. Analyses by state (Minnesota [MN], Nebraska [NE], or South Carolina [SC]) showed that, given our design and replication, power to detect a 10% difference in herbivory between treatments was: 0.63 in MN, 0.99 in NE, and 0.76 in SC. To detect a 20% difference in herbivory between treatments, power was very high: 0.99 in MN, 0.99 in NE, and 0.99 in SC. For biomass, a secondary issue, our power was more variable due to large differences in the residual term for each state. These differences were caused by large variation in the size of the plants in NE and SC, which ranged from 7 to 93 g. In any case, power to detect a 20% difference in biomass between treatments was: 0.90 in MN, 0.26 in NE, and 0.56 in SC. However, power to detect a 30% difference in biomass was: >0.99 in MN, 0.50 in NE, and 0.88 in SC. The main point is that, for the evaluation of visitation effects on intensity of herbivory, we clearly had sufficient power to reject the null hypothesis of even a small difference in visitation effects.

Second, does an alternative, more aggregated analysis of the data, by geographic region, lead to a different conclusion on the minimal role of plant visitation in the measurement of either herbivory intensity or biomass of ''touched'' plants? In our initial analysis (Bradley et al. 2004) we again chose to be conservative, recognizing that species should differ in the level of herbivory, independent of treatment, and accepting the fact that sampling error could reflect subtle differences in data collection among three groups of investigators in three regions of the country. However, in response to the criticism we have reanalyzed as requested, while correcting for errors made by Cahill and colleagues in their attempt to reanalyze our data. The results reinforce our initial interpretation of there being minor, if any, visitation effects in the measurement of intensity of herbivory.

In the reanalysis by state (= geographic region), we used a mixed-model analysis of variance (SAS Institute 2001: Mixed procedure) to describe the sources of variability. As in Cahill et al. (2004), visitation and species were considered fixed effects and field was random. The mixed-model procedure is based on likelihood, rather than the more dated ANOVA orientation recommended by Cahill and colleagues (see Littell et al. 1996, Steel et al. 1997). Further, we revised the suggested analysis based on our knowledge of the data collection; specifically, three species from MN, and all five species from NE and from SC were in a completely randomized design replicated within each of two fields. Thus, a different analysis was required for the experimental design (see Table 1). In MN, data on two additional species were collected at different fields, and thus these species are analyzed separately (Table 2). We first checked for significant interactions and found none. When we used the statistical design consistent with the way the data were collected, we found no visitation $\times$ species interactions, nor any field $\times$ treatment interactions. (Here treatment is designated as the visitation–species combination). We thus are free to

TABLE 1. Results of mixed-model analysis for both leaf herbivory (arcsine transformed) and biomass responses at the end of three visitation experiments conducted in three U.S. states in 2002 (see Bradley et al. 2003).

	Herbivory‡			Biomass		
Source of variation†	df	*F* or variance	*P*	df	*F* or variance	*P*
Minnesota						
Visitation, V	1, 191	0.35	0.55	1, 201	0.95	0.33
Species, Spp.	2, 191	12.2	<0.0001	2, 201	28.1	<0.0001
V × Spp.	2, 191	0.71	0.49	2, 201	0.17	0.84
Field		0.26	0.40		0.08	0.47
Field × V × Spp.–V		1.21	0.11		0.72	0.24
Residual		9.79	<0.0001		10.1	<0.0001
Nebraska						
Visitation	1, 234	0.32	0.67	1, 258	0.05	0.83
Species	4, 234	29.2	<0.0001	4, 258	7.2	<0.0001
V × Spp.	4, 234	0.09	0.98	4, 258	0.44	0.78
Field		0.34	0.37		0.24	0.40
Field × V × Spp.–V		1.84	0.03		1.39	0.08
Residual		10.8	<0.0001		1.28	<0.0001
South Carolina						
Visitation	1, 220	0.24	0.62	1, 224	0.86	0.36
Species	3, 220	54.2	<0.0001	3, 224	5.77	0.0008
V × Spp.	3, 220	0.28	0.84	3, 224	0.32	0.81
Field		0.14	0.44		0.65	0.26
Field × V × Spp.–V		1.02	0.15		0.96	0.17
Residual		10.5	<0.0001		10.6	<0.0001

Notes: Treatment and species are fixed effects, and field is a random effect. *F* statistics are presented for the fixed effects, while variance is presented for the random effects. Numerator df are listed first, then denominator df. Three species were analyzed separately (see Table 2).

† One treatment of interest is variation in the species-specific response to the visitation treatment; we defined a variable that represents the combined effects of visitation treatment and specific species, Spp.–V.

‡ Measured as the proportion of leaves damaged.

TABLE 2. Results of mixed-model analysis for both intensity of leaf herbivory (arcsine transformed) and biomass responses for *Potentilla recta*, *Solidago gigantea* (both at Minnesota), and *Tephrosia florida* (at South Carolina).

	Herbivory†			Biomass		
Source of variation	df	*F* or variance	*P*	df	*F* or variance	*P*
Potentilla recta						
Visitation	1, 61	1.72	0.19	1, 68	7.20	0.009
Field		0.62	0.27		0.58	0.28
Residual		5.52	<0.0001		5.83	<0.0001
Solidago gigantea						
Visitation	1, 67	10.51	0.002	1, 69	0.05	0.82
Field		0.66	0.26		0.20	0.42
Residual		5.79	<0.0001		5.87	<0.0001
Tephrosia florida						
Visitation	1, 57	1.64	0.21	1, 57	4.24	0.04
Field		0.70	0.24		0.55	0.29
Residual		5.34	<0.0001		5.34	<0.0001

Note: These three species were analyzed separately: the two in Minnesota because of experimental design, and the one in South Carolina because of significant visitation effects due to deer browsing (see Bradley et al. 2003).

† Measured as the proportion of leaves damaged.

consider the main effects of visitation (treatment) and species.

In the main analysis (Table 1), we found no significant visitation effects in *any* state (MN, NE, or SC) once one species (*Tephrosia florida*) was removed (see below). For the two species in MN that were analyzed separately (Table 2), one species (*Solidago gigantea*) had a significant effect of visitation on herbivory, but not on biomass. One hypothesis for this finding reflects the fact that all ramets measured belonged to only two clones of *S. gigantea*, so individual ramets were not completely independent of each other. The other two significant responses were in biomass. A visitation effect on *Potentilla recta* biomass (Table 2) likely reflects the deer browsing (vertebrate, not insect, herbivory); visited *P. recta* plants were shorter, and so weighed less, than unvisited plants, and they had evidence of deer browsing (see Bradley et al. 2003). A visitation effect on *Tephrosia florida* biomass (Table 2) likely reflects plant stature, since this species is a procumbent species; repeated visitation may have led to trampling of some individuals. However, visitation had no significant effect on the amount of herbivory recorded on either of these species (Table 2). Together, our results and those of the Cahill et al. (2004) reanalysis are consistent. Visitation also was not found as a main effect on herbivore damage in either Cahill et al. (2001: 309: Table 2) or Cahill et al. (2002:1405: Table 3). In another paper (Hik et al. 2003), they found some evidence of visitation as a main effect on level of herbivory, but only in one of three sites for only 2 of 16 species (see below); thus, their findings also are more consistent with the alternative spatial-heterogeneity hypothesis than with their visitation hypothesis.

Despite this, Cahill and colleagues persist in arguing for some biological significance to their findings of weak, infrequent visitation effects. All of the reanalyses of our data, both here and Cahill et al. (2004), support our original interpretation. Cahill et al. (2004) also concluded that ''. . . the results are not indicative of widespread and strong visitation effects.'' Yet, they argue that their reanalysis of our data is ''consistent with our prior work in which about 1/3 of the species tested show *some sort of visitation effect*'' (Cahill et al. 2004; italics ours). We suggest that even this interpretation is overstated based on the data available. The study by the Cahill group with the largest number of species (Hik et al. 2003) included 16 species at three sites within one region, or 48 species–site combinations. In that study they found a significant visitation effect at two sites and only for 3 of 16 species (18.8% of species, 6.25% of species–site combinations). Further, only two of the three species differed in intensity of herbivory (leaf area damaged) between control and visited groups; the third species had lower biomass in the visited group (Hik et al. 2003). Thus, for evaluation of the hypothesis that visitation alters herbivory, only 2 of 48 species–site combinations (4.2%) had any evidence of a visitation effect on measurement of herbivory. Adding the main effect on biomass on one species at one site (2.1% of the species–site combinations possible) seems to be an attempt to stretch the data to fit a preconceived notion that visitation must be having some sort of effect. Note, also, that Hik et al. (2003) did not evaluate the consequences of the herbivory, or of the statistically elevated estimate of herbivory in response to visitation for the two species at one site, on the biological issue—plant performance. In summary, >80% of the species studied and >95% of the species–site combinations evaluated even in that study showed *no herbivory response to visitation* (Hik et al. 2003). Clearly, even in their largest data set, the intensity of herbivory was rarely affected by visitation, if at all, for the large majority of plant species. We conclude that visitation effects are real: *real small*.

Finally, we have no argument with the suggestion that measurement *might* alter the subject being measured (Heisenberg 1927). The real question in this dispute, however, is whether this artifact is important in the quantification of the intensity of herbivory, where ''herbivory'' is defined as leaf consumption by herbivores. The evidence does not support the inference of significant effect of measurement. To date, five published studies have been designed specifically to examine the effects of visitation on herbivory. These studies included 46 different species in a variety of habitats, including grasslands, old fields, tundra, and boreal forests. Across all studies, only five different species (10.9%) showed any effect of visitation on an estimate of herbivory, while for the other 41 species visitation had no detectable effect on the intensity of herbivory. We conclude that the preponderance of data show that visitation has little effect on the key phenomenon, the intensity of herbivory and its influence on plant performance, regardless of the specific statistical analysis used.

Despite concurring that visitation in general has low effect sizes, Cahill and colleagues continue to imply that every subsequent study of herbivory (and, in fact, every field study of any aspect of plant performance!) requires a visitation-measurement control (Cahill et al. 2001, 2004). We strongly disagree. All field data collections are constrained by logistical considerations (time, money). Since the data across all studies show that the visitation effects are uncommon and small, the constraint of logistics argues against adding ''visitation control'' plots. A better investment of the effort would be further replication to quantify the variation that has been found to be significant, the variation among species, sites, and habitats (e.g., Louda 1982, 1983, Louda

et al. 1987*a*, *b*, *c*, Hunter et al. 1992, Louda and Rodman 1996, Strauss and Zangerl 2002, Bradley et al. 2003, Myers and Bazely 2003).

Clear, immediate rejection of this unsupported hypothesis is important, since its fate will affect the allocation of effort in future studies of herbivory. In 1890, Chamberlain (1964:755) clearly and eloquently cautioned against the pursuit of a pet hypothesis:

> *The moment one has offered an original explanation for a phenomenon which seems satisfactory, that moment affection for his intellectual child springs into existence, and as the explanation grows into a definite theory his parental affections cluster about his offspring and it grows more and more dear to him. . . . There springs up also unwittingly a pressing of the theory to make it fit the facts and a pressing of the facts to make them fit the theory. . . .*

In the interests of progress in the study of herbivory, we suggest that further defense of an ''herbivory uncertainty principle,'' or more precisely the visitation hypothesis that has not been well substantiated by field data, should be abandoned. The evidence does not support the putative ''principle.'' Instead, more effort should be placed on increasing our mechanistic, predictive understanding of the influence of environmental context on the intensity of herbivory, the biological mechanisms driving the spatial and temporal variation in herbivory, and the consequences of herbivory on plant abundance, distribution, and dynamics.

Acknowledgments

We thank all our colleagues who contributed to the original work and also to those colleagues who have taken the time to discuss the issues, help handle the data, and comment on drafts of our response, especially Mick Crawley, Nick Haddad, Heather Hager, Tony Joern, Tom Miller, Amy Kochsiek, Ramesh Laungani, and Jayne Jonas.

Literature Cited

Bradley, K. L., E. I. Damschen, L. M. Young, D. Kuefler, S. Went, G. Wray, N. M. Haddad, J. M. H. Knops, and S. M. Louda. 2003. Spatial heterogeneity, not visitation bias, dominates variation in herbivory. Ecology **84**:2214–2221.

Cahill, J. F., Jr., B. B. Casper, and D. S. Hik. 2004. Spatial heterogeneity, not visitation bias, dominates variation in herbivory: comment. Ecology **85**:2701–2706.

Cahill, J. F., Jr., J. P. Castelli, and B. B. Casper. 2001. The herbivory uncertainty principle: visiting plants can alter herbivory. Ecology **82**:307–312.

Cahill, J. F., Jr., J. P. Castelli, and B. B. Casper. 2002. Separate effects of human visitation and touch on plant growth and herbivory in an old-field community. American Journal of Botany **89**:1401–1409.

Chamberlain, T. C. 1965. The method of multiple working hypotheses. [Reprint of 1890 Science article.] Science **148**: 754–759.

Heisenberg, W. 1927. Über den anschaulichen inhalt der quantentheorestischen kinematik und mechanik. Zeitschrift für Physik **43**:172–198.

Hik, D. S., M. Brown, A. Dabros, J. Weir, and J. F. Cahill, Jr. 2003. Prevalence and predictability of handling effects in field studies: results from field experiments and a meta-analysis. American Journal of Botany **90**:270–277.

Hunter, M., T. Ohgushi, and P. W. Price. 1992. Effects of resource distribution on animal–plant interactions. Academic Press, San Diego, California, USA.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS system for mixed models. SAS Institute, Cary, North Carolina, USA.

Louda, S. M. 1982. Distribution ecology: variation in plant recruitment in relation to insect seed predation. Ecological Monographs **52**:25–41.

Louda, S. M. 1983. Seed predation and seedling mortality in the recruitment of a shrub, *Haplopappus venetus* (Asteraceae), along a climatic gradient. Ecology **64**:511–521.

Louda, S. M., P. M. Dixon, and N. J. Huntly. 1987*a*. Herbivory in sun and shade at a natural meadow–woodland ecotone in the Rocky Mountains. Vegetatio **72**:141–149.

Louda, S. M., M. A. Farris, and M. J. Blua. 1987*b*. Variation in methylglucosinolate and insect damage to *Cleome serrulata* (Capparaceae) along a natural soil moisture gradient. Journal of Chemical Ecology **13**:569–581.

Louda, S. M., N. J. Huntly, and P. M. Dixon. 1987*c*. Insect herbivory across a sun–shade gradient: response to experimentally-induced in situ plant stress. Acta Oecologia, Oecologia Generalis **8**:357–363.

Louda, S. M., and J. E. Rodman. 1996. Insect herbivory as a major factor in the shade distribution of a native crucifer (*Cardamine cordifolia* A. Gray, bittercress). Journal of Ecology **84**:229–238.

McGuinness, K. A. 1997. Tests for artefacts in some methods used to study herbivory and predation in mangrove forests. Marine Ecology Progress Series **153**:37–44.

Myers, J. H., and D. Bazely. 2003. Ecology and control of introduced plants. Cambridge University Press, Cambridge, UK.

SAS Institute. 2001. SAS user's guide. Version 8.2. Cary, North Carolina, USA.

Schnitzer, S. A., P. Reich, B. Bergner, and W. P. Carson. 2002. Herbivore and pathogen damage on grassland and woodland plants: a test of the herbivore uncertainty principle. Ecology Letters **5**:531–539.

Sokal, R. R., and F. J. Rohlf. 1997. Biometry. Third edition. W. H. Freeman and Company, New York, New York, USA.

Steel, R. G. D., J. H. Torrie, and D. A. Dickey. 1997. Principles and procedures of statistics: a biometrical approach. McGraw-Hill, New York, New York, USA.

Strauss, S. Y., and A. Zangerl. 2002. Plant–insect interactions in terrestrial ecosystems. Pages 77–106 *in* C. M. Herrera and O. Pellmyr, editors. Plant–animal interactions: an evolutionary approach. Blackwell Scientific, Oxford, UK.