



THE UNIVERSITY OF CHICAGO PRESS JOURNALS

---

A Meta-Analysis of Competition in Field Experiments

Author(s): Jessica Gurevitch, Laura L. Morrow, Alison Wallace and Joseph S. Walsh

Source: *The American Naturalist*, Vol. 140, No. 4 (Oct., 1992), pp. 539-572

Published by: The University of Chicago Press for The American Society of Naturalists

Stable URL: <https://www.jstor.org/stable/2462913>

Accessed: 19-11-2019 01:25 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*The American Society of Naturalists, The University of Chicago Press* are collaborating with JSTOR to digitize, preserve and extend access to *The American Naturalist*

## A META-ANALYSIS OF COMPETITION IN FIELD EXPERIMENTS

JESSICA GUREVITCH, LAURA L. MORROW, ALISON WALLACE,  
AND JOSEPH S. WALSH\*Department of Ecology and Evolution, State University of New York,  
Stony Brook, New York 11794-5245*Submitted February 11, 1991; Revised October 4, 1991; Accepted October 25, 1991*

**Abstract.**—A meta-analysis was conducted on field-competition experiments published in six journals over a 10-yr period. We analyzed the effects of competition on the biomass of organisms belonging to 93 species in a wide variety of habitats. Competition had a large effect overall, with a great deal of heterogeneity in that effect among organisms. There were large differences among trophic levels in competitive effects, but the relative magnitude of competition at different trophic levels was contrary to the predictions of ecological theory. Primary producers and carnivores displayed small to medium effects. In these two groups, interspecific effects did not differ from intraspecific effects, nor did effects differ in terrestrial versus aquatic habitats. The effects of competition on herbivores ranged from large effects on anurans and lotic arthropods, to medium effects on marine mollusks and echinoderms, to effects that were not statistically distinguishable from zero for terrestrial arthropods. Interspecific competitive effects among these herbivore groups were generally less than intraspecific effects. Among primary producers, the effects of competition were not different in high- and low-productivity habitats. Across all taxa, large organisms did not experience greater competitive effects than small organisms, and competitive effects did not depend on the size attained in the absence of competitors. The effects of competition were weakly density-dependent in cases in which it was possible to examine the effects of the density of neighbors. Experiments conducted on caged organisms resulted in greater competitive effects than those with free-roaming and unenclosed organisms. Experiments with small sample sizes, short durations, and poor experimental design were more variable than experiments that were larger, longer, and better planned.

Competition holds a central place in ecological and evolutionary theory. Early attempts to demonstrate the effects of competition were based mainly on laboratory or greenhouse experiments and on quantification of field observations. However, evidence for the ecological significance of competition in natural systems has been equivocal. In the last 10 yr, hundreds of field-competition experiments have been conducted on a wide diversity of organisms in a wide diversity of systems. Because of the limitations of other approaches, manipulative experiments carried out in the field may provide the best opportunity to resolve questions about competition between organisms in nature. It is well known that field experiments must be designed and interpreted with care (see, e.g., Connell 1983, 1990; Bender et al. 1984; Hairston 1989; and Goldberg and Barton 1992 for caveats), or their potential contribution is substantially diminished.

\* Present address: Ecology and Evolution, University of Chicago, 940 E. 57th Street, Chicago, Illinois 60637.

Nine years ago two highly influential papers attempted to synthesize the results of earlier field-competition experiments (Connell 1983; Schoener 1983). Yet many questions remain unanswered regarding the frequency and magnitude of competitive effects, differences in those effects among taxa, and the consequences of various artifacts for the outcome of experiments. We believe that one of the major reasons for the uncertainty surrounding the effects of competition in nature is the lack of adequate means for synthesizing the wealth of experimental results that have been gathered. This article explores the use of a new statistical tool, *meta-analysis*, to analyze field-competition experiments published over the past 10 yr (1980–1989). By allowing the synthesis of data from primary experiments, meta-analysis promises to advance our understanding of competition in nature.

#### BACKGROUND: OVERVIEW OF META-ANALYSIS AND ITS USE IN ECOLOGY

Meta-analysis permits the statistical synthesis of research results from a set of primary studies. The modern techniques have roots in the combined probability tests first proposed by Tippet (1931), Fisher (1932), and Pearson (1933), but the new methods are both more powerful and more useful than these “omnibus” tests (Hedges and Olkin 1985). Beginning with the work of Rosenthal (1963, 1976) and Glass (1976, 1978), current approaches have been developed primarily in medicine (Mann 1990) and in the social sciences, particularly in educational psychology (Glass et al. 1981; Hunter et al. 1982; Light and Pillemer 1984; Rosenthal 1984; Hedges and Olkin 1985; Wolf 1986). While it is controversial, meta-analysis has had a profound influence in resolving longstanding debates in these fields (see, e.g., Mann 1990). Meta-analysis has not previously been applied to ecological data.

In a meta-analysis, the original data are *not* reanalyzed. Instead, the outcomes of the different studies are examined to test whether they are consistent with one another and whether together they demonstrate an effect that is large, moderate, small, or not significantly different from zero. As the name implies, meta-analysis is the analysis of analyses. A number of different meta-analytic methods have been devised. They most commonly depend on obtaining an estimate of the *effect size*—essentially the magnitude of the effect of interest—from every experiment. The effect may be measured with different units in each study, but if the measurements can be equated linearly, it may be possible to standardize them to a single scale. One useful measure of the effect size is the difference between the means of two groups (e.g., an experimental group and a control group) standardized by dividing by the pooled standard deviation (Cohen 1969). (Other measures of effect size, such as correlation coefficients, are sometimes used.) This standardized mean difference,  $d$ , is conventionally considered to be large if it is greater than or equal to about 0.8 (i.e., the experimental group mean is eight-tenths of a standard deviation greater than that of the control group), medium at  $d$  equals 0.5, and small when  $d$  equals 0.2 (Cohen 1969). If the confidence interval (CI) for  $d$  overlaps zero, there is no statistically significant difference between the two groups. Cohen (1969) points out that a medium effect size of 0.5 is perceptible on casual observation: equivalent, for example, to the average difference in height

between 14-yr-old girls and 18-yr-olds. A small effect size of 0.2, for example, the mean height difference between 15- and 16-yr-old girls, would not be immediately apparent.

Meta-analysis has the potential to fundamentally change the way we draw conclusions from the outcomes of experiments. Our understanding of how nature is organized is usually founded on individual experiments that have become "text-book examples." We may also depend on more objective but statistically flawed "vote counts" (see below) in reviewing the literature on a particular topic. The field of meta-analysis encourages a very different perspective. If we imagine that any one experiment is an attempt to measure a particular, real effect, it is clear that there will be some error associated with estimating that effect. Together a set of studies can be analyzed to separate real effects from the random error, or "noise," surrounding the true effects. In fields like ecology, in which experiments cannot be exactly replicated, meta-analysis may illuminate the results of a body of data by quantitatively synthesizing separate studies.

Meta-analysis offers two major advantages for research syntheses in ecology and evolution. First, it makes a powerful set of tools available for answering the questions that are of greatest interest in reviewing and synthesizing primary literature. In reviews of ecological studies, what one most often wishes to know is the overall magnitude of an effect (Does competition have a large effect in nature?), whether that effect differs among contrasting categories of studies (Do carnivores compete more than herbivores?), and how much variation there is in the effect within and among categories. Meta-analysis is the only way to obtain this information. Second, meta-analysis is important because it offers an alternative to the ordinary approaches used in ecological research reviews. Some of the conventional approaches are logically and statistically flawed and are likely to result in misleading or incorrect conclusions.

Traditionally, reviews convey the expert opinion of the reviewer, which is necessarily subjective. While this approach cannot be faulted statistically, the power to reach unbiased and general conclusions is limited. In recent years many ecological reviews have relied on what meta-analysts call "vote counting." For example, Connell (1983), Schoener (1983), and Sih et al. (1985) used vote counting to evaluate the importance of competition and of predation in field experiments. In a vote count, studies are sorted into categories according to whether the outcome was positive and statistically significant, not statistically significant, or negative and statistically significant. Each study "casts a vote" for the importance of the overall effect, and the proportion of studies "voting" for and against the effect are compared. The decision that an effect is "real" may be based on "winning" a majority of votes. Vote tallies may be compared among categories of studies (herbivores vs. carnivores, blue- vs. white-collar males, etc.) to determine the circumstances under which the effect is more prevalent. Vote-count reviews are both intuitively appealing and easy to carry out (Hedges and Olkin 1980).

Unfortunately, the outcome of comparisons based on vote counting is likely to be the result of other factors in addition to the actual frequency and magnitude of the effect of interest. Significance levels alone are not an adequate basis for

deciding whether two studies have the same outcome. The treatment effects of two experiments may both be positive and both be statistically significant, but the two studies may be significantly different from one another. Alternatively, one study may demonstrate significant treatment effects while another does not, but the two may not be significantly different from each other (Hedges and Olkin 1985). The expected number of significant results is a function of both sample sizes and effect sizes. If the number of statistically significant outcomes is greater in one category of studies than another, that could be due to real differences in the effect of interest or it could be due to differences in sample sizes in the studies in the two categories. Vote counting will not always reliably detect real treatment effects and may fail to accurately detect agreement between studies in treatment outcome (Hunter et al. 1982). Meta-analysis does not suffer from these problems because effect size (unlike significance level) is not dependent on sample size.

It is especially problematic that vote counting leads to a strong bias toward finding no effect, making it overly conservative. Small sample sizes and small to moderate effects, both the norm in ecological studies, result in low statistical power. With low power, the error associated with estimating the effect is high, and the probability that any given study will fail to detect a significant result is high even when the true effect is nonzero. Therefore, regardless of the prevalence of the effect, the proportion of studies yielding significant results is expected to be low. Additional studies will not improve the chances that the vote count will detect an effect if they also have small samples (L. V. Hedges, personal communication). In fact, the tendency to make the wrong decision mounts as the number of studies becomes greater (Hedges and Olkin 1980). More studies like the ones available will not increase the proportion of significant results because, if new studies have the same power as old ones, they will have the same chance of producing a significant result. It is not surprising, therefore, when reviewers using vote counting on somewhat different sets of studies fail to reach similar conclusions, although other factors may also contribute to disagreements (e.g., Schoener 1985; Ferson et al. 1986; Abrami et al. 1988).

### *Ecological Questions*

The meta-analysis we performed addressed a number of questions regarding the effects of competition in field experiments, some of which have been asked in previous reviews and some of which were especially suited to a meta-analysis. It was not possible to answer all questions using the entire data set; instead, particular groups of organisms provided the only or the best answers to many of them, as explained below. These questions were as follows:

1. What is the overall magnitude of the effect of competition in field experiments, and how much does that effect vary? If competition is not generally experienced by organisms in nature or if the effects of competition are weak or fleeting, then it should be difficult to detect an overall effect that is significantly different from zero if competitive effects are measured for a large group of organisms in a great many systems. Field experiments offer the best opportunity to assess the intensity of competition in nature and the variation in the effects of competition.
2. Do competitive effects differ among trophic levels? Various predictions have

been made regarding the intensity of competition at different trophic levels (e.g., Hairston et al. 1960; Slobodkin et al. 1967; Menge and Sutherland 1976; Connell 1983; Schoener 1983; Oksanen 1988; Hairston 1989). In brief, Hairston and colleagues predicted that in terrestrial systems (phytophagous) herbivores do not compete (i.e., the effect size is not different from zero) while competition among carnivores and among producers is strong (the effect size is large). According to the authors, these predictions applied to the magnitude of the effects of competition on dominant organisms, not to a vote count of organisms. This makes meta-analysis more appropriate for testing these predictions than were previous approaches. Menge and Sutherland (1976) predicted that plants compete least, predators compete most, and herbivores compete at an intermediate level. We categorized organisms to trophic level as well as was possible, using information provided in the publications where available. In the case of mixed feeders the trophic level was based on the predominant food source if known. Some categorizations were unavoidably arbitrary.

3. Did the duration of experiments affect their outcomes?

4. Did the number of replicates affect experimental outcome?

5. Are the effects of competition dependent on the potential for growth without competitors? It has been suggested that some of the disagreement in predictions regarding the intensity of competitive effects at different resource levels may be caused by different definitions or different ways of measuring competitive effects (Grace 1990; D. Tilman, personal communication). In particular, it has been suggested that the intensity of competition should be expressed relative to performance without neighbors (Wilson and Tilman 1991). D. Tilman (personal communication) has suggested that the measurement of competitive effects should be scaled by the potential for growth in the absence of neighbors because absolute competitive effects depend on the capacity for growth without competitors. For example, Wilson and Tilman (1991) define competition intensity as the difference between growth rate with no neighbors present and growth rate in the presence of all neighbors, divided by growth rate with no neighbors.

6. Do large organisms experience greater competitive effects than small organisms? Schoener (1974, 1983) and Connell (1975, 1983) suggest that small animals generally should compete less than large ones either because they are more vulnerable to predators or because they may be more subject to harsh conditions or to disturbance.

7. Did the effects of competition in caged, stocked experiments differ from those in which the natural densities of uncaged organisms were manipulated? Numerous artifacts may change the outcome of an experiment. Experiments in which organisms are caged are at one end of the gradient of field experiments in natural systems. Rather than omit them from this study (and thereby severely compromise the representation of certain categories of organisms and systems), we evaluated the effects of caging on the level of competition by comparing caged and uncaged organisms.

8. Within trophic levels, does competition differ among terrestrial, marine, and freshwater systems?

9. Are the effects of intraspecific competition greater than those of interspecific

competition at all trophic levels? Because much of ecological theory is couched in terms of the relative intensities of intra- and interspecific competition, it is of interest to compare the relative magnitude of their effects. While it is often assumed that individuals of the same species compete most intensely, this generalization has been questioned with regard to terrestrial plants.

10. Is competition among terrestrial plants more intense in high-productivity habitats in comparison with low-productivity habitats? This has been a much-debated controversy, with some ecologists predicting or reporting greater competitive effects in productive habitats (e.g., Newman 1973; Grime 1977, 1987; Gurevitch 1986; Gurevitch and Unnasch 1989) but with others believing that competitive effects are no greater in productive areas (Tilman 1982, 1988; Grace 1990; Grace and Tilman 1990; Wilson and Tilman 1991).

11. Were the results in studies that were subject to pseudoreplication or other moderately severe design problems different in magnitude or variability (heterogeneity) than studies based on better experimental designs? Poor experimental designs may lead to unreliable or misleading results. We looked at how these problems affected the magnitude and variability of estimates of the effects of competition.

The interesting and substantive problems associated with bias on the part of people conducting and reporting on competition experiments will be examined elsewhere (J. Gurevitch and L. L. Morrow, unpublished manuscript). Meta-analysts have developed a number of creative approaches to detect bias and estimate its magnitude quantitatively (see, e.g., Light and Pillemer 1984; Rosenthal 1984; Hedges and Olkin 1985). These issues are as important in ecology as they are in other fields because there may be reason to suspect that people are more likely to carry out experiments in which they expect to find competition and that papers reporting on the discovery of statistically significant effects are more likely to be published than those finding no statistically significant effects. Since it has become fashionable to debunk the importance of competition in recent years, this bias may no longer be as great as it may have once been. The results reported here attempt to summarize the available evidence rather than to reveal a complete or unbiased picture of competition in nature.

## METHODS

### *Literature Search*

We chose six major journals that we judged would publish most recent field experiments on competition. It was not possible to conduct a more thorough search to find every field-competition experiment that has been performed. While it is impossible to know whether an expanded search would have substantially altered the nature of the data or the conclusions, we sought to make our method for selecting studies objective and unbiased. We examined issues of *American Naturalist*, *Copeia*, *Ecological Monographs*, *Ecology*, *Journal of Experimental Marine Biology and Ecology*, and *Oecologia* published between 1980 and 1989 for appropriate articles. Recent years of *Marine Biology*, *Proceedings in Marine*

*Ecology*, *Auk*, and *Journal of Mammology* were scanned, but we found no articles that could be used.

Criteria were devised a priori to avoid personal bias in selecting articles. Articles were included if all criteria applied to at least one comparison in any one experiment. By "comparison" we mean the comparison of a single experimental group with its control, for example, a group of organisms grown with and without competitors. Our criteria for including studies in the meta-analysis are described in Appendix A.

### *Extracting Data from Published Results*

To conduct the meta-analysis, we needed only the means for the experimental and control groups, the standard deviations, and sample sizes. This basic information was, however, often missing or at least well disguised. If these numbers could not be inferred or divined from other information in the article, the study could not be used. Many articles presented means and some measure of variance in the form of graphs. To use these data, graphs were digitized with a Science Accessories Corporation Graf/Pen and Graf/Bar acoustic digitizer with DS-DIGIT software (Slice 1990). The numbers obtained were highly repeatable (error was approximately 0.3%) and agreed well with unpublished data from which graphs were constructed (where available). This software is capable of capturing data scaled in inches, natural logarithms, common logarithms, or user-defined linear scales. Some articles were useless because unconventional graphing techniques, poorly defined or incomprehensible axes, and so forth made it impossible to obtain data reliably.

Obtaining an estimate of the standard deviation for experimental and control groups was especially complicated, requiring conversions based on various assumptions. An abbreviated description of how data were treated, based on the unit being measured, the equality of sample sizes, and the data available from each set of results can be found in Appendix B. A more complete description is available (Morrow 1990).

### *Calculation of Effect Sizes*

Data were analyzed on Microsoft Excel spreadsheets. For a given comparison the unbiased effect size,  $d$ , for a treatment is the standardized difference between the means of the experimental and control groups (Hedges and Olkin 1985):

$$d = \frac{\bar{Y}_e - \bar{Y}_c}{s} J(m), \quad (1)$$

where  $\bar{Y}_c$  is the mean of the control group,  $\bar{Y}_e$  is the mean of the experimental group,  $N_c$  is the total number of individuals,  $\sum n_{ij}$ , in the control group,  $N_e$  is the total number of individuals,  $\sum n_{ij}$ , in the experimental group,  $n_{ij}$  is the number of organisms in the  $j$ th plot of the  $i$ th treatment or control,  $m = N_e + N_c - 2$ ,  $s$  is the pooled standard deviation of the control and treatment groups, such that

$$s = \sqrt{\frac{(N_e - 1)(s_e)^2 + (N_c - 1)(s_c)^2}{N_e + N_c - 2}}, \quad (2)$$



$s_c$  is the standard deviation of the individuals in the control group, and  $s_e$  is the standard deviation of the individuals in the experimental group;  $J(m)$  is a correction term that removes small-sample-size bias of the standardized difference of means, where  $J$  is evaluated at  $m$  as

$$J(m) = 1 - \frac{3}{4m - 1} \quad (3)$$

(Hedges and Olkin 1985). The value of  $J(m)$  approaches one as sample size increases.

We defined the control group as that in which the density of organisms was closest to natural density, as specified by the author (see App. A). The sign of  $d$  depends on the difference between the mean value for the experimental group and the value for the control. If increased densities were indeed detrimental to an organism's fitness, reductions in the density of competitors should result in a positive effect size, and enhancements should result in a negative effect size. To pool effect sizes, the sign of  $d$  for all experimental enhancements in density was changed and the sign for reduction (removal) experiments let stand. Thus, a positive effect size indicated a detrimental effect of competition while a negative effect size indicated a beneficial effect of increased density.

The asymptotic distribution of  $d$  is normal with mean  $\delta$  and variance  $\sigma(\delta)$ . The variance is approximated by  $\hat{s}^2(d)$ , where

$$\hat{s}^2(d) = \frac{N_e + N_c}{N_e N_c} + \frac{d^2}{2(N_e + N_c)} \quad (4)$$

(Hedges and Olkin 1985).

The cumulated effect size across studies,  $d_+$ , is a weighted estimator of  $\delta$ , the true effect size, and is given by the following formula:

$$d_+ = \left[ \sum_{i=1}^k \frac{d_i}{\hat{s}^2(d_i)} \right] / \left[ \sum_{i=1}^k \frac{1}{\hat{s}^2(d_i)} \right], \quad (5)$$

where  $k$  equals the number of comparisons (Hedges and Olkin 1985). In calculating  $d_+$ , larger studies are weighted more heavily than smaller studies. This is reasonable because it is assumed that larger sample sizes will yield more precise results (Hedges and Olkin 1985).

The variance in  $d_+$ ,  $\hat{s}^2(d_+)$ , is

$$\hat{s}^2(d_+) = 1 / \left[ \sum_{i=1}^k \frac{1}{\hat{s}^2(d_i)} \right]. \quad (6)$$

The lower and upper limits for the 95% CI for  $d_+$ ,  $d_l$  and  $d_u$ , respectively, are

$$d_l = d_+ - [C_{\alpha/2} \hat{s}(d_+)] \quad (7a)$$

and

$$d_u = d_+ + [C_{\alpha/2} \hat{s}(d_+)], \quad (7b)$$

where  $C$  is the two-tailed critical value of the standard normal distribution.

### *A Schema for Meta-analysis*

The estimates of  $d_+$  and  $\hat{s}^2(d_+)$  are based on a fixed effects model that assumes that comparisons differ only in response scale and sample size. That is, it is assumed that all comparisons being averaged share the same true effect size and differ only as a result of sampling error (Hedges and Olkin 1985). As a first approach, we used this fixed effect model because a mixed model, which would have been preferable, has not yet been developed for meta-analysis. In a number of cases, where the data suggested that a fixed effects model was not really appropriate, we also estimated the effect size for particular groups using a random effects model (see below) so that the results of the two approaches could be compared.

Hedges and Olkin (1985) suggest a general approach for fitting models to data from a set of experiments in cases in which there are categorical independent variables, as an analogue to ANOVA. (Ordinary ANOVA cannot be justified when the variables being analyzed are effect-size estimates; see Hedges and Olkin 1985, p. 148.) Assume that there are a priori hypotheses that the reviewer wishes to test regarding differences among categories of studies. The first step is to test whether all studies, irrespective of category, share a common effect size. If the hypothesis that all effect sizes are equal (except for chance variation due to sampling error) is rejected, then the studies are broken into groups. The between-class homogeneity tests the extent to which effect sizes differ across classes. If the null hypothesis that all classes share a common effect size is rejected, the process continues by testing the within-class fit (the fit to a single effect size) within each class. If the within-class fit to a single effect size is rejected, categories can be further subdivided. The process ends when it is not possible to reject the hypothesis that all studies within a category share a common effect size. Alternatively, a random effects estimate of the effect size for the studies within a category could be calculated.

A statistical test for the homogeneity of effect sizes tests the null hypothesis that all effect sizes are equal versus the alternative hypothesis that at least one of the true effect sizes in a series of comparisons differs from the rest (Hedges and Olkin 1985; recall that we use "comparison" to mean a single effect size). A large sample test for homogeneity of effect sizes is based on the statistic  $Q$ , which has an asymptotic  $\chi^2$  distribution with  $k - 1$  degrees of freedom. The greater the value of  $Q$ , the greater the heterogeneity in effect sizes among comparisons.

The total heterogeneity for a group of comparisons,  $Q_T$ , can be partitioned into within-class heterogeneity,  $Q_W$ , and between-class heterogeneity,  $Q_B$ , which is analogous to the practice of partitioning variation in an ANOVA:

$$Q_B + Q_W = Q_T. \quad (8)$$

The total heterogeneity for a group of comparisons is calculated as

$$Q_T = \sum_{i=1}^p \sum_{j=1}^{m_i} \frac{(d_{ij} - d_{++})^2}{\hat{s}^2(d_{ij})}, \quad (9)$$

with  $(pm - 1)$  degrees of freedom and comparisons grouped into  $p$  classes, each class including  $m$  comparisons. For a single level of classification, the symbol  $d_{++}$  is the grand average weighted effect size across all classes (for two levels,  $d_{+++}$ , and so on). The overall within-class heterogeneity,  $Q_w$ , where

$$Q_w = \sum_{i=1}^p \sum_{j=1}^{m_i} \frac{(d_{ij} - d_{i+})^2}{\hat{s}^2(d_{ij})}, \quad (10)$$

is the sum of the heterogeneities from all classes. Heterogeneity within each class,  $Q_{wi}$ , is

$$Q_{wi} = \sum_{j=1}^{m_i} \frac{(d_{ij} - d_{i+})^2}{\hat{s}^2(d_{ij})}, \quad (11)$$

with  $(m - 1)$  degrees of freedom. The between-class heterogeneity,  $Q_B$ , is a measure of the variation between-class mean effect sizes, in which

$$Q_B = \sum_{i=1}^p \sum_{j=1}^{m_i} \frac{(d_{i+} - d_{++})^2}{\hat{s}^2(d_{ij})}, \quad (12)$$

with  $(p - 1)$  degrees of freedom (Hedges and Olkin 1985).

### *Independence of Comparisons*

Meta-analysis, like any attempt at synthesizing or summarizing results, makes the assumption that the elements being synthesized are independent. Violation of this assumption may alter the structure of the data and inflate significance levels for statistical tests (Abrami et al. 1988). Nonindependence among effect sizes included in a meta-analysis may exist at a number of conceptual levels, making it essentially impossible to avoid completely. At the simplest level, repeated measurements taken on the same individuals over the course of time, or with different measures of outcome, are not independent (e.g., Gurevitch and Chester 1986). When this problem exists, composite measurements may be a useful alternative under some circumstances (Hedges and Olkin 1985, chap. 10; Abrami et al. 1988). We did not include this type of repeated measures on the same individuals in our meta-analysis. At the next level, different groups from a factorial design experiment, or even different experiments reported in the same publication, are not necessarily independent of one another. The loss of information caused by omission of such effects may lead to much more serious distortions of the results than those caused by their nonindependence (Connell 1983; Hedges and Olkin 1985; Abrami et al. 1988). We included such effects in our meta-analysis, judging this to be the prudent middle course. For this reason, it may be wise to choose a more conservative significance level than  $P < .05$ . Somewhat more abstractly, there is also nonindependence among publications by a single author, whether each publication reports the results of an independent experiment or the results from a single experiment are split up and published in separate

articles. There is nonindependence among students and their major professors, among schools of thought, among phylogenies of organisms (Felsenstein 1985; Donoghue 1989), among the environments in which experiments are conducted, and among many other categories that one can imagine. Meta-analysis has spurred the exploration of the consequences of the nonindependence of effects, but the matter is far from resolved.

## RESULTS AND DISCUSSION

### *Overview of Field-Competition Experiments, 1980–1989*

Organisms respond in many ways to competition. We felt that combining all measures of competition in evaluating its overall effects would be unnecessarily confusing and would increase the variability (heterogeneity) among studies. The most common measure of response across all competition experiments was the biomass of organisms exposed to different levels of competition. Consequently we calculated the effect size in terms of biomass responses to competition. Other measures of response, such as responses in density and mortality, are examined elsewhere (Morrow 1990; J. Gurevitch, unpublished manuscript). Clearly the results will depend on the measure of outcome, and it must be kept in mind that we are discussing biomass responses here. We found 137 articles on competition in the field with 697 comparisons that could be used (Morrow 1990). Of these, 46 articles with 217 comparisons had biomass as a measure of response and are analyzed here (and listed in App. C). All subsequent results reported refer to these “biomass” studies.

Connell (1983) noted striking differences in the frequency of competition between articles reporting only a single comparison and those that reported results for more than one comparison. Seven of the 46 articles reported only a single comparison that we used, although most of those studies reported other results that were not relevant to our meta-analysis. Clearly the bias introduced when only a single datum is reported in an article is no longer a serious concern in competition studies.

Authors worked with a wide diversity of organisms, with 93 species' responses reported. Studies were generally small and brief. The duration of experiments ranged from 1 mo to 36 mo, with a median (among the 217 comparisons) of 4 mo. Of those 217, 163 lasted for less than 1 yr. The number of replicates ranged from one (that is, no replication in one of the treatments) to 1,455, with a median of 10 replicates. One hundred three comparisons had fewer than 10 replicates of the control, and 93 comparisons had fewer than 10 replicates for experimentals. A substantial number of studies had unbalanced designs (i.e.,  $N_c \neq N_e$ ), with some having greater than a 50:1 ratio for the number of organisms in control groups as compared with experimentals or vice versa. Few authors reported responses on a per-neighbor or per-unit biomass of neighbor basis (Goldberg and Werner 1983; Goldberg 1990). Since these data were not generally available, we could not directly examine per-unit responses. We did examine several issues indirectly related to per-unit measures of response (see below). Competition effect sizes as

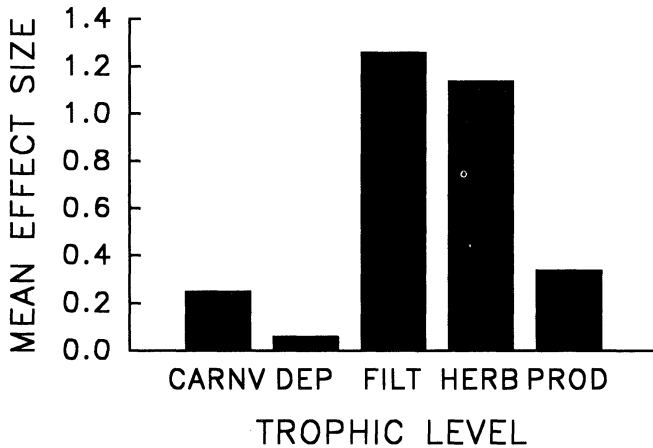


FIG. 1.—Mean effect size ( $d_+$ ) of competition for carnivores (CARNV), deposit feeders (DEP), filter feeders (FILT), herbivores (HERB), and primary producers (PROD) in manipulative field experiments. Responses to competition were measured as differences in biomass for organisms exposed to different levels of competition (see text).

defined here are therefore not directly equivalent to the familiar alphas from Lotka-Volterra competition equations.

### *Ecological Answers*

1. *What was the overall effect of competition?*—When all studies were considered together, competition had a large effect on biomass, with a grand mean effect size of  $d_+ = 0.80$  and a 95% CI of 0.77–0.83. Total heterogeneity among studies was extraordinarily high;  $Q_T = 4,955.24$  with 213 df (this value is not tabled, but it is statistically significant at far less than  $P = .001$ ). Thus while there is a large effect of competition overall and the CI for that effect does not overlap zero, we cannot be assured that all studies share a common effect size because of the substantial heterogeneity. This was as expected; one of the major purposes of this study was to examine the causes for differences in the effects of competition among studies. Three comparisons were omitted from the calculation of mean values and heterogeneities but are included in other analyses. These three comparisons were extreme outliers ( $d = 10.8, 6.2,$  and  $14.5$ ) and had fairly severe design problems from our perspective (pseudoreplicated and nonreplicated treatments and extreme imbalance in sample sizes among treatments), and it was felt that they distorted the overall values when included.

2. *Did competitive effects differ among trophic levels?*—The effect of competition differed substantially among trophic levels (fig. 1), with herbivores and filter feeders experiencing the greatest average effects, producers moderate effects, and carnivores and deposit feeders small effects on average. The differences between trophic levels were highly significant (table 1). The number of studies concerned with filter feeders and deposit feeders was too small for further analy-

TABLE 1  
EFFECTS OF COMPETITION AT FIVE TROPHIC LEVELS

| Trophic Level   | $d_+$ | 95% CI    | df  | $Q$      | $P$   |
|-----------------|-------|-----------|-----|----------|-------|
| Producers       | .34   | .29-.39   | 73  | 367.77   | <.001 |
| Herbivores      | 1.14  | 1.10-1.18 | 111 | 3,628.05 | <.001 |
| Deposit feeders | .06   | -.08-.19  | 2   | 34.87    | <.001 |
| Filter feeders  | 1.26  | .62-1.90  | 2   | 14.35    | <.001 |
| Carnivores      | .25   | .11-.38   | 21  | 116.77   | <.001 |
| All             | .80   | .77-.83   | 213 | 4,955.24 | <.001 |

NOTE.—Values reported are mean effect size ( $d_+$ ), 95% confidence interval (CI), degrees of freedom (df), heterogeneity statistic ( $Q$ ), and probability level ( $P$ );  $Q$  between trophic levels, 793.43; df, 4;  $P < .001$ .

sis. The effects of competition on primary producers, herbivores, and carnivores are examined in detail below.

These results refute theoretical predictions about the relative strength of competitive effects at different trophic levels. The modest levels of competition among carnivores and the high levels of competition among herbivores are especially surprising. The disagreement between data and theory may be because the theories are incorrect, but it may be because the data available do not provide an adequate test of the theories. In particular, the herbivore taxa that were the subject of most experimental studies may not have been what the theorists had in mind (see below for details on the herbivores). The scanty data that exist for terrestrial arthropods seem to support the predictions of Hairston et al. (1960).

3. *Were the measured effects of competition greater in experiments of longer duration?*—There was no tendency for larger effect sizes in studies of longer duration (fig. 2). There appears to be more scatter among values for short experiments ( $\leq 12$  mo), with few extreme values for  $d$  (positive or negative) reported for longer experiments. The greater variability among short studies might be because short studies are less reliable indicators of true effect sizes. Alternatively, because most studies are short, a greater diversity of organisms and systems are represented by short studies than by those of longer durations, which therefore might vary more.

4. *Did the number of replicates affect the outcome of experiments? How did density affect the measurement of competitive effects?*—There was also a tendency for experiments with smaller total sample sizes to vary more than those with larger samples (fig. 3). Few studies with large numbers of replicates had extreme (positive or negative) effect sizes. This is called a “funnel effect” because, as  $N$  increases, variation due to sampling error decreases and the values “funnel” down toward the true effect size (Light and Pillemer 1984, pp. 63–64).

It was difficult to examine density effects directly because most articles do not report the density of neighbors and because it is unclear how to equate densities of different kinds of organisms in interspecific competition. One group offered the chance to examine the relationship between density and competitive effects

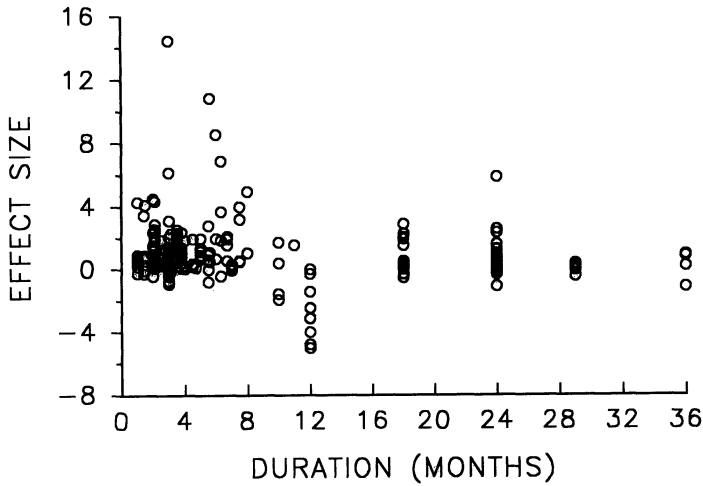


FIG. 2.—Magnitude of competitive effect size ( $d$ ) vs. the duration of each experiment. Each point represents a single comparison (see text).

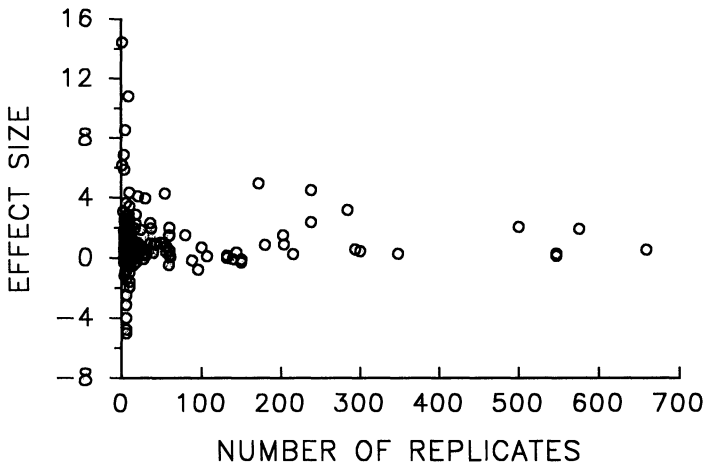


FIG. 3.—Magnitude of competitive effect size ( $d$ ) vs. the number of replicates used in the control or experimental group ( $N_c$  or  $N_e$ , whichever was smaller). Each point represents a single comparison.

in a straightforward manner. These were the experiments on intraspecific competition among herbivores stocked in enclosures. We defined relative density as the density closest to natural density divided by density at some level that was higher or lower than natural. Thus, a relative density less than 1.0 indicates a comparison with an enhanced density, and one greater than 1.0, a comparison of organisms at natural densities with those at reduced density. Density dependence exists (fig. 4) but is not very strong ( $r^2 = 0.17$ , slope = 0.33,  $P = .002$ ,  $k = 51$  data points).

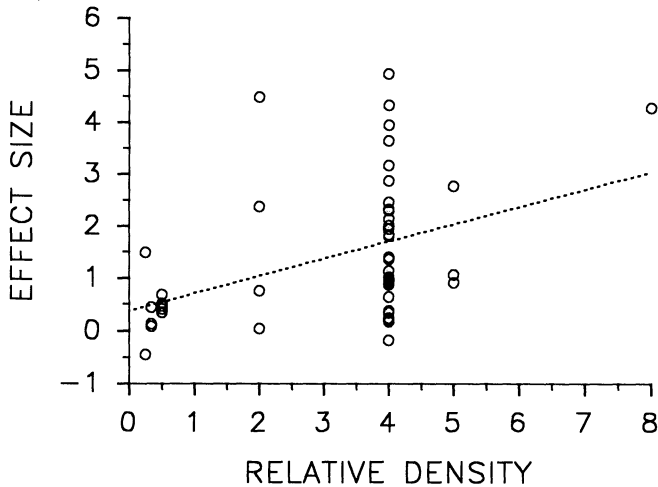


FIG. 4.—Competitive effect sizes ( $d$ ) for herbivores with regard to the relative intraspecific density at which they were maintained (1.0 = natural density).

This sort of relationship has been termed “disorderly” density dependence (Fowler 1990a).

It should be noted that the assumptions of standard regression procedures may not be met in synthesizing the results of independent studies and may result in inflated significance levels. Therefore, they are not strictly valid in meta-analysis (e.g., Hedges and Olkin 1985, pp. 11–12). However, this is not likely to overturn a conclusion of significance based on a probability level as small as  $P = .002$ , nor is it likely to result in accepting a false null hypothesis. We relied on them to give us ball park estimates of relationships because they are much easier to use than the general linear models developed for meta-analysis.

5. *Are the effects of competition dependent on the potential for growth without competitors?*—We examined the dependence of competitive effect size (defined above) on the size (mass) of organisms with no neighbors for experiments in which this information was available. We found no relationship between the effect of competition and size attained in the absence of competition (fig. 5). The lack of any dependence of competitive effects on size in the absence of competition is reflected in the low values of the correlation coefficient ( $r^2 = 0.005$ ) and slope (0.06,  $P = .65$ ,  $k = 43$ ; calculated for log (biomass); results were similar for nontransformed biomass values). The effect sizes commonly used in meta-analysis are an attempt to put different measures on a single scale by dividing by their standard deviations. It may be that this standardization accomplishes a purpose similar to that of measures proposed by Wilson and Tilman (1991)—putting the responses of larger or faster-growing organisms on the same scale as the responses of smaller ones.

6. *Do competing neighbors have a greater effect on large organisms than on small organisms?*—This question is related to the one above. To test it, we



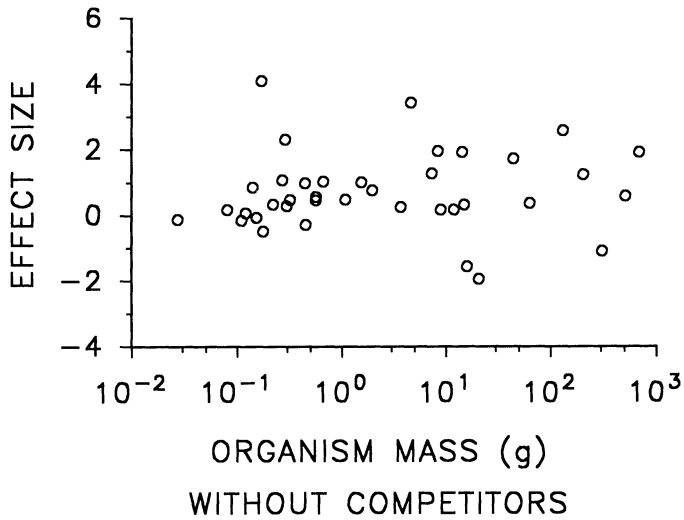


FIG. 5.—Effect sizes ( $d$ ) in relation to the mean mass in grams of the (target) organism in cases in which organisms were grown without competing neighbors (i.e., neighbor density was zero). Note that the X-axis is on a log scale.

regressed  $d$  on organism size (log of mean organism mass) at natural density. There was no relationship between organism mass and the magnitude of competitive effects (fig. 6;  $r^2 = 0.01$ ,  $P = .25$ ,  $k = 140$ ; recall that standard regression statistics may be indicative of relationships among variables but are not strictly valid in meta-analysis).

7. *Did the results in experiments in which organisms were enclosed differ from those in which they were not?*—There was a large difference in competitive effects between organisms that were free ranging or unenclosed, in comparison with those that were caged, stocked in enclosures, or similarly constrained (fig. 7). The difference between organisms in the two groups was statistically significant ( $Q_B = 584.58$ , 1 df,  $P \ll .001$ ). There are several explanations for this difference. Enclosed organisms may be forced to compete more than those that are free to move, or they may be more likely to be protected from predators that otherwise alleviate the effects of competition. However, as this comparison included all organisms, there is a strong bias in taxa for the two categories: primary producers were always unenclosed, for example, and freshwater herbivores were all enclosed. Therefore, the difference between caged and uncaged organisms may reflect only the differences among taxa. To correct for this, we also examined cage effects in marine mollusks, the single group that included experiments on both caged and uncaged animals.

Mollusk experiments offered a unique opportunity to compare the effects of competition for stocked, caged animals with those of a similar group of uncaged animals. There were sufficient numbers of interspecific comparisons with and without cages to evaluate cage effects per se (all intraspecific comparisons on

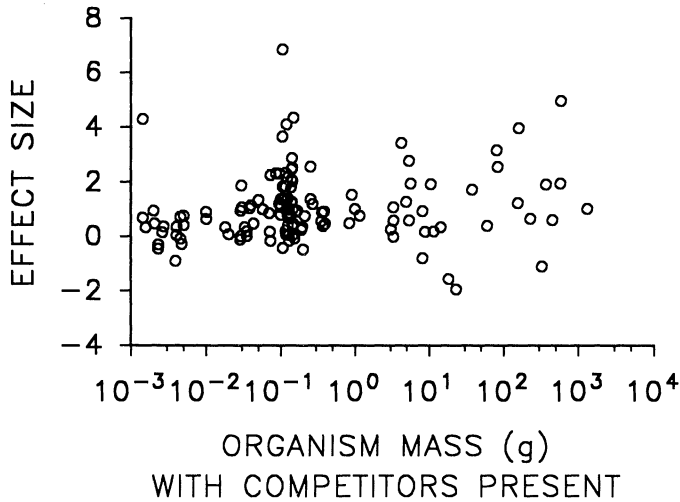


FIG. 6.—Effect sizes ( $d$ ) in relation to the mean mass in grams of the (target) organism in cases in which organisms were grown with approximately natural densities of competitors.

mollusks were made with stocked enclosures). Without cages, there was no detectable effect of competition on average; the mean effect of hypothesized competitors was not significantly different from zero (fig. 7). Caged animals, in contrast, experienced a medium effect of competition that was significantly greater than zero (fig. 7). The differences in the effects of interspecific competition among mollusks in caged versus unenclosed experiments were statistically highly significant ( $Q_B = 14.88$ , 1 df,  $P < .001$ ).

8. *Within trophic levels, did competitive effects differ among terrestrial, freshwater, and marine systems?*

9. *Did they differ between intra- and interspecific competition?*—To address these questions and several that follow, we examine in depth the effects of competition on primary producers and on carnivores, which were fairly straightforward, and then look at herbivores, which behaved in a more complex fashion.

*The effects of competition on primary producers.*—The moderate effects of competition on primary producers (table 1) did not differ significantly among terrestrial ( $k = 46$  comparisons), freshwater ( $k = 3$ ), and marine systems ( $k = 25$ ;  $Q_B = 0.42$ ,  $P > .50$ ). Intraspecific competitive effects ( $k = 15$  comparisons) did not differ from interspecific ( $k = 59$ ) effects ( $Q_B = 0.62$ ,  $P > .50$ ) in agreement with the findings of Goldberg and Barton (1992). When there is substantial heterogeneity among studies that cannot be explained by prior hypotheses, as there is for primary producers (table 1,  $Q_W = 367.77$ ), it is necessary to examine whether it is reasonable to believe that they share a common effect size (estimated by  $d_+$ ). If the experiments estimate a common effect, then effects will funnel to a mean value as sample size increases (Light and Pillemer 1984). The effect sizes for primary producers create a striking funnel (fig. 8). A second piece of evidence suggesting that it might be reasonable to characterize the responses of primary

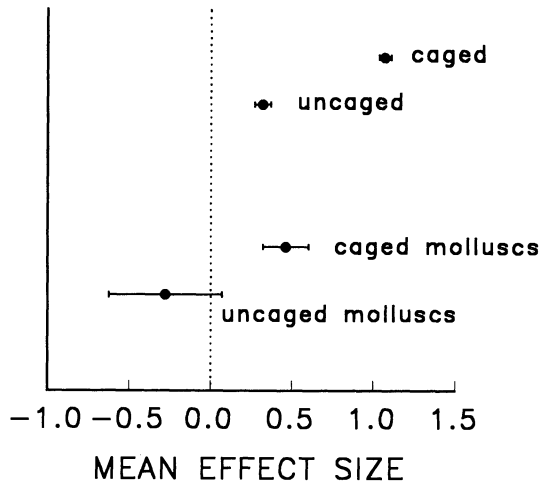


FIG. 7.—Mean effect size ( $d_+$ ) and 95% CI of competition for all caged or enclosed organisms in contrast with all uncaged organisms (*top*) and mean effect size ( $d_+$ ) of interspecific competition for marine mollusks in caged vs. uncaged trials (*bottom*).

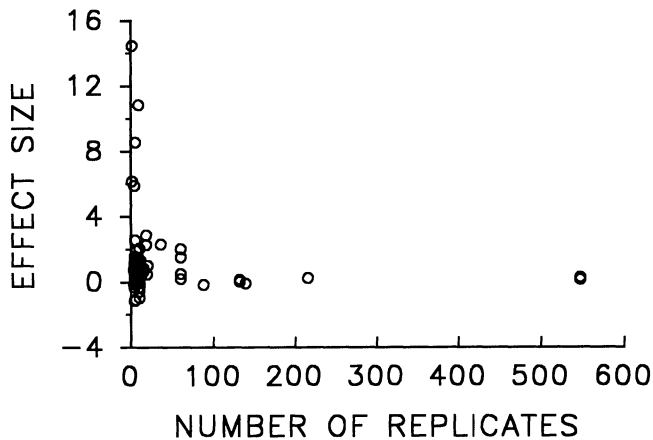


FIG. 8.—Effect sizes ( $d$ ) of competition for primary producers with respect to the number of replicates used in the control or experimental group ( $N_c$  or  $N_e$ , whichever was smaller). Each point represents a single comparison. Note that many points are obscured from view because they are clustered at low replicate numbers and slightly positive effect sizes.

producers with a common mean is the histogram of effects (fig. 9). Five comparisons (of 74 total) had very large effects ( $d \geq 4.0$ ); four of these had very small sample sizes (fig. 8) and the fifth had a serious design flaw (discussed above). Excepting these five, it is possible that the histogram (fig. 9) represents a distribution with a single mean and variance.

To estimate the mean and CI of  $d$  for primary producers more accurately, we

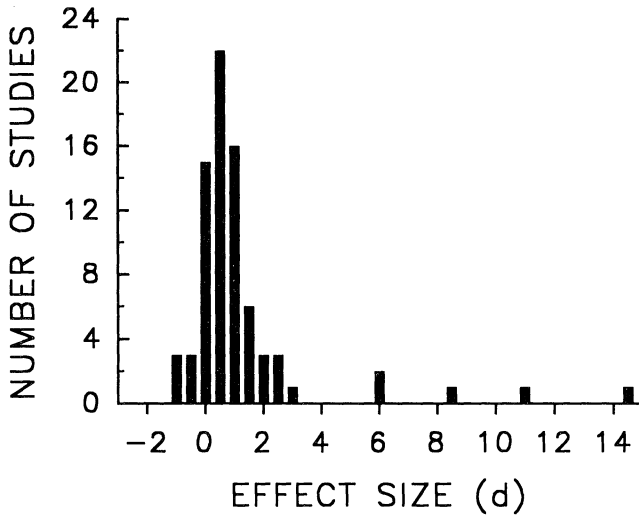


FIG. 9.—Histogram of effect sizes of competition among primary producers

eliminated all comparisons in which the number of replicates in either the controls or the experimental treatments was less than 10. The value of  $d_+$  for the remaining 41 comparisons was essentially unchanged ( $d_+ = 0.32$  with a 95% CI of 0.27–0.37). Eliminating small samples also eliminated two of the four lowest values (i.e.,  $< -0.8$ ) and nine of the eleven highest values (i.e.,  $> 1.6$ ) but did not greatly reduce the heterogeneity ( $Q_W = 300.48$ , 40 df). If we assume that the studies within a category do not share a single true effect size but rather are samples from a distribution of true effects, it is more appropriate to calculate the mean effect using a random effects model (Hedges 1983; Hedges and Olkin 1985). Calculating the mean effect size,  $\bar{d}$ , using a random effects model (Hedges and Olkin 1985, pp. 193–199), we obtained a mean effect size of  $\bar{d} = 0.53$  with a 95% CI of 0.36–0.71 and 40 df. We can conclude that the effect of competition on primary producers is moderate and statistically significant (the CI does not overlap zero). These results are substantively in agreement with those of Goldberg and Barton (1992). Except for small sample sizes and unbalanced designs, plant ecologists committed few errors in experimental design.

It is not surprising that primary producers appear to share a common, moderate effect of competition. All plants have the same basic requirements for resources, and on average the intensity of competition may not depend on the species identity of neighbors. Coexisting species will not coexist for long if the effects of competition are extremely high. While in some cases the effects are very small, the removal of neighbors (or equivalent experiments) appears to have on average a medium effect on plant biomass.

10. *Was competition more intense for primary producers in high- than in low-productivity habitats?*—We tested the hypothesis that competitive effects are greater among terrestrial plants in productive systems by comparing experiments

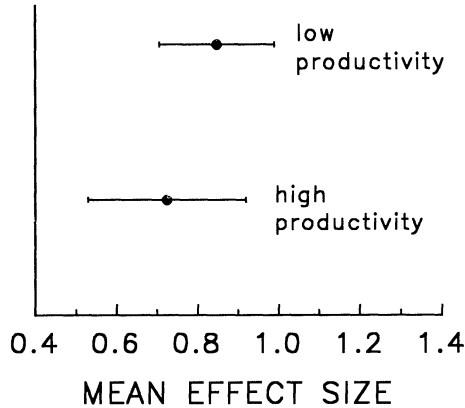


FIG. 10.—Mean effect size ( $d_+$ ) and 95% CI for primary producers in low-productivity habitats in contrast with high-productivity habitats.

performed in deserts and arctic systems ( $k = 11$ ) with those in prairies, meadows, and old fields ( $k = 23$ ). Competition actually had a smaller effect in the high-productivity systems than in low-productivity systems, although there were large effects in both (fig. 10), and the differences between them were not statistically significant ( $Q_B = 0.97$ , 1 df).

*The effects of competition on carnivores.*—While there was significant heterogeneity in the effects of competition among carnivores ( $Q = 116.77$ , 21 df), species appear to possibly share a common effect (fig. 11). There were two outliers with high values ( $>1.9$ ) and two outliers with low values ( $<-1.5$ ). All four outliers had pseudoreplicated designs. Following the recommendation of Hedges and Olkin (1985, pp. 248–252), we examined the results when these four outliers were omitted. For the remaining comparisons,  $d_+ = 0.32$  with a 95% CI of 0.18–0.47. This reduced the differences among studies in effect sizes to a marginally significant value ( $Q = 33.91$ , 17 df,  $P = .01$ ; recall that we are being very conservative about significance levels, as per the discussion above).

The data available on competition among carnivores included five species of terrestrial carnivores (spiders, scorpions, and marsupials) studied by two authors and reported in three publications and seven species of freshwater carnivores (insects and fish) published in five articles. Carnivores were studied much less frequently than primary producers or herbivores. Experiments on carnivores were conducted in terrestrial ( $k = 9$ ) and freshwater systems ( $k = 13$ ). All experiments on terrestrial carnivores were on interspecific competition while both inter- and intraspecific effects were studied in freshwater systems. The differences between terrestrial carnivores and freshwater carnivores were not statistically significant (with outliers removed as above,  $Q_B = 3.32$ , 1 df). There were no significant differences for inter- versus intraspecific competition in freshwater systems ( $Q_B = 2.18$ , 1 df).

11. *Did results differ between studies that were well designed and those with problems in experimental design?*—Carnivore experiments offered an opportu-

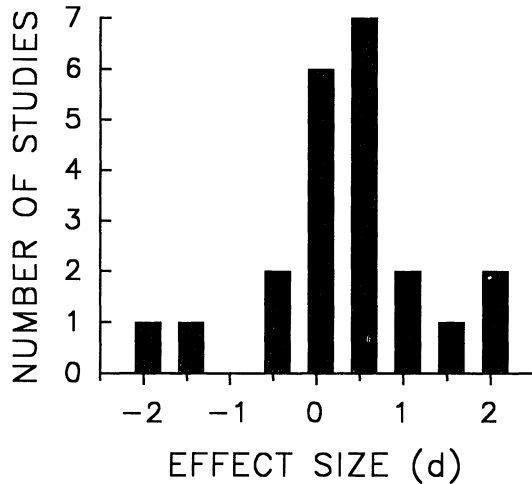


FIG. 11.—Histogram of effect sizes of competition among carnivores

nity to test the effects of problems in experimental design since there were a substantial number of both studies with good designs and also studies with problems in design. (However, the studies with design problems were all on vertebrates, and those without serious problems were all on arthropods.) Competition had a small but statistically significant effect on carnivores in experiments with good design or only minor design problems (fig. 12). As a group, in studies that were pseudoreplicated or in which the controls were not replicated, the average effects of competition were not distinguishable from zero (fig. 10). That is, removing the more poorly designed studies resulted in an improvement in the ability to detect the real effects of competition among carnivores, which were otherwise obscured. Furthermore, the heterogeneity was more than four times as great among poorly designed studies ( $Q_W = 96.40$ , 9 df) as among studies with good design ( $Q_W = 20.23$ , 11 df). The differences between mean effect size in well- and poorly designed studies were not statistically significant ( $Q_B = 0.13$ , 1 df).

*The effects of competition on herbivores.*—While the effect of competition on the biomass of herbivores was large ( $d_+ = 1.14$ ), the heterogeneity in this group was very high—more than an order of magnitude greater than in other trophic levels (table 1), which suggests that all herbivores did not behave alike and prompts a somewhat different approach to analyzing the data. In examining the distribution of effect sizes, it became apparent that it made more sense to separate responses into taxa within systems than to divide herbivores strictly into terrestrial, freshwater, and marine systems. That was because the number of taxa studied within each system was very limited, and lumping the few taxa within each made little sense biologically or statistically. The herbivores studied fell into four broad groups: anurans, mollusks, echinoderms, and arthropods. The differences among the four groups were large and statistically significant (table 2).

Anurans (frogs and toads) experienced large competitive effects (table 2; one

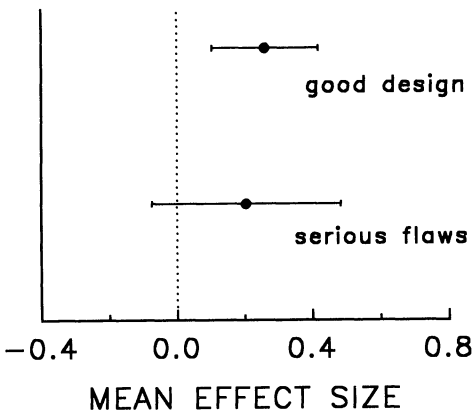


FIG. 12.—Mean effect size ( $d_+$ ) and 95% CI for carnivores in experiments with good experimental designs or only minor design problems in contrast with those in experiments with serious problems in experimental design.

TABLE 2  
EFFECTS OF COMPETITION FOR DIFFERENT HERBIVORE GROUPS

| Organisms     | $d_+$ | 95% CI    | df | $Q$      | $P$   |
|---------------|-------|-----------|----|----------|-------|
| Anurans       | 1.15  | 1.10–1.19 | 14 | 1,710.95 | <.001 |
| Mollusks:     |       |           |    |          |       |
| Intraspecific | .69   | .56–.82   | 33 | 151.19   | <.001 |
| Interspecific | .36   | .23–.49   | 45 | 360.49   | <.001 |
| Echinoderms   | .60   | .27–.93   | 3  | 7.59     | <.01  |
| Arthropods:   |       |           |    |          |       |
| Freshwater    | 2.13  | 2.00–2.25 | 5  | 707.98   | <.001 |
| Terrestrial   | –.04  | –.16–.08  | 5  | 13.76    | <.025 |

NOTE.—Abbreviations as in table 1;  $Q$  between taxa, 340.26; df, 3;  $P$  < .001.

study was omitted before analysis because of a highly unbalanced design). Effect sizes ranged from 0.29 to 4.94, and sample sizes were large (100–1,350). Responses for eight species in four genera were studied. All experiments were conducted in stocked artificial tanks, and lasted from 1 mo to just over 6 mo. Heterogeneity among effects was large and statistically significant (table 2). When the heterogeneity is this great, it is reasonable to question whether the mean and variance of the group really convey useful information or whether the responses are so diverse that the organisms are not behaving in such a way that they can effectively be grouped together. For example, if the effects are strongly bimodal or if the distribution of effect sizes is essentially “flat” (i.e., there is a uniform distribution of effect sizes), the mean may be a misleading indicator of the behavior of the group. The histogram of the effect sizes for this group, however, is fairly reassuring. The mean and confidence limits of these effect sizes provide a reasonable summary of their distribution, and it appears that it is sensible to

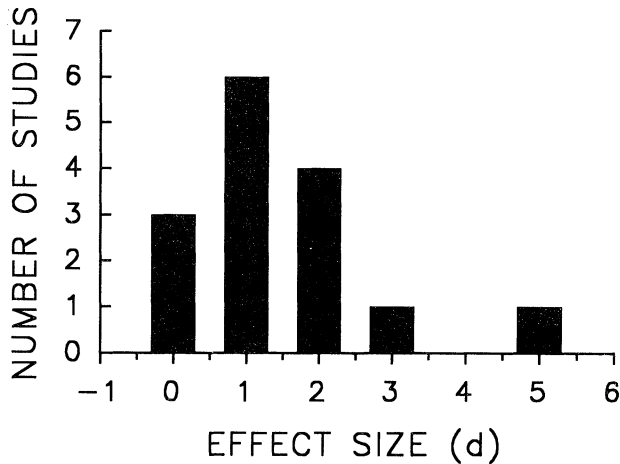


FIG. 13.—Histogram of effect sizes of competition in anurans

speak of the performance of frogs and toads as a group (fig. 13). The large value obtained for the heterogeneity statistic is in one sense an artifact of the large sample sizes in this group: even relatively small differences in effects will result in high heterogeneity when sample sizes are large (Hedges and Olkin 1985).

Although there was a statistically significant difference between inter- and intra-specific competitive effects ( $Q_B = 67.88$ , 1 df; intraspecific,  $d_+ = 1.43 \pm 0.08$ ; interspecific,  $d_+ = 1.02 \pm 0.05$ ), the distributions of these subgroups did not indicate that they were “good” groups (results not shown). Removing a single outlying value (with  $d \approx 5$ ) substantially reduced the difference between them ( $Q_B = 7.64$ ,  $.01 < P < .005$ ; intraspecific,  $d_+ = 1.157 \pm 0.17$ ; interspecific was unchanged), as well as reducing the mean effect size for anurans slightly ( $d_+ = 1.06$  with a CI of  $1.01$ – $1.11$ ). Consequently it seems most reasonable to calculate a mean effect size for all frogs and toads (except for the two comparisons omitted as discussed above) with a random effects model (Hedges 1983; Hedges and Olkin 1985); that is, we should assume that the individual measures of effects are samples from a distribution of true effects. The values thus obtained for anurans as a group are  $\bar{d} = 1.178$ , with  $k = 14$  and a CI of  $0.737$ – $1.619$ . This is a very large effect that is very similar to the estimate made with the fixed effects model.

Mollusks as a group experienced moderate competitive effects (table 2). All organisms lived in marine habitats. Sample sizes were predominately small (3 to ca. 120 replicates per treatment). The histogram of effect sizes could conceivably describe a single coherent group, but there are very long tails and possible multiple peaks (fig. 14). This is reflected in the substantial and statistically significant heterogeneity value (table 2). Intraspecific competitive effects were significantly greater than interspecific effects (table 2;  $Q_B = 12.22$ ,  $P < .001$ ). But breaking mollusks into inter- and intraspecific comparisons did not resolve the large heterogeneity among values within each group. Interspecific effect sizes ranged from  $-5.055$  to  $6.850$ . There was only a single negative value for an intraspecific



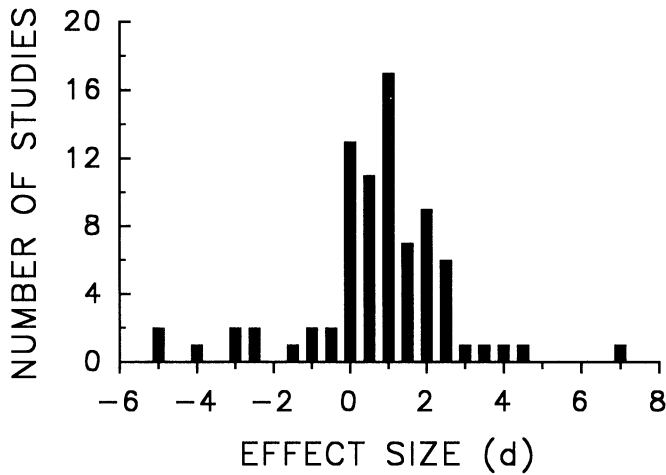


FIG. 14.—Histogram of effect sizes of competition in mollusks, including inter- and intra-specific comparisons.

comparison ( $-0.17$ ) and only two very large values ( $>3.0$ ). Using a random effects model to calculate the means and variances, we found a moderate average effect of interspecific competition ( $\bar{d} = 0.301$ ) with a 95% CI that overlapped zero ( $-0.317$  to  $0.919$ ). Intraspecific competition, in contrast, had a very large effect ( $\bar{d} = 1.175$ ) that was significantly greater than zero (95% CI,  $0.795$ – $1.556$ ).

Experiments on mollusks suffered from a number of problems in experimental design and reporting of data. Problems that were especially common in mollusk experiments were small number of replicates, no treatments at natural density for comparison (i.e., no real “control”), and measurements pooled for all experimental organisms without consideration of the cage from which they were taken (resulting in an incorrect estimate of the standard deviation and thus a potentially inaccurate estimate of  $d$ ). Most troubling was that in many cases it was necessary to estimate sample sizes from assorted “clues” in the articles because these basic data were not provided by the authors. Undoubtedly this increased the difficulty (and heterogeneity) in evaluating the results of the experiments.

Only four measures of competitive effects on echinoderms were found. These organisms had a moderate mean effect size (table 2) that was significantly greater than zero. The  $Q$  statistic indicates marginally significant heterogeneity ( $.005 < P < .01$ ), the result of three effect sizes very close to  $d = 0.5$  and a single high value of  $d = 2.78$ . Two comparisons had moderate sample sizes ( $N = 24$  or  $48$ ) and two (including the one with a large effect size) had very small sample sizes ( $N = 6$ ). Based on this small group of experiments it is reasonable to conclude that echinoderms experience a medium effect of competition on biomass.

Arthropods belonged to two distinct groups: larval forms living in freshwater streams and terrestrial forms. There was a highly significant difference between these groups ( $Q_B = 152.98$ ;  $P \ll .001$ ), and we examined them separately. Both groups were represented by only a small number of comparisons. Lotic arthro-

pods had a very large mean effect size (table 2). Values for individual effect sizes ranged from  $-0.449$  to  $+4.495$ . Sample sizes ranged from medium to large (ca. 26–475), although as with mollusk studies it was necessary to estimate these numbers because in many cases they were not reported directly. The experiments examined intraspecific competitive effects for two species in two publications. All organisms were stocked in enclosures and, as is appropriate considering their brief lives, were studied in experiments of very short duration (1–2 mo).

In spite of the large theoretical interest in the effects of competition on terrestrial herbivores, we found only a single study that experimentally investigated biomass effects for this group (Karban 1989). Results for the single insect species studied, *Philaenus spumarius* males and females, revealed a wide range of effects ( $-0.966$  to  $+3.566$ ), with a mean effect that was not distinguishable from zero (table 2). The marginally significant heterogeneity among studies (at  $.01 < P < .025$ ) is due to a single large effect size, which was based on a small number of replicates ( $N_e = 2$ ,  $N_c = 6$ ). Other sample sizes were small to moderate (5–24). These experiments lasted for 3 mo.

### *Causes of Heterogeneity in Effect Sizes*

As discussed above, heterogeneity among studies may be troubling because it raises questions about whether all experiments are truly estimating a single underlying mean value. On the other hand, investigating causes for heterogeneity among studies may be of even greater interest than coming up with precise estimates of mean effect size (Light and Pillemer 1984). Abrami et al. (1988) point out that tests of heterogeneity are especially sensitive to differences among effect sizes when the number of study outcomes and total sample sizes are large, as they were in the present study.

The statistics we used assume normal distributions of the primary data that are the subject of the meta-analysis (i.e., it is assumed that  $\bar{Y}_e$ ,  $\bar{Y}_c$ ,  $s_e$ , and  $s_c$  were calculated by the original authors on normally distributed data). Although essentially all parametric statistics share this assumption, most authors do not report whether they tested their data for meeting this assumption, whether the data were transformed prior to analysis, or whether reported means and standard deviations were calculated on transformed data and back transformed. Unfortunately, it is not known how robust meta-analytic statistics are to violations of these assumptions, but this must surely add to the noise.

As our results show, poor experimental design and analysis and unclear reporting of data can inflate heterogeneity. Studies with small samples and short durations also may increase variability. But heterogeneity in the effects of competition in field experiments can reflect real (as opposed to artifactual) differences experienced by the organisms. We chose data to include a wide range of conditions under which organisms compete. Studies included experiments over a range of densities, at different seasons, at different water depths, in fertilized and unfertilized plots, with and without predators, and so on to test theoretical predictions or to elucidate the full range of responses. Certainly the data reflect that variability. In some cases, as for some of the groups examined above, it makes more sense to view the effect size as a random variable. The range of competitive

effects for a group of organisms may be revealing when it is presented along with the mean and CI.

While heterogeneity may be “explained” satisfactorily in many instances, sometimes it alerts the reviewer to patterns of variation that suggest that she or he is not looking at a group that really belongs together. Meta-analysis is a unique instrument for quantifying the variation among studies so that such questions can be legitimately addressed.

#### OBSERVATIONS AND CONCLUSIONS

This study makes it clear that more organisms need to be studied in underrepresented systems. We know quite a lot about competition among terrestrial plants and among marine mollusks and almost nothing about terrestrial herbivores or carnivores of any kind. Intraspecific competition needs to be examined much more often. We need to devise ways of studying mobile animals, particularly in aquatic systems, without enclosing them. The relationship between how much organisms compete in nature and how much they compete in field experiments remains unknown for these and other reasons discussed above.

Meta-analysis can reveal some of the gaps and inadequacies in a field of literature perhaps more starkly than ordinary reviews. It is clear from our study that the quality of the ecological literature will increase substantially if there are standard requirements for reporting basic statistical data, such as including standard deviations (or some measure of variance) and sample sizes when reporting means, unless there is a good reason not to do so. One does not have to surmise on what pages a cited article is to be found, and one should not have to surmise what the sample sizes were when means are reported (Fowler 1990*b*). Published graphs are usually the only public record of the outcome of an experiment. They should convey information in as straightforward a manner as possible, including the use of standard scales and easy-to-read axes. Dazzling three-dimensional designs and imaginative scales of measurement make it difficult or impossible to use the data presented. A number of studies used extremely unbalanced designs, with 30–50 times as many replicates in one treatment as another. Perhaps in some cases this cannot be avoided, but the consequence is that the performance of one group is known with much greater precision than the other, and therefore the accuracy of the comparison between them is compromised. Careful thought to the allocation of replicates would improve the reliability of ecological data.

Ecological experiments have relied increasingly on more and more complex experimental designs: greater numbers of species, locations, densities, and interactions among these factors. Sometimes such designs offer more revealing information and provide better tests of theory. However, financial and other limitations mean that complex experiments typically have small sample sizes. In a very real sense, we are often getting less and less information from ecological experiments as they become more complex, because our ability to differentiate among the responses to the treatments is diminished when sample sizes become small. Meta-analysis offers an alternative, emphasizing the value of large, simple studies (Mann 1990). No one experimentalist can study all organisms in all sys-

tems under all conditions. If we can synthesize the results of independent studies, the collective efforts of all experimental ecologists may prove a much more powerful means by which major questions can be resolved in ecology.

#### ACKNOWLEDGMENTS

We thank D. Slice for sharing his expert help with digitizing published graphs, and F. J. Rohlf for the use of his digitizing equipment. E. Bauder, A. Desrochers, R. Karban, W. Morris, and R. Ryti generously provided additional data not available in their articles. We appreciate the advice, comments, and encouragement of colleagues, including F. S. Chapin, D. Dykhuisen, D. E. Goldberg, L. V. Hedges, R. Karban, J. A. Morrison, T. Postol, J. Thomson, and an anonymous reviewer. This research was supported by National Science Foundation grants BSR 89-08112 and BSR 91-13065 to J.G. This is contribution 800 in Ecology and Evolution at the State University of New York at Stony Brook.

#### APPENDIX A

##### SELECTION OF ARTICLES FOR INCLUSION IN THE META-ANALYSIS

The criteria for including data from an article in the meta-analysis were as follows:

1. *It was a field experiment.*—Our criteria were similar to those of Schoener (1983). A field experiment is subject to natural variability in biotic and abiotic interactions. Outdoor "garden" experiments were excluded unless the focus organisms were heterotrophs that were living or feeding on the cultivated plants. Experiments occurring in outdoor artificial ponds stocked with animals, plants, and other matter characteristic of natural pond habitats were included, as were studies in which animals were stocked in cages in their natural habitats (but see text for a test of stocked vs. unstocked experiments).

2. *The experimenter directly manipulated densities of at least one presumed competitor species.*—The two types of manipulations were reduction, in which the natural density of a species was reduced either partially or completely, and enhancement, which involved either an intraspecific- or an interspecific-density increase. If a density manipulation was independently established in a variety of places, during different times, under various resource and predator regimes, or some combination of these, we included all appropriate comparisons in an effort to represent the variety of conditions under which competition may naturally occur.

The meta-analysis models we used are designed for simple two-category comparisons, which complicated the use of studies with factorial designs. We quantified only the effects of density manipulations and not the effects of interactions with other treatments. If competition was tested under different resource or predator regimes, a control for competitor density manipulations must have been included at each resource or predator level. Each such manipulation was treated as a separate comparison (see text for a discussion of nonindependence of comparisons). Experiments in which *only* resource or predator levels were increased or decreased were excluded, as were those in which entire communities were transplanted without directly manipulating the densities of hypothesized competitors.

3. *A response was measured on a specific organism that was identified to genus and species or at least to the same unknown species.*—Experiments that measured one response for a group of organisms, such as "mosses" or "grasses," were excluded.

4. *The response was presented as either means and some measure of variance (which could be converted to a standard deviation) with a known sample size, or the actual data from which means, standard deviations, and sample sizes could be determined.*—The data were obtained from text, tables, and figures. Additionally, a few authors of recent articles

(1987 through 1989) furnished unpublished data, such as means, standard deviations and errors, and sample sizes. Results that were presented only as correlation coefficients, regression coefficients, medians, test statistics, and probability values were not used.

5. *The response measured was one that is clearly indicative of an effect on the organism's fitness.*—For simplicity we excluded comparisons in which interpreting the response required further assumptions, for example, responses based on optimality or niche theory, such as size of home range or size of individual prey. The data were sorted according to the type of response measured. For this meta-analysis, only responses expressed in terms of the biomass (including percent cover of mature plants) of the responding organisms were used. This included the largest number of comparisons by a substantial margin. The results of other responses to competition (density, mortality, length, growth, reproductive output, recruitment, feeding, and physiological performance) are examined elsewhere (J. Gurevitch, unpublished manuscript; see Morrow 1990).

6. *The treatment had an appropriate control that was performed simultaneously and in the same place.*—The control was the density closest to natural conditions, as specified by the author. If two types of organisms naturally co-occurred, "coexistence" was used as the "control" and (experimentally established) "noncoexistence" as the "experimental" treatment.

When a series of manipulations were performed in which densities were enhanced or reduced at various levels (e.g.,  $0.1 \times$ ,  $0.5 \times$ ,  $2 \times$ , and  $4 \times$  natural densities), the lowest reduction and the highest enhancement were compared with the control (the average natural density). We assumed that the authors chose relevant experimental densities for their particular research organisms and made appropriate measurements on them. If all densities within an experiment were within the natural range, the lowest and highest densities were treated as the "control" and "experimental" densities, respectively. Our "control" was not always the one identified as such by the author(s), particularly when the density manipulation was not the chief focus of the study.

Two or more comparisons may have shared the same control. Duplicate use of "controls" occurred because an experiment involved both a density enhancement and a density reduction or because the effects of more than one hypothesized competing species was tested (see text for a discussion of nonindependence of observations).

Sometimes investigators repeatedly sampled the individuals or plots over the course of the experiment. The data used in such cases were only those measured at the conclusion of the experiment or at the end of the density manipulation.

7. *The experiment was free of severe design problems.*—Two such problems were changing competitor identity but not density and failing to replicate at all. Pseudoreplicated studies (Hurlbert 1984) and studies with other moderate design problems were included if there were no other design problems. Such studies were flagged for further examination (see text).

## APPENDIX B

### DETERMINATION OF STANDARD DEVIATIONS FROM INDIVIDUAL EXPERIMENTS

Values used for standard deviations in the calculation of effect sizes were calculated as follows:

When the data reported for the control and for the experimental treatment were means, standard deviations or standard errors, and numbers of individuals measured, then  $n_{ij}$  is the number of organisms in the  $j$ th plot of the  $i$ th treatment or control and  $a_i$  is the number of plots assigned to the  $i$ th treatment or control. If all measured individuals in a treatment were combined without regard to cages or plots, the mean and error were based on  $\sum_{j=1}^a n_{ij}$  individuals. First, if the *mean and standard deviation* of each treatment were reported, these data were used directly. Second, if the *mean and the standard error* of

each treatment were reported, we calculated the standard deviation:

$$s_Y = \text{SE} \left( \sum_{j=1}^a n_{ij} \right)^{1/2}, \quad (\text{B1})$$

where  $\text{SE} \approx s_{\bar{y}_i}$ , which is an estimate of the standard error of the individual measurements.

In other cases, individuals were averaged for each plot separately, and  $n_{i1} = n_{i2} = \dots = n_{ia}$ . First, when the *mean and standard deviation or standard error* of each plot were reported, we calculated the grand mean,  $\bar{\bar{Y}}_i$ , for each treatment (i.e., the mean of the plot means), as

$$\bar{\bar{Y}}_i = (\bar{Y}_{i1} + \bar{Y}_{i2} + \dots + \bar{Y}_{ia})/a_i, \quad (\text{B2})$$

and the standard deviation,  $s_{\bar{y}_i}$ , of the plot means,  $\bar{Y}_{i1} \dots \bar{Y}_{ia}$ , as

$$s_{\bar{y}_i} = [\Sigma(\bar{Y}_{ij} - \bar{\bar{Y}}_i)^2/(a_i - 1)]^{1/2}. \quad (\text{B3})$$

The standard deviation of individual measurements,  $s_Y$ , was estimated by multiplying  $s_{\bar{y}_i}$  by the square root of the number of individuals per plot:

$$s_Y = s_{\bar{y}_i}(n_{ij})^{1/2}. \quad (\text{B4})$$

Second, some authors reported the *grand mean and standard error, but the standard error was calculated on the basis of the number of plots*. Although the response was clearly measured on individual organisms (e.g., biomass per individual, shell length, growth), the  $n$  on which standard error was calculated in a number of studies corresponded to the number of plots or cages for the treatment rather than to the number of individuals. The reported standard error was actually an approximation of the standard error of the means,  $s_{\bar{y}_i}$ . To obtain the standard deviation of *individual (organism)* measurements, the reported standard error was multiplied by the square root of the number of plots to get the standard error of the individual measurements,  $s_{\bar{y}}$ . This quantity was then multiplied by the square root of the number of individuals per plot:

$$s_Y = s_{\bar{y}}(a_i n_{ij})^{1/2}. \quad (\text{B5})$$

The *means and an ANOVA table* were reported in some articles, and the degrees of freedom corresponded to the number of individuals measured. The square root of the within-group mean square was used to calculate an estimate of the standard deviation:

$$s_Y = (\text{within-group mean square})^{1/2}. \quad (\text{B6})$$

If each plot was averaged and at least one sample size was not equal to the others, a weighted grand mean,  $\bar{\bar{Y}}_{wi}$ , a weighted variance of the means,  $s_{\bar{y}^2_{wi}}$ , and a weighted average  $n$ ,  $n_{0i}$  (Sokal and Rohlf 1981, p. 214), were calculated.

## APPENDIX C

TABLE C1

SOURCES OF DATA USED IN THE META-ANALYSIS

| Source                      | Journal  |
|-----------------------------|--|
| Alford 1989                 | Ecology  |
| Alford and Wilbur 1985      | Ecology  |
| Armesto and Pickett 1985    | Ecology  |
| Bauder 1989                 | Ecology  |
| Chapin et al. 1989          | Oecologia  |
| Choat and Andrews 1986      | Oecologia (Berlin)                                 |
| Chow 1989                   | Journal of Experimental Marine Biology and Ecology |
| Creese and Underwood 1982   | Oecologia (Berlin)                                 |
| Crowley et al. 1987         | Oecologia (Berlin)                                 |
| Dickman 1986                | Ecological Monographs                              |
| Ehleringer 1984             | Oecologia (Berlin)                                 |
| Ellison 1987 <sup>a</sup>   | Ecology  |
| Ellison 1987 <sup>b</sup>   | Ecology  |
| Fletcher 1987               | Ecological Monographs                              |
| Fletcher and Underwood 1987 | Ecology  |
| Fowler 1986                 | Ecology  |
| Heil and Bruggink 1987      | Oecologia (Berlin)                                 |
| Hill and Knight 1987        | Ecology  |
| Jara and Moreno 1984        | Ecology  |
| Jernakoff 1985              | Journal of Experimental Marine Biology and Ecology |
| Johansson 1987              | Oecologia (Berlin)                                 |
| Johnson and Mann 1988       | Ecological Monographs                              |
| Karban 1989                 | Ecology  |
| Kastendiek 1983             | Journal of Experimental Marine Biology and Ecology |
| Lamberti and Resh 1987      | Oecologia (Berlin)                                 |
| Lee and Bazazz 1982         | Ecology  |
| Levitan 1989                | Ecology  |
| Lively and Raimondi 1987    | Oecologia (Berlin)                                 |
| McConaughay and Bazzaz 1987 | Ecology  |
| McGraw and Chapin 1989      | Ecology  |
| Morin 1986                  | Ecology  |
| Morin et al. 1988           | Ecology  |
| Morrisey 1987               | Journal of Experimental Marine Biology and Ecology |
| Ortega 1985                 | Journal of Experimental Marine Biology and Ecology |
| Persson 1987                | Oecologia (Berlin)                                 |
| Petratis 1989               | Journal of Experimental Marine Biology and Ecology |
| Pierce et al. 1985          | Ecology  |
| Polis and McCormick 1986    | Oecologia (Berlin)                                 |
| Polis and McCormick 1987    | Ecology  |
| Quinn and Ryan 1989         | Journal of Experimental Marine Biology and Ecology |
| Spence 1986                 | Oecologia (Berlin)                                 |
| Sutherland and Ortega 1986  | Journal of Experimental Marine Biology and Ecology |
| Titus and Stephens 1983     | Oecologia (Berlin)                                 |
| Turner and Lucas 1985       | Journal of Experimental Marine Biology and Ecology |
| Underwood 1984              | Oecologia (Berlin)                                 |
| Wilbur and Alford 1985      | Ecology  |

## LITERATURE CITED

- Abrami, P. C., P. A. Cohen, and S. d'Appollonia. 1988. Implementation problems in meta-analysis. *Review of Educational Research* 58:151-179.
- Alford, R. A. 1989. Variation in predator phenology affects predator performance and prey community composition. *Ecology* 70:206-219.
- Alford, R. A., and H. W. Wilbur. 1985. Priority effects in experimental pond communities: competition between *Bufo* and *Rana*. *Ecology* 66:1097-1105.
- Armesto, J. J., and S. T. A. Pickett. 1985. Experiments on disturbance in old-field plant communities: impact on species richness and abundance. *Ecology* 66:230-240.
- Bauder, E. 1989. Drought stress and competition effects on the local distribution of *Pogogyne abramsii*. *Ecology* 70:1083-1089.
- Bender, E. A., T. J. Case, and M. E. Gilpin. 1984. Perturbation experiments in community ecology: theory and practice. *Ecology* 65:1-13.
- Chapin, F. S., J. B. McGraw, and G. R. Shaver. 1989. Competition causes regular spacing of alder in Alaskan shrub tundra. *Oecologia* 79:412-416.
- Choat, J. H., and N. L. Andrews. 1986. Interactions amongst species in a guild of subtidal benthic herbivores. *Oecologia (Berlin)* 68:387-394.
- Chow, V. 1989. Intraspecific competition in a fluctuating population of *Littorina plena* Gould (Gastropoda: Prosobranchia). *Journal of Experimental Marine Biology and Ecology* 130:147-165.
- Cohen, J. 1969. Statistical power analysis for the behavioral sciences. Academic Press, New York.
- Connell, J. H. 1975. Some mechanisms producing structure in natural communities: a model and evidence from field experiments. Pages 460-490 in M. Cody and J. Diamond, eds. *Ecology and evolution of communities*. Harvard University Press, Cambridge, Mass.
- . 1983. On the prevalence and relative importance of interspecific competition: evidence from field experiments. *American Naturalist* 122:661-696.
- . 1990. Apparent versus "real" competition in plants. Pages 9-26 in J. B. Grace and D. Tilman, eds. *Perspectives on plant competition*. Academic Press, New York.
- Creese, R. G., and A. J. Underwood. 1982. Analysis of inter- and intra-specific competition amongst intertidal limpets with different methods of feeding. *Oecologia (Berlin)* 53:337-346.
- Crowley, P. H., P. M. Dillon, D. M. Johnson, and C. N. Watson. 1987. Intra-specific interference among larvae in a semivoltine dragonfly population. *Oecologia (Berlin)* 71:447-456.
- Dickman, C. R. 1986. An experimental study of competition between two species of dasyurid marsupials. *Ecological Monographs* 56:221-241.
- Donoghue, M. J. 1989. Phylogenies and the analysis of evolutionary sequences, with examples from seed plants. *Evolution* 43:1137-1156.
- Ehleringer, J. R. 1984. Intraspecific competitive effects on water relations, growth, and reproduction in *Encelia farinosa*. *Oecologia (Berlin)* 63:153-158.
- Ellison, A. M. 1987a. Effects of competition, disturbance, and herbivory on *Salicornia europaea*. *Ecology* 68:576-586.
- . 1987b. Density-dependent dynamics of *Salicornia europaea* monocultures. *Ecology* 68:737-741.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125:1-15.
- Ferson, S., P. Downey, P. Klerks, M. Weissberg, I. Kroot, S. Stewart, G. Jacquez, J. Ssemakula, R. Malenky, and K. Anderson. 1986. Competing reviews, or why do Connell and Schoener disagree? *American Naturalist* 127:571-576.
- Fisher, R. A. 1932. Statistical methods for research workers. 4th ed. Oliver & Boyd, London.
- Fletcher, W. J. 1987. Interactions among subtidal Australian sea urchins, gastropods, and algae: effects of experimental removals. *Ecological Monographs* 57:89-109.
- Fletcher, W. J., and A. J. Underwood. 1987. Interspecific competition among subtidal limpets: effect of substratum heterogeneity. *Ecology* 68:387-400.
- Fowler, N. L. 1986. Density-dependent population regulation in a Texas grassland. *Ecology* 67:545-554.
- . 1990a. Disorderliness in plant communities: comparisons, causes and consequences. Pages



- 291–306 in J. B. Grace and D. Tilman, eds. *Perspectives on plant competition*. Academic Press, New York.
- . 1990b. The 10 most common statistical errors. *Bulletin of the Ecological Society of America* 71:161–164.
- Glass, G. V. 1976. Primary, secondary, and meta-analysis of research. *Educational Researcher* 5:3–8.
- . 1978. Integrating findings: the meta-analysis of research. Pages 351–379 in L. S. Shulman, ed. *Review of research in education*. Vol. 5. Peacock, Itasca, Ill.
- Glass, G. V., B. McGaw, and M. L. Smith. 1981. *Meta-analysis in social research*. Sage, Beverly Hills, Calif.
- Goldberg, D. E. 1990. Components of resource competition in plant communities. Pages 27–49 in J. B. Grace and D. Tilman, eds. *Perspectives on plant competition*. Academic Press, New York.
- Goldberg, D. E., and A. M. Barton. 1992. Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. *American Naturalist* 139:771–801.
- Goldberg, D. E., and P. A. Werner. 1983. Equivalence of competitors in plant communities: a null hypothesis and a field experimental approach. *American Journal of Botany* 70:1098–1104.
- Grace, J. B. 1990. On the relationship between plant traits and competitive ability. Pages 51–66 in J. B. Grace and D. Tilman, eds. *Perspectives on plant competition*. Academic Press, New York.
- Grace, J. B., and D. Tilman, eds. 1990. *Perspectives on plant competition*. Academic Press, New York.
- Grime, J. P. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *American Naturalist* 111:1169–1194.
- . 1987. Dominant and subordinate components of plant communities: implications for succession, stability and diversity. Pages 413–428 in A. J. Gray, M. J. Crawley, and P. J. Edwards, eds. *Colonization, succession and stability*. Blackwell Scientific, Oxford.
- Gurevitch, J. 1986. Competition and the local distribution of the grass *Stipa neomexicana*. *Ecology* 67:46–57.
- Gurevitch, J., and S. T. Chester. 1986. Analysis of repeated measures experiments. *Ecology* 67:251–255.
- Gurevitch, J., and R. S. Unnasch. 1989. The effects of competition on plant community structure at two levels of soil resources. *Canadian Journal of Botany* 67:3470–3477.
- Hairston, N. G. 1989. *Ecological experiments: purpose, design and execution*. Cambridge University Press, Cambridge.
- Hairston, N. G., F. E. Smith, and L. B. Slobodkin. 1960. Community structure, population control, and competition. *American Naturalist* 94:421–425.
- Hedges, L. V. 1983. A random effects model for effect sizes. *Psychological Bulletin* 93:388–395.
- Hedges, L. V., and I. Olkin. 1980. Vote-counting methods in research synthesis. *Psychological Bulletin* 88:359–369.
- . 1985. *Statistical methods for meta-analysis*. Academic Press, Orlando, Fla.
- Heil, G. W., and M. Bruggink. 1987. Competition for nutrients between *Calluna vulgaris* (L.) Hull and *Molina caerulea* (L.) Moench. *Oecologia* (Berlin) 73:105–108.
- Hill, W. R., and A. W. Knight. 1987. Experimental analysis of the grazing interaction between a mayfly and stream algae. *Ecology* 68:1955–1965.
- Hunter, J. E., F. L. Schmidt, and G. B. Jackson. 1982. *Meta-analysis: cumulating research findings across studies*. Sage, Beverly Hills, Calif.
- Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211.
- Jara, H. F., and C. A. Moreno. 1984. Herbivory and structure in a midlittoral rocky community: a case in southern Chile. *Ecology* 65:28–38.
- Jernakoff, P. 1985. The effect of overgrowth by algae on the survival of the intertidal barnacle *Tessieropora rosea* Krauss. *Journal of Experimental Marine Biology and Ecology* 94:89–97.
- Johansson, L. 1987. Experimental evidence for interactive habitat segregation between roach (*Rutilus*

- rutilus*) and rudd (*Scardinius erythrophthalmus*) in a shallow eutrophic lake. *Oecologia* (Berlin) 73:21–27.
- Johnson, C. R., and K. H. Mann. 1988. Diversity, patterns of adaptation, and stability of Nova Scotian kelp beds. *Ecological Monographs* 58:129–154.
- Karban, R. 1989. Community organization of *Erigeron glaucus* folivores: effects of competition, predation, and host plant. *Ecology* 70:1028–1039.
- Kastendiek, J. 1983. Competitor-mediated coexistence: interactions among three species of benthic macroalgae. *Journal of Experimental Marine Biology and Ecology* 62:201–210.
- Lamberti, G. A., and V. H. Resh. 1987. Herbivory and intraspecific competition in a stream caddisfly population. *Oecologia* (Berlin) 73:75–81.
- Lee, T. D., and F. A. Bazzaz. 1982. Regulation of fruit and seed production in an annual legume, *Cassia fasciculata*. *Ecology* 63:1363–1373.
- Levitan, D. R. 1989. Density-dependent size regulation in *Diadema antillarum*: effects on fecundity and survivorship. *Ecology* 70:1414–1424.
- Light, R. J., and D. B. Pillemer. 1984. Summing up: the science of reviewing research. Harvard University Press, Cambridge, Mass.
- Lively, C. M., and P. T. Raimondi. 1987. Dessication, predation, and mussel-barnacle interactions in the northern Gulf of California. *Oecologia* (Berlin) 74:304–309.
- Mann, C. 1990. Meta-analysis in the breech. *Science* (Washington, D.C.) 249:476–480.
- McConaughay, K. D. M., and F. A. Bazzaz. 1987. The relationship between gap size and performance of several colonizing annuals. *Ecology* 68:411–416.
- McGraw, J. B., and F. S. Chapin. 1989. Competitive ability and adaptation to fertile and infertile soils in two *Eriophorum* species. *Ecology* 70:736–749.
- Menge, B. A., and J. P. Sutherland. 1976. Species diversity gradients: synthesis of the roles of predation, competition and temporal heterogeneity. *American Naturalist* 110:351–369.
- Morin, P. J. 1986. Interactions between intraspecific competition and predation in an amphibian predator-prey system. *Ecology* 67:713–720.
- Morin, P. J., S. P. Lawler, and E. A. Johnson. 1988. Competition between aquatic insects and vertebrates: interaction strength and higher order interactions. *Ecology* 69:1401–1409.
- Morrisey, D. J. 1987. Effect of population density and presence of a potential competitor on the growth rate of the mud snail *Hydrobia ulvae* (Pennant). *Journal of Experimental Marine Biology and Ecology* 108:275–295.
- Morrow, L. L. 1990. Field experiments on competition: a meta-analysis. Master's thesis. State University of New York at Stony Brook, Stony Brook.
- Newman, E. I. 1973. Competition and diversity in herbaceous vegetation. *Nature* (London) 244:310–311.
- Oksanen, L. 1988. Ecosystem organization: mutualism and cybernetics or plain Darwinian struggle for existence? *American Naturalist* 131:424–444.
- Ortega, S. 1985. Competitive interactions among tropical intertidal limpets. *Journal of Experimental Marine Biology and Ecology* 90:11–25.
- Pearson, K. 1933. On a method of determining whether a sample of size  $n$  supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika* 25:379–410.
- Persson, L. 1987. Effects of habitat and season on competitive interactions between roach (*Rutilus rutilus*) and perch (*Perca fluviatilis*). *Oecologia* (Berlin) 73:170–177.
- Petratis, P. S. 1989. Effects of theperiwinkle *Littorina littorea* (L.) and of intraspecific competition on growth and survivorship of the limpet *Notoacmea testudinalis* (Muller). *Journal of Experimental Marine Biology and Ecology* 125:99–115.
- Pierce, C. L., P. H. Crowley, and D. M. Johnson. 1985. Behavior and ecological interactions of larval odonata. *Ecology* 66:1504–1512.
- Polis, G. A., and S. J. McCormick. 1986. Scorpions, spiders and solpugids: predation and competition among distantly related taxa. *Oecologia* (Berlin) 71:111–116.
- . 1987. Intraguild predation and competition among desert scorpions. *Ecology* 68:332–343.
- Quinn, G. P., and N. R. Ryan. 1989. Competitive interactions between two species of intertidal

- herbivorous gastropods from Victoria, Australia. *Journal of Experimental Marine Biology and Ecology* 125:1–12.
- Rosenthal, R. 1963. On the social psychology of the psychological experiment: the experimenter's hypothesis as unintended determinant of experimental results. *American Scientist* 51: 268–283.
- . 1976. *Experimenter effects in behavioral research*. Irvington, New York.
- . 1984. *Meta-analytic procedures for social research*. Sage, Beverly Hills, Calif.
- Schoener, T. W. 1974. Resource partitioning in ecological communities. *Science* (Washington, D.C.) 185:27–39.
- . 1983. Field experiments on interspecific competition. *American Naturalist* 122:240–285.
- . 1985. Some comments on Connell's and my reviews of field experiments on interspecific competition. *American Naturalist* 125:730–740.
- Sih, A., P. Crowley, M. McPeck, J. Petranka, and K. Strohmeier. 1985. Predation, competition, and prey communities: a review of field experiments. *Annual Review of Ecology and Systematics* 16:269–311.
- Slice, D. E. 1990. DS-DIGIT: basic digitizing software. Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook.
- Slobodkin, L. B., F. E. Smith, and N. G. Hairston. 1967. Regulation in terrestrial ecosystems, and the implied balance of nature. *American Naturalist* 101:109–124.
- Sokal, R. R., and F. J. Rohlf. 1981. *Biometry*. W. H. Freeman, New York.
- Spence, J. R. 1986. Relative impacts of mortality factors in field populations of the waterstrider *Gerris buenoi* Kirkaldy (Heteroptera: Gerridae). *Oecologia* (Berlin) 70:68–76.
- Sutherland, J. P., and S. Ortega. 1986. Competition conditional on recruitment and temporary escape from predators on a tropical rocky shore. *Journal of Experimental Marine Biology and Ecology* 95:155–166.
- Tilman, D. 1982. *Resource competition and community structure*. Princeton University Press, Princeton, N.J.
- . 1988. *Dynamics and structure of plant communities*. Princeton University Press, Princeton, N.J.
- Tippett, L. H. C. 1931. *The method of statistics*. Williams & Norgate, London.
- Titus, J. E., and M. D. Stephens. 1983. Neighbor influences and seasonal growth patterns for *Vallisneria americana* in a mesotrophic lake. *Oecologia* (Berlin) 56:23–29.
- Turner, T., and J. Lucas. 1985. Differences and similarities in the community roles of three rocky intertidal surfgrasses. *Journal of Experimental Marine Biology and Ecology* 89:175–189.
- Underwood, A. J. 1984. Vertical and seasonal patterns in competition for microalgae between intertidal gastropods. *Oecologia* (Berlin) 64:211–222.
- Wilbur, H. W., and R. A. Alford. 1985. Priority effect in experimental pond communities: response of *Hyla* to *Bufo* and *Rana*. *Ecology* 66:1106–1114.
- Wilson, S. D., and D. Tilman. 1991. Components of plant competition along an experimental gradient of nitrogen availability. *Ecology* 72:1050–1065.
- Wolf, F. M. 1986. *Meta-analysis: Quantitative methods for research synthesis*. Sage, Beverly Hills, Calif.

Associate Editor: Thomas R. Meagher