

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/250076062>

The Importance of Data-Selection Criteria: Meta-Analyses of Stream Predation Experiments

Article in *Ecology* · June 1999

DOI: 10.1890/0012-9658(1999)080[1132:TIODSC]2.0.CO;2

CITATIONS

159

READS

209

3 authors, including:



[Göran Englund](#)

Umeå University

80 PUBLICATIONS 3,270 CITATIONS

[SEE PROFILE](#)



[Scott Cooper](#)

University of California, Santa Barbara

77 PUBLICATIONS 4,669 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Climate change induced regime shifts in Northern lake ecosystems [View project](#)



Population dynamics and epidemic dynamics [View project](#)

All content following this page was uploaded by [Göran Englund](#) on 20 May 2017.

The user has requested enhancement of the downloaded file.



The Importance of Data-Selection Criteria: Meta-Analyses of Stream Predation Experiments

Goran Englund; Orlando Sarnelle; Scott D. Cooper

Ecology, Vol. 80, No. 4. (Jun., 1999), pp. 1132-1141.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9658%28199906%2980%3A4%3C1132%3ATIODCM%3E2.0.CO%3B2-N>

Ecology is currently published by The Ecological Society of America.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/esa.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

THE IMPORTANCE OF DATA-SELECTION CRITERIA: META-ANALYSES OF STREAM PREDATION EXPERIMENTS

GÖRAN ENGLUND,^{1,4} ORLANDO SARNELLE,^{2,5} AND SCOTT D. COOPER³

¹ Department of Animal Ecology, Umeå University, 901 87 Umeå, Sweden

² National Center for Ecological Analysis and Synthesis, University of California,
Santa Barbara, California 93101 USA

³ Department of Ecology, Evolution and Marine Biology, Santa Barbara, California 93106 USA

Abstract. The value of meta-analysis in ecology hinges on the reproducibility of patterns generated by quantitative synthesis. Meta-analysts will vary in the criteria they use to screen studies and select data within studies, even when addressing exactly the same question. We summarize some of the many decisions that an ecologist must make in deciding what data to include in a synthesis. We then show, using multiple meta-analyses taken from the same literature on stream predation experiments, that meta-analytic conclusions can be colored by selection criteria that are not specifically a function of the relevance of the data. As a consequence, we recommend that meta-analysts perform several meta-analyses using different selection criteria to examine the robustness of reported findings. We also advise ecological meta-analysts to minimize use of selection criteria that are based on judgments of study quality when extracting data from the literature, because of the potential for unconscious bias. The influence of quality criteria on patterns in the data set can then be examined empirically. Our comparisons of mean effect size, for studies included vs. excluded on the basis of “quality” criteria, provided no evidence that rejected studies were aberrant or more variable than “acceptable” studies. One result of excluding such studies was a loss of statistical power. We urge ecologists to be more explicit about how data are selected for a meta-analysis, to examine the robustness of the patterns they report, and to conduct meta-analyses to describe as well as to infer.

Key words: data selection criteria; meta-analysis; non-independence; predation experiments; predator effects in streams; stream predation experiments.

INTRODUCTION

Meta-analysis holds considerable promise for the development of quantitative generalizations in ecology based on the combined evidence of experimental studies (Gurevitch et al. 1992, Sarnelle 1992, Osenberg et al. 1997). While it should be expected that meta-analytic generalizations may change over time as the experimental literature grows, it is less appreciated that multiple meta-analyses asking the same question and performed within a short span of time (so that the available literature has changed little) may produce conflicting conclusions. This problem has fueled criticism of meta-analysis in other fields (Slavin 1984, Abrami et al. 1988). A familiar example in ecology is the lack of resolution provided by two simultaneous and conflicting vote-counting reviews dealing with the prevalence of competition (Connell 1983, Schoener 1983;

see Gurevitch and Hedges [1993] for discussion of problems with vote counting).

Many factors may lead to conflicting conclusions among contemporaneous meta-analyses having similar objectives. An obvious factor is the nature of the literature search, which can lead to large differences in the set of studies included in a meta-analysis (for example, compare the study lists of Connell [1983] and Schoener [1983]). To alleviate this problem, ecological meta-analysts are encouraged to consult widely available guidelines for thorough literature searching (Cooper 1984, Cooper and Hedges 1994). A second potential source of variability among meta-analyses arises from subtle differences in the questions posed. For example, Connell (1983) and Schoener (1983) ostensibly asked the same general question about competition (how frequently is it detected in field experiments?), but the metrics they used to quantify this frequency were subtly different. Just as in primary research, explicit formulation of the question is crucial in meta-analysis, in particular because of the close relationship between the question posed and the choice of effect-size metric, an issue discussed at length elsewhere in this special feature (Goldberg et al. 1999, Osenberg et al. 1999).

In this paper we focus on a third potential source of variability among meta-analyses with the same objec-

Manuscript received 11 November 1997; revised 2 July 1998; accepted 8 July 1998; final version received 17 September 1998. For reprints of this Special Feature, see footnote 1, p. 1103.

⁴ E-mail: goran.englund@animecol.umu.se

⁵ Present address: Department of Fisheries and Wildlife, Michigan State University, East Lansing, Michigan 48824-1222 USA.

tive, namely the way in which studies, and, in particular, data within studies, are selected for inclusion in the analysis. Differences in data-selection criteria appear to underlie differences in meta-analytic conclusions in other fields, to the extent that opposite conclusions from the same literature base are possible (Abrami et al. 1988). In ecology, data-selection criteria are likely to have a major impact on meta-analytic conclusions given the extremely heterogeneous nature of ecological systems and methodologies. This heterogeneity forces the meta-analyst to make a large number of decisions regarding what studies to include and how data should be extracted from those studies. When the available studies are few, such decisions can greatly alter the nature of the data set.

The importance of data-selection criteria in ecological meta-analysis is suggested by the contrasting conclusions reached by three published meta-analyses of predator effects in streams (Cooper et al. 1990, Wooster 1994, Dahl and Greenberg 1996). For example, Cooper et al. (1990; 14 studies) found a highly significant decrease in the magnitude of predator effects on prey as the mesh size of predator enclosures/exclosures increased. Mesh size was assumed to be correlated with the rate of prey exchange (immigration, emigration) between cages and the outside world (cages are porous to prey to varying degrees). Cooper et al.'s (1990) results suggested that higher rates of prey exchange in cages with larger mesh may mask the effects of predator consumption. In contrast, the meta-analyses of Wooster (1994; 20 studies) and Dahl and Greenberg (1996; 10 studies) did not detect a relationship between predator effect and mesh size. The relationship between per capita prey exchange (a function of mesh size plus other design and environmental variables) and the magnitude of predator impact is fundamental to the understanding of predator impacts in open systems (Cooper et al. 1990, Sih and Wooster 1994, Englund and Olsson 1996, Englund 1997, Nisbet et al. 1997), so resolution of these discrepancies is important. As will be the case with most contemporaneous meta-analyses, data were selected differently in each of the above meta-analyses.

In this paper, we first discuss the types of decisions that must be made when ecological data are selected for a meta-analysis. We then examine three patterns from the above-listed meta-analyses using a newly constructed data set, and ask whether these patterns are robust to changes in selection criteria. We show that some of the conclusions reached in published analyses are not robust, and suggest strategies to minimize what we consider to be unproductive controversy. In particular, we recommend that meta-analysts perform multiple analyses using different selection criteria and examine the robustness of their results.

TYPES OF SELECTION CRITERIA

The decisions that have to be made when selecting data for meta-analysis can be loosely categorized into

three types relating to the relevance, quality, and independence of data. Some of these decisions are driven by the question posed and so are concerned with the relevance of the available data. It is critical that the question be explicit, so that the relevance of the selected data is clear. Reference to a model can be helpful here, to ensure that the selected data are appropriate for both the question and the chosen metric of effect size (Osenberg et al. 1999). Some relevance criteria are obvious. For example, Dahl and Greenberg (1996) asked whether two types of predatory fish differ in their effects on benthic prey, so they rejected studies that examined the combined effects of both types of predators. Other criteria are less obvious, but no less critical for interpretation. For example, all three meta-analyses referred to above purported to examine the magnitude of predator effects on benthic prey in streams, yet only one (Cooper et al. 1990) provided any definition of "prey." Some of the responses reported in the reviewed studies were for species that, rather than being reduced by predator consumption, were probably affected indirectly and even positively by the manipulated predator. Clearly, an explicit explanation of what constitutes prey is more vital to the validity of a meta-analysis than a description of the physical means by which data are extracted from published figures, a detail that is often included in the methods of meta-analysis papers. Lastly, it sometimes may be ambiguous whether particular data are relevant to the question. For example, should studies conducted in isolated stream pools with little or no flow and a more or less lentic fauna be included in a meta-analysis of stream studies? Clearly, ecologists will differ in their judgment of whether such data should be included or not.

A second set of selection criteria can be loosely classified as those relating to judgments of study quality. Quality features that have been used to reject individual studies from stream predation meta-analyses are: confounded controls, lack of replication, lack of within-treatment variance estimates, and incomplete predator removal from predator areas (Wooster 1994, Dahl and Greenberg 1996). It has been argued cogently by non-ecological meta-analysts that the use of selection criteria based on quality features of studies should probably be avoided as much as possible (Hunter and Schmidt 1990). There are two fundamental reasons to be wary of quality criteria. First, reviewers are likely to vary considerably in their assessments of study quality (Cooper 1984). Whereas most ecologists would agree that a predator manipulation that lacked a concurrent control should not be included in a meta-analysis, many quality decisions are less clear (lack of replication, confounded controls) and their varied application will result in markedly different subsets of the available data being used in meta-analyses addressing the same question. If the results of a meta-analysis critically depend on such decisions, confidence in meta-analytic generalizations will be diminished.

A more serious problem with the use of quality criteria is the potential for bias to enter into the analysis (Hunter and Schmidt 1990). It has been shown that a reviewer is more likely to find methodological fault with studies that report results that conflict with the reviewer's predispositions (Mahoney 1977). The establishment of rigid quality criteria before the literature is searched does not guarantee that bias will be avoided because most meta-analysts in ecology will probably be very familiar with the literature before extracting any data and will have a biased expectation of the outcome of a quantitative review.

A third set of selection criteria concerns the definition of an independent observation in meta-analysis. The restriction that observations must be independent largely stems from the desire to apply inferential statistics to meta-analysis data (Gurevitch and Hedges 1993, Gurevitch and Hedges 1999). However, even if this statistical requirement is relaxed and the meta-analysis regarded as primarily a descriptive procedure, lack of independence of observations can produce biased patterns if a large proportion of the observations originate from a small subset of published studies or authors. In practice, the handling of multiple observations within a study is often not straightforward, and the meta-analyst is faced with additional subjective and arbitrary decisions. A rigorous statistical determination of which observations are independent within a particular study may not be possible, so ecologists must often use their intuition in extracting observations from studies with multiple observations. For example, Wooster (1994) only included experiments from the same study if they involved different predator species or were conducted in different seasons. One strategy for dealing with problems of non-independence is to identify groups of observations that may not be independent (data from the same study, researcher, geographic area, taxon, etc.), and then examine if these observations bias the overall estimate. An example is provided below (see *The influence of selection criteria: Results*). In all cases, it is critical that the description of data selection is explicit.

THE INFLUENCE OF SELECTION CRITERIA: META-ANALYSES OF STREAM PREDATION EXPERIMENTS

The three published meta-analyses on stream predation experiments differ in several respects in addition to how studies and data were selected, most notably in the metric of effect size, so we compiled a new data set with which to examine how meta-analytic patterns are influenced by changes in selection criteria. We organized selection criteria into four groups, corresponding to our attempts to reconstruct the criteria used by Wooster (1994, the WOOSTER subset) and Dahl and Greenberg (1996, the DAHL subset), plus two additional subsets. We then compared patterns generated by the complete data set (COMPLETE, the least re-

strictive set of criteria) and each of the four subsets of data. Three patterns explicitly addressed in previous meta-analyses were examined: the effects of invertebrate vs. vertebrate predators (Wooster 1994), the effects of drift-feeding vs. benthic-feeding fish (Dahl and Greenberg 1996), and the relationship between mesh size and predator impact (Cooper et al. 1990). In cases where selection criteria appeared to influence inferences about these patterns, we compared patterns generated by studies included or excluded by a particular set of criteria.

Methods

The COMPLETE data set comprised studies published between 1980 and 1996 on the effects of fish or invertebrate predators on benthic invertebrates in which at least one treatment with the predator and one without the predator (control) were maintained in field enclosures/exclosures or stream sections. Partial predator removals (Allan 1982, Reice 1991) were thus excluded. We conducted our own literature searches, in addition to using the study lists of Cooper et al. (1990), Wooster (1994), and Dahl and Greenberg (1996), so it is likely that we have surveyed nearly all relevant papers published in English during this period. For the purposes of comparison with the Wooster (1994) and Dahl and Greenberg (1996) studies, we examined responses of total invertebrate density (note that we refrain from using the term "prey" here because this term may have been used too loosely in previous meta-analyses). In addition, we examined responses of two categories of benthic invertebrates: "epibenthic taxa" and chironomids (Order Diptera, Family Chironomidae). Epibenthic taxa overwhelmingly consisted of mayflies (Order Ephemeroptera) and stoneflies (Order Plecoptera), mobile taxa that dominate the epibenthic fauna of most stony streams and are very common components of the diet of both fish and invertebrate predators. The remainder of the epibenthic category consisted of minor contributions from the following taxa (in decreasing order of prevalence in the data set): free-living caddisfly larvae (Order Trichoptera, Hydrobiosidae, *Rhyacophila*), amphipods, isopods, decapods, cladocerans, copepods, gastropods, and very low densities of larger invertebrate predators (Orders Odonata, Megaloptera, and Coleoptera). Of these minor taxa, the responses of large taxa (decapods, Odonata, etc.) were only reported in studies of fish predation. Chironomids were examined separately because they are small and cryptic, often numerically dominant, and can be one or more trophic levels removed from some of the predators typically manipulated in stream studies. A complete list of included studies and the data extracted from each study is given in the Appendix.

The magnitude of predator effect was quantified as: PI (predator impact) = $\ln(N_p/N_0)$, where N_p is mean prey densities in cages containing the predator and N_0 is mean prey density in cages lacking the predator

(Cooper et al. 1990). Note that a more negative value of PI indicates a larger negative effect of the predator. We chose this index based on observations of prey dynamics in studies that sampled cages over time (G. Englund, O. Sarnelle, and S. D. Cooper, *unpublished manuscript*), which suggested that prey equilibrate relatively quickly in most experiments. PI in this case simply represents a ratio of equilibrium densities (and as such is time invariant), with the log transformation providing a symmetrical scale and improved statistical properties (see Hedges et al. 1999). To maximize the match between data and metric, we used data from the final sampling for experiments that were sampled over time. For studies reporting the results of multiple manipulations, experiments were considered independent observations if factorial manipulations were applied, or if separate experiments involved different predators or were conducted in different streams, seasons, or years. The one exception was the multi-year, multi-season fish manipulations of Feltmate and Williams (1991). In this case, we followed the authors' lead by averaging across years. For studies in which multiple predator densities were maintained, we averaged responses across predator densities.

In addition to the criteria used to compile the COMPLETE data set, the WOOSTER subset excluded studies that compared predator exclosures with open cages (Dudgeon 1993), studies reporting responses in biomass (milligrams per square meter) rather than density (individuals per square meter) units, experiments conducted in arenas not isolated by mesh barriers (e.g., in naturally isolated stream sections or pools, Wiseman et al. 1993) and studies without estimates of within-treatment variance for individual taxa. The latter criterion was necessitated by Wooster's use of a weighted d as the metric of effect size. The DAHL subset excluded studies that compared exclosures with open cages, experiments conducted in arenas not isolated by mesh barriers, studies that lacked estimates of within-treatment variance for individual taxa, studies that reported responses of only one prey taxon (Feltmate and Williams 1989), and experiments conducted in artificial stream-side channels (Schlosser and Ebel 1989). Artificial channels placed on the natural streambed were included, following Dahl and Greenberg (1996). The VARIANCE subset only excluded studies (from the COMPLETE data set) that lacked estimates of within-treatment variance for individual taxa, whereas the MINIMAL subset excluded all studies excluded by DAHL plus studies reporting responses in biomass units.

Ideally, statistical tests on meta-analytic data should be performed on effect-size estimates that are weighted by within-study variance (Gurevitch and Hedges 1993). This generally was not possible in our analysis primarily because very few papers reported variances for the epibenthic category (one of our primary interests). As a consequence, we usually calculated effect size for

this category by summing the densities of several taxa. Instead of making untestable assumptions in calculating variances for these sums (to enable a limited, weighted analysis), we performed less restrictive, but unweighted analyses. Consequently, we caution that our P values are not exact and our tests are less powerful than if we were able to weight effect sizes (Gurevitch and Hedges 1999). To assess how unreliable our unweighted estimates and tests might be, we applied mixed-model meta-analysis (Rosenberg et al. 1997) to effect sizes (PI) for total invertebrate density using a subset of experiments for which variances of total invertebrate density were reported ($N = 44$). These analyses compared means for different predator types after applying the various sets of selection criteria. Weighted means from these analyses were highly correlated with unweighted means for the same data and selection criteria ($r = 0.91$, $N = 14$ sets of experiments), and in no case did the two analyses differ in their assessment of the null hypothesis. These analyses also suggested a positive relationship between the variance of PI and experiment duration ($r_s = 0.28$, $P < 0.05$, $N = 43$), after one experiment with unusually low variance was deleted), so an analysis weighted by the inverse of the variance of PI (Gurevitch and Hedges 1993) would potentially weight shorter studies more heavily than longer ones. Given that our metric is derived for equilibrium densities, we deemed weighting by the inverse of the variance to be inappropriate in this instance. In any case, our objective is to compare results across subsets of the data, so it is primarily the relative level of significance, which is probably fairly robust to the tests employed, that is of consequence in this paper.

Results

The COMPLETE data set contained a total of 41 published studies and 86 experiments, after we excluded a single large outlier (PI = -2.79 for a drift-feeding fish, Feltmate and Williams 1991). The PI for this outlier was more than five standard deviations greater in absolute magnitude than the next greatest effect size for drift-feeding fish (-0.99 , Table 1), leading us to exclude this datum from all analyses. This PI was also several times larger than other estimates for this predator species made by the same authors. The data sets for two of the predator types (invertebrates and drift-feeding fish) were dominated by particular authors or taxa. Most observations (66–75%, depending on response taxon) for invertebrate predators came from studies authored by two researchers, Peckarsky (Peckarsky 1985, 1991, Peckarsky and Dodson 1980, Peckarsky et al. 1990) and Lancaster (Lancaster 1990, Lancaster et al. 1990, 1991). However, mean PI's for experiments by these authors were similar to the remainder of the data set (ANOVA, $P > 0.88$). In addition, most observations for invertebrate predators (58–64%) were for species in the insect order Plecoptera (stoneflies). The data set for drift-feeding fish was

TABLE 1. Predator impact (PI) summary statistics for the COMPLETE data set, by predator type and response category.

Predator type	Response category	N	Mean	Variance	Min	Max
Fish	Total	46	-0.25	0.20	-1.38	0.76
	Epibenthic taxa	37	-0.30	0.17	-1.47	0.40
	Chironomids	24	-0.11	0.59	-1.49	2.00
Invertebrates	Total	38	-0.30	0.12	-1.36	0.51
	Epibenthic taxa	36	-0.49	0.19	-1.71	0.15
	Chironomids	24	-0.32	0.15	-1.36	0.34
Benthic-feeding fish	Total	23	-0.43	0.22	-1.38	0.34
	Epibenthic taxa	14	-0.40	0.263	-1.47	0.28
	Chironomids	13	-0.54	0.36	-1.49	0.46
Drift-feeding fish	Total	22	-0.08	0.13	-0.57	0.76
	Epibenthic taxa	22	-0.23	0.12	-0.99	0.40
	Chironomids	10	0.43	0.43	-0.25	2.00

Notes: Total = total invertebrates; N, = no. of experiments (out of 86 experiments in 41 published studies); Min and Max are the smallest and largest PI, respectively, in each category.

dominated by observations from the Family Salmonidae (73–90%), so any conclusion about this category of predator is essentially a statement about trout. It may also be noted that the entire data set was strongly dominated by studies performed in small temperate streams.

Effects of invertebrate vs. vertebrate predators.—Wooster (1994) found the effects of invertebrate predators to be significantly greater than the effects of vertebrate predators on invertebrate density (all taxa combined). We were unable to reproduce this pattern for any of the data sets we constructed (Table 2). Our analyses did not indicate any consistent trend for invertebrate effects on total invertebrate density to be larger than those of fish (all but one vertebrate observation in Wooster's analysis was from fish). We were also unable to reproduce Wooster's pattern when we limited the analysis to studies in our data set that were also used by Wooster (mean PI's: -0.213 for invertebrates, -0.262 for fish, not different at $P > 0.60$). As a consequence, we suspect that our inability to detect any

trace of a pattern is not related to the growth of the literature since Wooster's analysis. One possible explanation lies in the effect-size metrics used. The metric employed by both Wooster (1994) and Dahl and Greenberg (1996) is based on an arithmetic difference divided by a pooled standard deviation across the two treatments (d). Effects of invertebrates may appear larger when using d if such manipulations tend to produce responses with less within-treatment variation, and so smaller pooled standard deviations. The latter would not be surprising given that invertebrates tend to be manipulated in smaller enclosures that probably can be replicated more precisely (Osenberg et al. 1997). The data support this line of reasoning. The average pooled standard deviation for fish manipulations ($SD = 6329$) was nearly three times larger than that for invertebrate manipulations ($SD = 2170$). Consequently, the conclusion that invertebrates have larger effects than vertebrates (fish) on total invertebrate density (Wooster 1994) may be a consequence of inherent confounding

TABLE 2. Comparison of the impact of fish vs. invertebrate predators for three invertebrate categories.

Response category	Selection subset	Mean PI (1 SE)		<i>t</i>	df	<i>P</i>
		Fish predators	Invertebrate predators			
Total invertebrates	COMPLETE	-0.25 (0.07)	-0.30 (0.06)	0.60	82	0.55
	VARIANCE	-0.30 (0.08)	-0.26 (0.06)	0.37	64	0.71
	WOOSTER	-0.28 (0.09)	-0.24 (0.07)	0.37	53	0.71
	DAHL	-0.26 (0.10)	-0.24 (0.07)	0.17	47	0.86
	MINIMAL	-0.26 (0.10)	-0.20 (0.09)	0.40	40	0.70
Epibenthic taxa	COMPLETE	-0.29 (0.07)	-0.49 (0.07)	1.96	71	0.05
	VARIANCE	-0.32 (0.08)	-0.51 (0.08)	1.58	56	0.12
	WOOSTER	-0.30 (0.08)	-0.57 (0.10)	2.04	45	0.05
	DAHL	-0.26 (0.10)	-0.44 (0.09)	1.36	39	0.18
	MINIMAL	-0.26 (0.10)	-0.50 (0.12)	1.57	32	0.12
Chironomids	COMPLETE	-0.112 (0.157)	-0.32 (0.08)	1.20	46	0.24
	VARIANCE	-0.320 (0.152)	-0.24 (0.09)	0.46	34	0.65
	WOOSTER	-0.267 (0.152)	-0.30 (0.14)	0.13	24	0.90
	DAHL	-0.261 (0.164)	-0.236 (0.14)	0.11	22	0.92
	MINIMAL	-0.261 (0.164)	-0.440 (0.46)	0.44	15	0.66

Notes: Results are presented for the full data set (COMPLETE) and for four subsets of data selected using four sets of criteria (see *The influence of selection criteria: Methods* for definitions of criteria sets). Mean predator impacts (PI) for the two predator types were compared using a t test; data are means with 1 SE in parentheses).

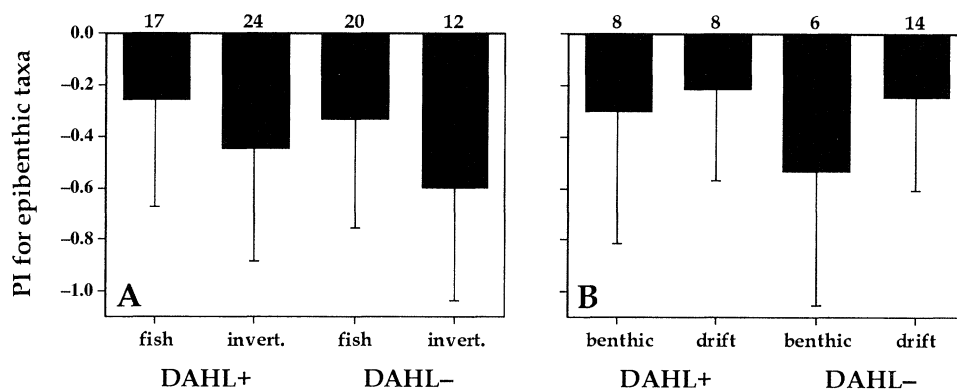


FIG. 1. Effect size of predation (PI, predator impact). Responses (means and 1 SD) of epibenthic taxa to the manipulation of different types of predators, using data that would be included (+) or excluded (-) by the DAHL set of selection criteria. Numbers at the top are the number of observations in each group. (A) Comparisons of fish vs. invertebrate predators. (B) Comparisons of benthic-feeding vs. drift-feeding fish.

associated with the use of *d*. Pooled standard deviations for benthic-feeding vs. drift-feeding fish were more similar ($SD = 5691$ for benthic, 7310 for drift-feeding fish).

In contrast to the lack of evidence for larger invertebrate effects on total invertebrate density, our analyses suggest that invertebrate predators do have larger effects than fish on the density of epibenthic taxa (Table 2). PI's for invertebrate predators on this group were consistently 1.5–2 times larger than fish PI's regardless of how the data were selected; however, the statistical significance of this difference was influenced by the selection criteria applied to the data set. Using the DAHL subset of criteria, we suspect most meta-analysts would conclude that there was no difference in the effects of the two predator types on epibenthic taxa ($P > 0.18$), a conclusion driven by the loss of power resulting from a very restrictive set of selection criteria. Studies included or excluded by the DAHL criteria were quite similar in mean effect size (Fig. 1A). Notably, application of the stringent set of DAHL criteria did not eliminate studies that were more variable in effect size (Fig. 1A). We observed no consistent differences between the effects of the two predator types on chironomids, but this result is related to the pooling of benthic-feeding and drift-feeding fish in this analysis, an issue that we consider next.

Effects of drift-feeding vs. benthic-feeding fish.—Dahl and Greenberg (1996) found a highly significant difference between the effects of benthic-feeding vs. drift-feeding fish on invertebrate density (all species combined) for their small data set ($N = 10$ experiments). In this case, our analyses support their conclusion, but here again, the statistical significance of the difference between fish types was strongly affected by the exclusion of studies (Table 3). The studies excluded by the DAHL criteria (in Fig. 1B: DAHL-) were no more variable than those included (Fig. 1B: DAHL+). The only noticeable difference between

these sets of studies is the somewhat larger mean PI for benthic-feeding fish shown by excluded studies (Fig. 1B). Analyzing epibenthic and chironomid responses separately revealed that the difference between fish-feeding types is largely driven by a difference in effects on chironomids. Benthic-feeding fish consistently showed much larger effects on chironomids than drift-feeding fish, although here the difference between predator types shrinks and the statistical significance of the difference wanes as the data set is truncated by selection criteria (Table 3). Larger effects of benthic-feeding fish on chironomids is reassuring given that chironomids are small, relatively sedentary prey that should be encountered less by visually oriented fish that feed on drift (trout). The least-restrictive data set (COMPLETE) also provides evidence that drift-feeding fish may actually affect chironomids positively (*t* test of whether mean PI is non-zero, $P < 0.10$, $N = 10$ experiments). More experiments are needed to confirm this pattern, which has important implications for understanding trophic interactions in streams.

Relationship between mesh size and predator impact.—Cooper et al. (1990) found a statistically significant relationship between predator impact and the mesh size of experimental cages using unweighted analysis and PI, while Wooster (1994) and Dahl and Greenberg (1996), using weighted analyses and *d*, did not. Given that the data suggest fish and invertebrates may have contrasting effects on different categories of prey (Table 2), we analyzed relationships between predator impact and mesh size separately by both predator type and prey category. Epibenthic taxa and chironomids might be expected to show different responses because mayflies and stoneflies (the dominant taxa in the epibenthic category) tend to be more mobile than chironomids. The results of these analyses show a strong influence of selection criteria. We found a negative relationship between PI for chironomids and mesh size, for invertebrate predators (predator effects

TABLE 3. Comparison of the impact of benthic-feeding vs. drift-feeding fish on three invertebrate categories.

Response category	Selection subset	Mean PI (1 SE)		<i>t</i>	df	<i>P</i>
		Benthic-feeding fish	Drift-feeding fish			
Total invertebrates	COMPLETE	-0.43 (0.10)	-0.08 (0.08)	2.72	43	0.01
	VARIANCE	-0.43 (0.11)	-0.10 (0.12)	1.97	30	0.06
	WOOSTER	-0.40 (0.12)	-0.13 (0.12)	1.54	26	0.13
	DAHL	-0.38 (0.13)	-0.08 (0.16)	1.36	21	0.19
	MINIMAL	-0.38 (0.13)	-0.08 (0.16)	1.36	21	0.19
Epibenthic taxa	COMPLETE	-0.40 (0.14)	-0.23 (0.08)	1.16	34	0.26
	VARIANCE	-0.42 (0.14)	-0.20 (0.09)	1.28	23	0.21
	WOOSTER	-0.38 (0.16)	-0.22 (0.10)	0.89	19	0.39
	DAHL	-0.30 (0.18)	-0.21 (0.13)	0.39	14	0.70
	MINIMAL	-0.30 (0.18)	-0.21 (0.13)	0.39	14	0.70
Chironomids	COMPLETE	-0.54 (0.17)	0.43 (0.21)	3.70	21	0.001
	VARIANCE	-0.54 (0.17)	0.31 (0.18)	2.66	21	0.02
	WOOSTER	-0.44 (0.17)	0.26 (0.24)	1.95	12	0.08
	DAHL	-0.45 (0.19)	0.26 (0.24)	1.87	11	0.09
	MINIMAL	-0.45 (0.19)	0.26 (0.24)	1.87	11	0.09

Notes: Results are presented for the full data set (COMPLETE) and for four subsets of the data selected using four sets of criteria (see *The influence of selection criteria: Methods* for definitions of criteria sets). Mean predator impacts (PI) for the two fish types were tested by *t* test; data are means with 1 SE in parentheses. Benthic-feeding genera were *Hesperoleucas*, *Hybognathus*, *Phoxinus*, *Rhinichthys*, *Semotilus* (Cyprinidae), *Culea*, *Gasterosteus* (Gasterosteidae), *Bryconamericus*, *Cheirodon* (Characidae), *Cottus* (Cottidae), *Lepomis* (Centrarchidae), *Etheostoma* (Percidae), and *Oreonectes* (Cobitidae). Drift-feeding genera were *Salmo*, *Salvelinus*, *Oncorhynchus* (Salmonidae), and *Galaxias* (Galaxiidae).

were more negative as mesh size increased) using the COMPLETE data set ($r_s = -0.59$, $P = 0.002$). This relationship for chironomids drove the negative relationship between PI for total invertebrates and mesh size (Table 4). These negative relationships completely disappeared when the DAHL criteria were applied ($P > 0.42$). In the case of the chironomid relationship, observations in the DAHL subset included a very small range of mesh sizes relative to the entire data set (Fig. 2A), which partly may explain the lack of pattern in the DAHL data. In contrast to the results for chironomids, the relationship for invertebrate predators between PI for epibenthic taxa and mesh size was not

significant for the COMPLETE data set, but when the data were restricted by the WOOSTER criteria, a marginally significant positive relationship appeared ($P < 0.10$, Table 4).

Potentially contrasting relationships between mesh size (a surrogate for the exchange rate of prey between cages and the outside world) and PI for chironomids (negative) vs. epibenthic taxa (positive?) (Table 4, invertebrate predators) are of great interest in light of aforementioned differences in the migration rates of these prey groups, and, more generally, the role that behavioral responses of prey play in driving the responses of prey density to predation in open systems

TABLE 4. Relationships between the magnitude of predator impact (PI) and the mesh size of experimental enclosures for two predator types (fish vs. invertebrates) and three invertebrate categories.

Response category	Selection subset	Fish			Invertebrates		
		r_s	<i>N</i>	<i>P</i>	r_s	<i>N</i>	<i>P</i>
Total invertebrates	COMPLETE	0.23	42	0.14	-0.39	38	0.02
	VARIANCE	0.27	33	0.13	-0.27	33	0.13
	WOOSTER	0.23	29	0.24	-0.45	26	0.02
	DAHL	0.18	24	0.41	-0.17	25	0.42
	MINIMAL	0.18	24	0.41	-0.44	18	0.07
Epibenthic taxa	COMPLETE	0.18	33	0.31	0.17	36	0.33
	VARIANCE	0.09	26	0.67	0.20	32	0.27
	WOOSTER	0.02	22	0.91	0.34	25	0.10
	DAHL	-0.17	17	0.51	0.21	24	0.33
	MINIMAL	-0.17	17	0.51	0.32	17	0.21
Chironomids	COMPLETE	0.29	20	0.22	-0.59	24	0.002
	VARIANCE	0.38	18	0.12	-0.48	18	0.04
	WOOSTER	0.40	15	0.14	-0.51	11	0.11
	DAHL	0.42	14	0.13	0.11	10	0.76
	MINIMAL	0.42	14	0.13	0.87	3	0.33

Notes: Results are presented for the full data set (COMPLETE) and for four subsets of the data selected using four sets of criteria (see *The influence of selection criteria: Methods* for definitions of criteria sets). r_s = Spearman's correlation coefficient. *P* values are for two-tailed null-hypothesis tests.

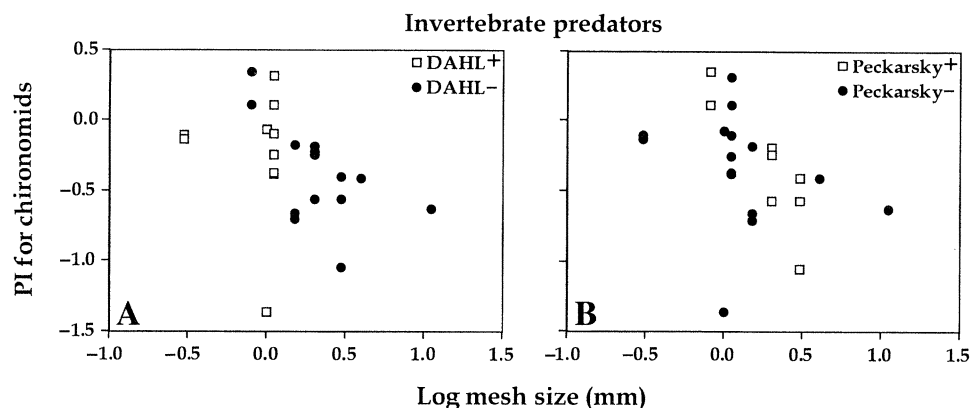


FIG. 2. Relationships between effect sizes (PI) measuring the response of chironomids to the manipulation of invertebrate predators and the mesh size of experimental enclosures. (A) Data coded with respect to whether they would be included (+) or excluded (-) based on the DAHL set of selection criteria. (B) Data coded with respect to whether they originated from a study by Peckarsky (+) or not (-).

(Sih and Wooster 1994, Englund 1997, Nisbet et al. 1997, G. Englund, O. Sarnelle, and S. D. Cooper, *unpublished manuscript*). There are, however, at least two caveats that should be considered with respect to this analysis. The first concerns the fact that much of the data for the chironomid relationship comes from experiments conducted by Peckarsky (9 of 24 observations, PI vs. mesh size: $r_s = -0.87$ for these 9 observations), which were conducted following a relatively uniform protocol. Fortunately, the data from Peckarsky and from other sources both show a negative relationship between predator impact and mesh size (Fig. 2B). The second caveat concerns the degree to which variation in mesh size provides an adequate representation of variation in per capita exchange rate of prey into and out of cages. In a subsequent paper, we suggest that mesh size by itself may not be a good index of per capita exchange because it is confounded with other variables (G. Englund, O. Sarnelle, and S. D. Cooper, *unpublished manuscript*). Our objective here was simply to show that selection criteria, which often may be tangential to the question posed, can have a major impact on the conclusions of an ecological meta-analysis.

CONCLUSIONS

If meta-analysis is to become a useful and trustworthy tool for generalizing about the results of ecological experiments, it is critical that meta-analytic patterns be convincing. If patterns appear and disappear depending on how studies from the same literature are selected for synthesis, a morass will result. A clear illustration of the potential problem with data selection is shown by the relationships between PI (predator impact) for chironomids and mesh size (Table 4, invertebrate predators). Conclusions about this pattern were strongly affected by selection decisions that were not always a function of the relevance of the data. The DAHL criteria included such factors as whether a study reported variances or was conducted in an artificial

channel that was placed beside, rather than within, the stream. It is difficult to see what specific relevance these criteria have to the question of whether different types of predators have different effects on stream invertebrates. Excluding data based on irrelevant criteria, without examining the effects of such exclusion, will only lead to unproductive and unnecessary controversy. At the same time, it must be emphasized that certain selection criteria are essential to construction of a meaningful meta-analysis. An obvious example is the selection of studies and data with respect to time-scale considerations (Downing et al. 1999). In this case, some data are clearly appropriate for the question and some data are not.

If meta-analytic patterns are affected by arbitrary and tangential decisions about data selection as we have shown, it makes sense to examine the influence of these decisions on the results. Of particular concern are selection criteria based on judgments of quality, in particular because of the potential for unconscious bias (Hunter and Schmidt 1990). As an alternative to rejecting studies in a pseudo-a priori way (we assume that meta-analysts generally have an idea of what studies contain before selecting data), we advocate that the effect of quality and other selection criteria on meta-analytic conclusions be treated as an empirical question (Glass et al. 1981). Relevant studies should be initially included with minimal prejudice as to methodological features. The data set can then be examined for differences in effect size among studies coded for various methodological attributes, as we have illustrated (Fig. 2). There was no evidence from the experiments we surveyed that rejected studies were aberrant or more variable than included ones.

The disadvantage of taking an empirical approach to study quality and other selection criteria is that uncertainty is increased, because the statistical and biological significance of patterns may fluctuate depending

on which data are included. We suggest that it is better to be thorough and uncertain than to be pseudo-rigorous and misled. In addition, uncertainty may be less of a problem if ecological meta-analysis is viewed primarily as a descriptive tool used for the generation of new hypotheses (Glass et al. 1981, Hunter and Schmidt 1990), rather than as a rigorous inferential tool. Many important ecological questions can only be addressed by non-exact meta-analysis due to the nature of the available data (Downing et al. 1999). In our case, construction of the epibenthic category precluded weighted statistics but provided ecological insights and new hypotheses for testing. We feel it would be foolish to avoid summarizing experimental data because of standards about statistical rigor that are too strict. Where more rigor is needed is in the reporting of meta-analytic protocols. Given the potential problems that may arise from too-liberal use of a priori selection criteria, we recommend that ecological meta-analysts be more explicit about how and why data were selected and analyzed, as is the norm in the experimental papers surveyed.

ACKNOWLEDGMENTS

This work was conducted as part of the Meta-analysis Working Group (Meta-analysis, interaction strength and effect size: application of biological models to the synthesis of experimental data) supported by the National Center for Ecological Analysis and Synthesis (NCEAS), a center funded by NSF (DEB 94-21535), the University of California, Santa Barbara, and the State of California. G. Englund was supported by a post-doctoral fellowship from the Swedish Natural Research Council, and O. Sarnelle by a post-doctoral fellowship from NCEAS. This research also was partially supported by a NSF grant DEB-9407591 to S. D. Cooper. The development of our ideas benefitted greatly from interactions with the entire Working Group. We also wish to thank the NCEAS staff for their support and hospitality.

LITERATURE CITED

- Abrami, P. C., P. A. Cohen, and S. d'Apollonia. 1988. Implementation problems in meta-analysis. *Review of Educational Research* **58**:151–179.
- Allan, J. D. 1982. The effects of reduction in trout density on the invertebrate community in a mountain stream. *Ecology* **63**:1444–1455.
- Connell, J. H. 1983. On the prevalence and relative importance of interspecific competition: evidence from field experiments. *American Naturalist* **122**:661–696.
- Cooper, H. M. 1984. The integrative research review: a systematic approach. Sage Publications, Beverly Hills, California, USA.
- Cooper, H. M., and L. V. Hedges. 1994. The handbook of research synthesis. The Russell Sage Foundation, New York, New York, USA.
- Cooper, S. D., S. J. Walde, and B. L. Peckarsky. 1990. Prey exchange rates and the impact of predators on prey populations in streams. *Ecology* **71**:1503–1514.
- Dahl, J., and L. Greenberg. 1996. Impact on stream benthic prey by benthic vs drift-feeding predators: a meta-analysis. *Oikos* **77**:177–181.
- Downing, J. A., C. W. Osenberg, and O. Sarnelle. 1999. Meta-analysis of marine nutrient-enrichment experiments: variation in the magnitude of nutrient limitation. *Ecology* **80**:1157–1167.
- Dudgeon, D. 1993. The effects of spate-induced disturbance, predation and environmental complexity on macroinvertebrates in a tropical stream. *Freshwater Biology* **30**:189–197.
- Englund, G. 1997. Effects of spatial scale and prey movements in predator caging experiments. *Ecology* **78**:2316–2325.
- Englund, G., and T. I. Olsson. 1996. Treatment effects in a predator caging experiment: influence of predation rate and prey movements. *Oikos* **77**:519–528.
- Feltmate, B. W., and D. D. Williams. 1989. Influence of rainbow trout (*Oncorhynchus mykiss*) on density and feeding behavior of a perlid stonefly. *Canadian Journal of Fisheries and Aquatic Sciences* **46**:1575–1580.
- Feltmate, B. W., and D. D. Williams. 1991. Evaluation of predator-induced stress on field populations of stoneflies (Plecoptera). *Ecology* **72**:1800–1806.
- Glass, G. V., B. McGaw, and M. L. Smith. 1981. Meta-analysis in social research. Sage Publications, Beverly Hills, California, USA.
- Goldberg, D. E., T. Rajaniemi, J. Gurevitch, and A. Stewart-Oathen. 1999. Empirical approaches to quantifying interaction intensity: competition and facilitation along productivity gradients. *Ecology* **80**:1118–1131.
- Gurevitch, J., and L. V. Hedges. 1993. Meta-analysis: combining the results of independent studies in experimental ecology. Pages 378–398 in S. Scheiner and J. Gurevitch, editors. The design and analysis of ecological experiments. Chapman & Hall, New York, New York, USA.
- Gurevitch, J., and L. V. Hedges. 1999. Statistical issues in ecological meta-analyses. *Ecology* **80**:1142–1149.
- Gurevitch, J., L. L. Morrow, A. Wallace, and J. S. Walsh. 1992. A meta-analysis of competition in field experiments. *American Naturalist* **140**:539–572.
- Hedges, L. V., J. Gurevitch, and P. Curtis. 1999. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**:1150–1156.
- Hunter, J. E., and F. L. Schmidt. 1990. Methods of meta-analysis: correcting error and bias in research findings. Sage Publications, Newbury Park, California, USA.
- Lancaster, J. 1990. Predation and drift of lotic macroinvertebrates during colonization. *Oecologia* **85**:48–56.
- Lancaster, J., A. G. Hildrew, and C. R. Townsend. 1990. Stream flow and predation effects on the spatial dynamics of benthic invertebrates. *Hydrobiologia* **203**:177–190.
- Lancaster, J., A. G. Hildrew, and C. R. Townsend. 1991. Invertebrate predation on patchy and mobile prey in streams. *Journal of Animal Ecology* **60**:625–641.
- Mahoney, M. J. 1977. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research* **1**:161–175.
- Nisbet, R. M., S. Diehl, W. G. Wilson, S. D. Cooper, D. D. Donaldson, and K. Kratz. 1997. Primary-productivity gradients and short-term population dynamics in open systems. *Ecological Monographs* **67**:535–553.
- Osenberg, C. W., O. Sarnelle, and S. D. Cooper. 1997. Meta-analysis, interaction strength, and effect size in ecological experiments: the need for biological models. *American Naturalist* **150**:798–812.
- Osenberg, C. W., O. Sarnelle, S. D. Cooper, and R. D. Holt. 1999. Resolving ecological questions through meta-analysis: goals, metrics, and models. *Ecology* **80**:1105–1117.
- Peckarsky, B. L. 1985. Do predaceous stoneflies and siltation affect the structure of stream insect communities colonizing enclosures? *Canadian Journal of Zoology* **63**:1519–1530.
- . 1991. A field test of resource depression by predatory stonefly larvae. *Oikos* **61**:3–10.
- Peckarsky, B. L., and S. I. Dodson. 1980. Do predaceous stoneflies influence benthic distributions in streams. *Ecology* **61**:1275–1282.
- Peckarsky, B. L., S. C. Horn, and B. Statzner. 1990. Stonefly

- predation along a hydraulic gradient: a field test of the harsh-benign hypothesis. *Freshwater Biology* **24**:181–191.
- Reice, S. R. 1991. Effects of detritus loading and fish predation on leafpack breakdown and benthic macroinvertebrates in a woodland stream. *Journal of the North American Benthological Society* **10**:42–57.
- Rosenberg, M. S., D. C. Adams, and J. Gurevitch. 1997. MetaWin: Statistical software for meta-analysis with re-sampling tests. Sinauer Associates, Sunderland, Massachusetts, USA.
- Sarnelle, O. 1992. Nutrient enrichment and grazer effects on phytoplankton in lakes. *Ecology* **74**:551–560.
- Schlosser, I. J., and K. K. Ebel. 1989. Effects of flow regime and cyprinid predation on a headwater stream. *Ecological Monographs* **59**:41–57.
- Schoener, T. W. 1983. Field experiments on interspecific competition. *American Naturalist* **122**:240–285.
- Sih, A., and D. Wooster. 1994. Prey behavior, prey dispersal, and predator impacts on stream prey. *Ecology* **75**:1199–1207.
- Slavin, R. E. 1984. Meta-analysis in education: how has it been used? *Educational Researcher* **13**(8):6–15.
- Wiseman, S. W., S. D. Cooper, and T. L. Dudley. 1993. The effects of trout on epibenthic odonate naiads in stream pools. *Freshwater Biology* **130**:133–145.
- Wooster, D. 1994. Predator impacts on stream benthic prey. *Oecologia* **99**:7–15.

APPENDIX

The complete list of stream predation studies used in our meta-analysis and the data extracted from them are available in ESA's Electronic Data Archive: *Ecological Archives* E080-007.