BULLETIN OF MARINE SCIENCE. 53(l): 128-153. 1993

EXPERIMENTS WITH FRESHWATER INVERTEBRATE ZOOPLANKTIVORES: QUALITY OF STATISTICAL ANALYSES

Stuart H. Hurlbert and Michael D. White

ABSTRACT

We examined the statistical analyses of experimental data in 95 papers published during 1966-1990 on the ecology, physiology, and behavior of freshwater invertebrate zooplanktivores.

Serious statistical errors were found in 51% of the papers. The frequencies of particular types of errors were as follows: sacrificial pseudoreplication (31%), simple pseudoreplication (11%), temporal pseudoreplication (7%), pseudofactorialism (5%), misuse of one-tailed tests (4%), and failure to use log-transformation in multi-way ANOVA when appropriate (9%). Several of these frequencies are much higher when calculated on the basis of the number of papers where a given type of error is possible. For example, pseudofactorialism occurred in 28% of the papers using multi-way ANOVA, the only type of paper for which this error is possible. We hope that this exhaustive review of a very specific but important area of plankton research will create a positive feedback network among this core group of researchers. This could so improve the use of statistics in future plankton studies that this field could serve as a model for other areas of biology equally plagued by improper statistical analysis.

The quality of statistical analyses in the biological sciences is poor, as numerous critical reviews have shown. Indeed, if the frequencies of statistical error reported by reviews of the ecological (Hurlbert, 1984), marine biological (Underwood, 1981), behavioral (Machlis et al., 1985; Kroodsma, 1989), and biomedical (Schor and Karten, 1966; Glantz, 1980; O'Fallon et al., 1978) literature are representative, the *majority* of papers that employ statistical analyses contain serious statistical errors. The errors are diverse in kind, and their consequences for the validity and strength of conclusions are thus also diverse. In many disciplines one of the commonest types of error is pseudoreplication (sensu Hurlbert, 1984). Its most usual consequence is the underestimation of *P* values in significance tests. These values are often one to several orders of magnitude lower than the *P* values that would have been yielded by correct statistical procedures. The acceptance of the conclusions of such papers by trusting editors and readers has been based on false premises. Reanalysis of many specific findings will show them to be statistically unsupported. Such error may not permanently bias our understanding of any general phenomena (e.g., the ability of substances produced by *Chaoborus* to induce production of protective spines and helmets in certain prey species) that ultimately are studied in multiple experiments by multiple investigators. It does seem likely to slow progress and to lead us on wild goose chases occasionally. The greatest immediate consequence, however, is the tremendous burden it places on conscientious editors, reviewers, thesis advisors, and statisticians. The morass of incorrect statistical analyses in the literature creates a Sisyphean task for them. It provides an abundance of negative models that continually are undoing their instructional efforts.

The immediate sources of this problem are many, but they all may derive chiefly from lack of conceptual and terminological clarity in many key areas of statistics. If so, attempts to clarify concepts and terminology through critiques of large numbers of specific examples of statistical malpractice may be the most effective countermeasure. This was the premise of an earlier review (Hurlbert, 1984), the favorable reception of which has reinforced our faith in that approach.

Whereas that earlier review was based on a sampling of the ecological literature, primarily for the period of 1960-1980, the present paper has a narrower focus. We undertake *to review the statistical analyses of most of the experimental papers concerning the ecology, physiology and behavior of freshwater invertebrate zooplanktivores published during the period* 1966-1990. This is a small corpus of work that represents one of the most productive and dynamic areas of ecology over the last two decades.

Our hope is this: to create a solid "statistical conscience" among the still small number of people studying invertebrate zooplanktivores and to so elevate the quality of statistical analysis in the future literature on this topic that it can serve as a model for other areas of biology.

This objective is feasible. The most frequent errors relate to simple, if widely misunderstood, statistical concepts. The number of researchers in this area is still small. Many personal and professional relationships exist among these researchers. In particular, anonymously or otherwise, they do a lot of reviewing of each others' manuscripts. These invertebrate zooplanktivore researchers thus constitute a natural network. If this "self-improvement" network has the desired effect, then each individual in it can become a node for good statistical advice to students and colleagues in the other networks, especially academic departments, in which they operate.

Finally, lest Popperian readers disparage the scientific nature of this contribution, we state one of our objectives as a testable hypothesis. In our earlier review, which found a rather low frequency (12% of all papers, and 26% of those applying inferential statistics) of pseudoreplication in the plankton literature, we concluded that planktologists were "relatively virginal" (Hurlbert, 1984). Our hypothesis therefore shall be that this is true, specifically that the frequency of pseudoreplication does not significantly exceed 12%. The alternative hypothesis then is that planktologists have an error frequency in excess of 12%, in which case they shall merit some more severe appellation.

METHODS

As indicated, we have attempted to examine the majority of papers published during the period 1966-1990 that report experimental work on freshwater invertebrate zooplanktivores. We guess that the 95 papers examined represent perhaps 70-80% of those published during 1966-1990 on this topic. We have further restricted our survey to papers published in English. This is not a severe limitation. Relatively few of the experimental studies in this area have been published in other languages, and, of those that have, only a small percentage have used inferential statistics to analyze their data.

For every major experiment in a paper, we examined the statistical procedures, if any, that were used to analyze it. We looked only for certain types of errors that have high potential for causing large errors in the estimation of P values. These are defined as follows. Examples of each will be discussed in the Results section.

Types of Errors

1. Simple Pseudoreplication (Hurlbert, 1984). — There is a single experimental unit per treatment, but multiple measurements are made on each experimental unit at each monitoring period. These multiple measurements are then treated statistically as if each represented a separate experimental unit. In essence this represents confusion between the experimental unit and the *evaluation unit*. The latter is defined as that element of an experimental unit on which an individual measurement is made (Urquhart, 1981). Where the experimental unit is an aquarium, for example, an evaluation unit might be an individual copepod or water sample from the aquarium.

2. *Temporal Pseudoreplication (Hurlbert, 1984).* — In its simplest version, this is similar to simple pseudoreplication except that the multiple measurements on an experimental unit are taken successively in time. If the successive measurements on a given unit are treated statistically as if each represented a different experimental unit, temporal pseudoreplication is the result. The flawed analysis often is in the form of a t-test, or a 1-way ANOVA that reports $a(t - 1)$ error degrees of freedom. Or it may be in the form of a 2-way ANOVA with time treated as a blocking factor and with, in the case of a completely randomized design, $(a - 1)(t - 1)$ or $at(n - 1)$ error degrees of freedom (a = no. of treatments, n = no. replicate experimental units per treatment, t = no. of monitoring times per experimental unit).

3. *Sacrificial Pseudoreplication (Hurlbert, 1984).* — When the number of experimental units (n) per treatment is 2 or more and when the number of evaluation units (k) measured per experimental unit is 2 or more, some analyses ignore the structure in the set of nk measurements per treatment and treat each measurement as if it represented an independent replicate of the treatment. Such analyses constitute sacrificial pseudoreplication. They ignore or "sacrifice" the opportunity to partition the total variance into "among experimental units" and "within experimental units" components and hence to carry out a valid analysis. If the replicate measurements made on each experimental unit are made at successive points in time, then analyses that ignore the structure in the nk measurements per treatment may be said to exemplify both sacrificial and temporal pseudoreplication.

4. *Pseudofactorialism.* — This is a new term for an increasingly common error that is discussed at length elsewhere (Hurlbert, in prep.). It is defined as the invalid statistical analysis that results from the misidentification of two or more *response variables* as representing different levels of an *experimental variable*. Most often the invalid analysis consists of the use of an $(n + 1)$-way ANOVA in a situation where two or more n-way ANOVAs would be the appropriate approach.

For example, an experiment is conducted to assess the effects of three notonectid densities (high, medium, low) on the July abundance of six zooplankton species co-occurring in experimental ponds. This calls for a separate 1-way ANOVA for each zooplankton species. If it is viewed as a 3×6 factorial experiment, however, and analyzed with the corresponding 2-way ANOVA, then pseudofactorialism is being committed.

5. *Metric-interaction Mismatch.* — This is a new term for another common error that is defined and discussed at length elsewhere (Hurlbert and White, in prep.). It concerns the appropriate *metric* or scale for expressing the effect of a factor (treatment or blocking), and it concerns the interpretation of *interaction* effects (or lack thereof) in multi-way ANOVAs. We claim that for most response variables, "magnitude of effect" is most appropriately and meaningfully measured as *percent* change rather than as *absolute* change. When this is true, assessments of factor interaction in multi-way ANOVAs are meaningful only when the data are log-transformed prior to analysis. This failure to log transform in these situations we have deemed an error, a metric-interaction mismatch.

6. Misuse of 1-tailed Tests. — Despite the confusing advice inmost statistics books (Lombardi and Hurlbert, in prep.), biologists generally make scant use of 1-tailed tests. Most uses in the biological literature constitute error. The decisive criterion usually should have nothing to do with the scientist's suspicions (or lack thereof) as to what the direction or sign of the expected effect will be. The criterion rather should be what the scientist's attitude likely would be if, after an a priori decision to use a 1-tailed test, the effect found is strongly in the opposite direction from that expected. *If it seems likely that the scientist would be tempted to carry out some further analysis,* e.g., a 2-tailed test or the other 1-tailed test, *then this is sufficient to allow classification of the original 1-tailed test as inappropriate and an error.* Such an assessment is not as subjective as it may seem. In most circumstances the scientist *will* be so tempted, because the only alternatives to further analysis, namely the discarding of information and/or the repeating of the experiment, will be unattractive.

7. Insufficient Information. — In many instances our evaluation of experiments was hindered by lack of fundamental information. This lack did not constitute statistical error, though in some cases it may have prevented us from detecting such. Regardless, the lack of this basic information must be regarded as a deficiency where it occurs.

We recognized two categories under this heading. In one that we call "Evaluation impossible" (Ei), the existence of treatment effects is impossible to assess formally either because treatments were unreplicated or because no information was provided on variability among replicates within treatments, neither in the form of statistical tests nor in the form of standard errors or deviations, nor in the form of data for individual experimental units.

In the other category, termed "Description inadequate" (Di), we placed experiments that were analyzed statistically but without sufficient information on the experimental design and/or procedures to allow for determination of whether the procedures used were appropriate.

Finally, there were several situations where circumstantial evidence suggested the presence of a particular error, but where it was not possible to be certain. In these instances, a status of description inadequate could be justified but we have opted instead to indicate the suspected error with the appropriate symbol followed by "?" to indicate the uncertainty.

8. Everything Fine. — Experiments which were adequately described, and which contained none of the errors on which we were focusing, were rated as "fine" as symbolized by an asterisk (*) in our tabulation. Many of these experiments made no use of inferential statistics, which, by itself, is not necessarily a weakness. The asterisk (*) has no wider meaning than the above. We do not intend it as an indicator of the overall quality or value of an experiment or statistical analysis.

Conduct of the Literature Survey

Our survey covered both field and laboratory studies. Most concerned physiological, behavioral, or ecological phenomena.

Our original intent was to focus exclusively on manipulative experiments (sensu Hurlbert, 1984), and we generally did not review papers that reported only mensurative experiments and observational studies. However, many papers reported both mensurative and manipulative experiments, and we have sometimes analyzed the mensurative experiments in such papers when these were subject to statistical analysis. Though the several types of errors listed above have been defined in terms of manipulative experiments, they do, in most cases, have their counterparts in mensurative experiments.

The distinction between manipulative and mensurative experiments is often subtle. For example, if we set up 5 aquaria, put one *Mesocyclops* female and 20 individuals of each of 3 prey species into each, and compare the percent survival at 24 h of the 3 prey species, we have *not* conducted a manipulative experiment. We cannot identify different experimental units that have received different treatments. Though we have suggested it may be useful to designate such studies as mensurative experiments (Hurlbert, 1984), biologists and statisticians should both be aware that within the discipline of statistics, such studies are conventionally regarded as "observational," a category of very wide scope.

When more than one experiment was reported in a given article, we have been flexible in how we have reported our findings (Table 1). When the different experiments used the same statistical procedures and these all were fine or all committed the same error, we have not listed the experiments separately. On the other hand, when the different experiments used different statistical tests and/or represented different errors, we have analyzed and listed the experiments separately. In several instances where different response variables in a single experiment were analyzed by different statistical procedures, we have evaluated the analysis for each response variable separately.

The frequencies of the different types of error were calculated on a per paper basis (Table 2). That is, we determined the number of papers that contained one or more examples of a given type of error, and then divided by the total number (95) of papers examined. These are conservative indices of the extent of statistical malpractice. We would have obtained much higher frequencies if we had based our calculation only on the number of papers where a given type of error was possible. For example, temporal pseudoreplication, by definition, was a possibility only in experiments that were monitored on more than one date or time period.

RESULTS AND DISCUSSION

The results of our analyses of the individual papers are given in Table 1. The frequencies of the different types of errors, including a partial breakdown by time period, are reported in Table 2. Examples of the different types of errors have been schematized in Figures 1-6.

The most general finding is that the overall percentage of papers containing one or more serious statistical errors is very high: 51%. Pseudoreplication, in particular sacrificial pseudoreplication, was the most common error found. Occurring in 41% of the papers, pseudoreplication was almost four times more frequent than in the freshwater plankton studies reviewed earlier (Hurlbert, 1984) where its frequency was 12%. The difference is significant ($\chi^2 = 11.1$, $P < 0.001$), so we must reject our null hypothesis. Freshwater planktologists are not "relatively virginal" in matters statistical. "Virginoid" seems too mild an appellation, given the 51% overall error rate. Perhaps "biologists of the night"?

We emphasize again the conservative nature of our error estimates. Many of these would be 2 to 10 times higher if calculated only on the basis of the number of studies where they were possible.

When not only lack of error but also adequacy of information (Ei, Di) was considered, only 36% of the 95 papers were judged satisfactory or "everything fine" (Table 2).

The only positive aspect of the data is the *slight* indication that during recent years the frequency of pseudoreplication has declined and that of "everything fine" papers has increased relative to their frequencies during the preceding two decades.

Pseudoreplication. — SIMPLE PSEUDOREPLICATION. This was found in papers by Gilbert (1967, 1975), Gilbert and Kirk (1988), Gilbert and Stemberger (1984), Grant and Bayly (1981), Hebert and Grewe (1985), Kerfoot (1975, 1987), Murdoch et al. (1984), Stemberger and Gilbert (1984), and Wong et al. (1986), and probably was present in that by Li and Li (1979). The example shown in Figure 1 is typical. In it the individual aquaria represent the experimental units, and the individual *Daphnia* represent what Urquhart (1981) has very usefully termed the *evaluation units*. In all situations, the number of error df available for testing for treatment effects is a function of the number of experimental units per treatment and is independent of the number of evaluation units measured per

Table 1. Some statistical errors in experimental studies of invertebrate zooplanktivores. Types of errors are abbreviated as follows: Di = description of experimental design and/or statistical procedures insufficient to allow determination of whether appropriate procedures were used, Ei = no statistical errors present but evaluation of treatment effects *impossible* owing either to lack of treatment replication or to absence of information on variability among treatment replicates, M = metric-interaction mismatch, Pf = pseudofactorialism, Sa = sacrificial pseudoreplication, Si = simple pseudoreplication, T = temporal pseudoreplication, t1 inappropriate use of a one-tailed test, X = major error of some other sort (see text), * = none found

Article (author, year)	Nature of experiment (effect of...)	Statistical procedure	Errors detected
Addicott (1974)	Larval *Wyeomyia* density and plant size on protozoan abundance, diversity, evenness, species richness	Linear regression	*
Arts et al. (1981)	Acilius (dytiscid) on *Daphnia* density and vertical distribution	1. *t*-test 2. Wilcoxon test	Sa Sa
Barry and Bayly (1985)	I. *Anisops* density on *Daphnia* crest development	3-way ANOVA	Ti; Sa?
	2. *Anisops* "water" on *Daphnia* crest development	2-way ANOVA	*
	3. Predator species (9) on *Daphnia* crest development	multiple 2-way ANOVA	*
	4. *Daphnia* size and crest development and predator species on attack success	3-way ANOVA	Pf
	5. Temperature and *Daphnia* phenotype on *Anisops* attack success	2-way ANOVA	Pf
Brandl and Fernando (1974)	I. *Ceriodaphnia* size on consumption by Acanthocyclops	"Wilcoxon-White test"	T
	2. Prior diet regime on consumption of *Ceriodaphnia*	"Wilcoxon-White test"	T
Brandl and Fernando (1975)	Prey species on predation by Cyclops and *Mesocyclops*	None	*?
Brandl and Fernando (1978)	Prey species on cyclopoid copepod electivities	*?*	*Di*
Brandl and Fernando (1981)	Cyclopoid predation on prey densities	None	Ei
Confer (1971)	1. *Diaptomus* density on *Mesocyclops* feeding rate	None	*
	2. *Diaptomus* instar on *Mesocyclops* feeding rate	χ^2-test	Sa
Cooper (1983)	Prey species, light and predator size on predation by insects	Sign test; U-test	*
Cooper and Goldman (1980)	Prey type, prey abundance, *Mysis* abundance, availability of alternative prey on consumption rates of *Mysis*	Wilcoxon rank sum test, sign test, median test	*
Cooper and Smith (1982)	1. *Daphnia* species on predation by various insects	Sign test	*
	2. *Chaoborus* on *Daphnia* densities	*U*-test	*
Dodson (1974)	*Diaptomus* on *Daphnia* and *Chaoborus density*	Correlation	*?
Dodson (1988a)	*Chaoborus, Notonecta,* and *Lepomis* on *Daphnia* morphology	Nested ANOVA	*
Dodson (1988b)	1-2. *Chaoborus* and *Notonecta* on *Daphnia* vertical position	2-way ANOVA	T
	3-4. Predator-conditioned water on *Daphnia* vertical position	1 -way ANOVA	T?
Dodson (1989)	*Chaoborus, Notonecta* and *Lepomis* on *Daphnia*	1-way ANOVA	*

Table 1. Continued

Article (author, year)	Nature of experiment (effect of...)	Statistical procedure	Errors detected
Dodson and Cooper (1983)	*Craspedacusla* on zooplankter densities	*t*-test; *U*-test	*
Dodson and Havel (1988)	1. Notonecta and Daphnia clone on Daphnia body size (adults, neonates)	3-way ANOVA	M, Pf
	2. Chemical vs. physical presence of Notonecta on Daphnia	2-way ANOVA	M
	3. Notonecta and algal density on *Daphnia*	2-way ANOVA	M
Von Ende and Dempsey (1981)	1-3. Prey species on survival in presence of *Chaoborus*	χ^2-test	Sa
	4. Initial *Bosmina* density and *Chaoborus* presence on final Bosmina density	2-way ANOVA?	M, Di
Fedorenko (1975)	Prey density and temperature on *Chaoborus* predation rate	Graphical	*
Folt (1987)	1. Copepod species on survival rate	*t*-test	*
	2. Diaptomus density on filtration rate	Regression	*
	3. Time and prey ratio on *Mysis* preference	2-way ANOVA, LSD	*
	4. Total prey density on predation risk	1-way ANOVA, LSD	*
Folt et al. (1982)	Prey density on *Mysis*		
	1. Predation rate	Graphical	*
	2. Preference for *Espischura*	Graphical, *t*-test?	Di
Gilbert (1966)	*Asplanchna*-substance and age of media on *Brachionus* spine length	None	*
Gilbert (1967)	1. *Asplanchna*-substance on *Brachionus* spine length		
	1. Experiments 1-2	*t*-test	Si
	2. Experiment "3"	*t*-test	*
	2. Embryonic stage on spine inducibility	None	*
	3. Various factors on spine production	None	*
	4. Spine presence on avoidance of ingestion	None	*
Gilbert (1973a)	Humps on cannibalism by *Asplanchna*	χ^2-test	*
Gilbert (1973b)	Alpha-tocopherol and cannibalism on *Asplanchna* morphotype	None	*
Gilbert (1975)	1. Tocopherol on *Asplanchna* size and morphotype frequencies	1. *G*-test 2. *t*-test	Si, Sa Si
	2. Prey type on *Asplanchna* size and morphotype frequencies	*t*-test	*, Sa
Gilbert (l976a)	1-9. Prey type on *Asplanchna* feeding response	*G*-test	Sa
Gilbert (1976b)	*Asplanchna* morphotype on reproductive rate	*U*-test	t1
Gilbert (1 976c)	1-4. Sex and clone on susceptibility to *Asplanchna* predation	*G*-test	Sa
Gilbert (1977a)	1-3. Prey type on *Asplanchna* predatory behavior	*G*-test	Sa
Gilbert (1977b)	Chemical stimuli on *Asplanchna* feeding response	*G*-test	*
Gilbert (1988)	Rotifer species on *Daphnia*-induced mortality rate	*t*-test; graphical	*

Table 1. Continued

Article (author, year)	Nature of experiment (effect of...)	Statistical procedure	Errors detected
Gilbert (1989)	*Daphnia* on rotifer and ciliate densities and growth rates	1. *t*-test	*
		2. 1-way ANOVA	*
Gilbert and Kirk (1988)	1. *Keratella* species on responses to *Asplanchna*	1. *t*-test	Si
		2. *G*-test	Si
	2. *Keratella* species on responses to *Daphnia*	*G*-test	Si
Gilbert and Stemberger (1984)	1-2. *Asplanchna*-conditioned medium on *Keratella* morphology	*t*-test	Si
	3. Spine presence on susceptibility to *Asp/anchna* predation	*G*-test	Sa
Gilbert and Williamson (1979)	*Asplanchna* and Mesocyc/ops, singly in combination, on *Polyarthra* and *Keratella* survival	*G*-test	Sa
Grant and Bayly (1981)	1-4. *Anisops* on *Daphnia* crest development	Fieller's theorem; *t*-test	Si
	5. *Daphnia* crest on *Anisops* predation	*U*-test	*
Hanazato and Yasuno (1989)	*Chaoborus* and *Pseudorasbora* (fish) on zooplankton densities	None	*
Hanazato and Yasuno (1990)	*Chaoborus* and insecticide on zooplankter densities	None	*
Hanazato (1990)	*Chaoborus* on *Daphnia* morphology	*t*-test	*
Havel (l985a)	1. *Chaoborus* density on *Daphnia* spine development	None	*
	2. Temperature on *Daphnia* spine development	Linear regression?	Sa?
Havel (l985b)	1. *Daphnia* morph on escape from *Chaoborus*	Binomial	Sa
	2. *Daphnia* size on predation		
	1. by *Chaoborus*	*U*-test	Sa
	2. by *Lepiodora* and *Acanthocyclops*	Regression	*?
Havel and Dodson (1984)	Prey type on *Chaoborus* attack success	Binomial	Sa, t1
Havel and Dodson (1987)	*Daphnia* morphotype on life history parameters	1. *t*-test	t1
		2. 2-way ANOVA	*
		3. Friedman's test	*
Havens (1990)	*Chaoborus* on zooplankter densities	*t*-test	*
Hebert and Grewe (1985)	*Chaoborus* "factor" on helmet size in 6 *Daphnia* clones	3-way ANOVA	Si, M, Di
Hebert and Loaring (1980)	1. Ratio of prey types on consumption by *Heterocope*	2-way ANOVA	Pf, Di
	2. Prey type on consumption by *Heterocope*	χ^2-test	Sa
Hewett (1980a)	Prey size on		
	1. *Didinium* size	2-way ANOVA	Pf
	2. *Didinium* growth rate	I -way ANOVA	*
Hewett (l980b)	1. Prey density and prey species on *Didinium* capture rate and division time	1. 2-way ANOVA	Sa, T, M
		2. Regression	Sa, T
	2. Prey density on *Didinium* size	*U*-test	*
Hewett (1988)	*Didinium* size and *Paramecium* size on *Didinium* predation		
	1. predation behavior	3-way ANOVA	Pf
	2. prey captures/division	2-way ANOVA	M
	3. division time	2-way ANOVA	M
Janicki and Decosta (1990)	*Mesocyc/ops* on survival of prey types	χ^2-test	Sa

Table 1. Continued

Article (author, year)	Nature of experiment (effect of.....)	Statistical procedure	Errors detected
Kerfoot (1975)	*Epischura* on *Bosmina* morph frequencies		
	1. 1-2 day results	χ^2-test	Si
	2. 4-day results	χ^2-test	Si, Sa
Kerfoot (1977)	Prey species on *Epischura* predation rate	χ^2-test	Sa?
Kerfoot (1987)	*Epischura* predation		
	I. on *Bosmina* morphotype frequencies	χ^2-test	Sa, Si
	2. on *Bosmina* mucro length	Graphical	Sa, Si
Kerfoot and Peterson (1980)	1. Transfer (environment) on *Bosmina*	χ^2-test	Sa?
	2. Bosmina sex on susceptibility to *Cyclops* predation	χ^2-test	Sa
Krueger and Dodson (1981)	1. *Chaoborus* on *Daphnia* spine development	*U*-test	*
	2. Developmental stage on susceptibility to *Chaoborus* factor	Median test	*
Li and Li (1979)	1. Prey species on *Acanthocyclops* predation rate	χ^2-test	X
	2. *Acanthocyclops* on prey swimming speed	Wilcoxon rank sum test	Si?
Luecke and O'Brien (1983)	1. Zooplankter species and temperature on *Heterocope* feeding rate	2-way ANOVA	M
	2. *Heterocope* on *Daphnia* densities	None	*
Lunte and Luecke (1990)	*Leptodora* on prey densities	*t*-test	*
McQueen (1969)	Prey density on consumption by *Cyclops*	None	*
Moore (1988)	Total rotifer density on *Chaoborus* Selectivity	*U*-test; Kruskal-Wallis	*
Murdoch and Scott (1984)	1. *Nolonecta* instar, prey species and prey size on no. prey consumed	None	*
	2. *Notonecta* instar and total *Daphnia* density	1-way ANOVA	*
	3. *Notonecta* on *Daphnia* temporal variability	*t*-test	*
	4. *Notonecta* (instar) on		
	1. Daphnia density and egg ratio	2-way ANOVA	M, T
	2. *Daphnia* mean size and fecundity	2-way ANOVA	M, Sa, T
	3. Daphnia size distribution	χ^2-test	Sa
	4. percent *Daphnia* ovigerous	2-way ANOVA?	T
	5. *Daphnia* adults as percent and death rate	1. 2-way ANOVA 2. *t*-test	T Sa, T
	6. *Daphnia* biomass	*t*-test	*
	5. *Notonecta* (instar) on *Ceriodaphnia* density	1-way ANOVA, *t*-test	*
Murdoch et al.(1984)	1. Previous diet on *Notonecta* prey preference	*t*-test	*
	2. Temperature on predator attack rate, etc.	Regression	Di
	3. *Notonecta* on		
	1. mosquito density	Graphical	*
	2. mosquito size distribution	χ^2-test	Si
	3. zooplankton temporal variability	None	*
Murtaugh (1981)	Zooplankter species on *Neomysis* predation rate	1-tailed Wilcoxon test	t 1

Table 1. Continued

Article (author, year)	Nature of experiment (effect of...)	Statistical procedure	Errors detected
Neill (1981)	*Chaoborus* density on zooplankter		
	1. prey consumed, recruitment and mortality and natality rates	None	Ei
	2. prey densities	1-way ANOVA	*
	3. body length	*t*-tests	Sa?
Neil! (1984)	1. *Chaoborus* on rotifer densities	1-way ANOVA	Ei
	2. *Diaplomus* and *Daphnia* on rotifer densities	1-way ANOVA	Ei
Nero and Sprules (1986)	1-2. Source and type of prey on *Mysis* clearance rate	2-way ANOVA	M
	3-4. Prey type and mysid age on clearance rate	2-way ANOVA	M
	5. Prey type of *Mysis* clearance rate	1-way ANOVA	*
O'Brien and Schmidt (1979)	*Bosmina* origin on *Heterocope* predation rate	None	*
O'Brien et al.(1979)	*Daphnia* size on *Heterocope* predation rate	Regression	*
O'Brien and Vinyard (1978)	1. Prey type on *Anisops* predation rate	None	*
	2-3. Prey type on survival and reproductive rates	None	*
Pastorok (1980)	1. Prey density and predator hunger on *Chaoborus* feeding behavior	None	*
	2. Prey species on *Chaoborus* feeding behavior	*t*-test, *U*-test	*
	3. Prey species on *Chaoborus* growth rate	ANCOVA	Sa?, Di
Peacock (1982)	*Chaoborus* on Cyclops		
	1. clutch size	2-way ANOVA	T?
	2. density	*U*-test	T?
	3. percent females with eggs	*t*-test	T?
	4. several other variables	None	Ei
Peacock and Smyly (1983)	1. Copepod species on no. individuals eaten by *Chaoborus*	None	Ei
	2. Distance between copepod and *Chaoborus* on number of copepods eaten	None	Ei
	3. Cyclops on		
	1. density of various zooplankters	None	*
	2. *Tropocyclops* survivorship	None	Ei
Riessen et al. (1988)	1. Prey type on probability of capture by contact with *Chaoborus*	χ^2-test	Sa
	2. Prey species on consumption rate by *Chaoborus*	1-way ANOVA	*
	3. *Chaoborus* on zooplankter prey densities	*t*-test	*
Salt (1974)	Densities of *Didinium* and *Pararnec:-urn* on *Didinium* predation rate	2-way ANOVA	M, Di
Schuize and Folt (1989)	1. Nauplius stage on *Epischura* predation	*t*-test	*
	2. *Epischura* density and phytoplankton presence on predation on nauplii	2-way(?) ANOVA, *t*-test, *U*-test	*
	3. Prey density on *Epischura* predation rate	Regression	*

Table 1. Continued

Article (author, year)	Nature of experiment (effect of...)	Statistical procedure	Errors detected
Schuize and Folt (1990)	1. Food type on *Epischura*		
	1. survivorship	LIFETEST/SAS	*
	2. egg production	Kruskal-Wallis	*
	2. Food density on *Epischura*		*
	1. survivorship	LIFETEST/SAS	
	2. egg production	Kruskal-Wallis	*
	3. predation rate	Regression	*
	3. Temperature on *Epischura* egg production	Wilcoxon rank-sum	*
Scott and Murdoch (1983)	Zooplankter size and species on *Notonecta* predation rate	Linear regression; *t*-test	*
Soto (1985)	Cyclops and *Daphnia* on *Diapwmus* spp. densities	1-way ANOVA; *t*-test	*
Sprules (1972)	1. Pond on zooplankter survival	*t*-test	Sa, T
	2. Prey species on susceptibility		*
	1. to *Chaoborus predation*	None	
	2. to *Ambyswma* predation	χ^2-test	*
Stemberger (1985)	Prey type, prey type ratio, and starvation period on *Diacyclops* predation	1-way ANOVA, *t*-test	*
Stemberger and Gilbert (1984)	1-3. *Asplanchna-*, *Mesocyclops-*, and Tropocyclops-conditioned media on *Keratella* spine development	*G*-test; *t*-test?	Si, Sa
	4. Spine presence on ingestion rate by predator	*t*-test	*
Stemberger and Gilbert (1987)	1-16. Predator- and competitor-conditioned media on *Kera(ella* spine production	*G*-test	Sa
	17-2 1. Spine presence on *Keratella* survival in presence of "enemy"	*t*-test	*?
Stenson (1987)	*Chaoborus* presence on *Holopedium* capsule size	*t*-test	Di
Vanni (1988)	*Chaoborus* and *Daphnia* on other zooplankters	2-way ANOVA; *t*-test	*
Vinyard and Menger (1980)	1. Prey density on number killed by *Chaoborus*	Graphical	*
	2. Prey type on evasion and escape success	None	Ei
	3. Prey density and alternative prey type on number eaten	None	Ei
Vuorinen et al. (1989)	Chaoborus-conditioned water on *Daphnia*		*
	1. carapace size	1 -way ANOVA	
	2. reproduction	None	*
	3. spine production	None	*
Walls and Ketola (1989)	*Chaoborus on Daphnia:*		
	1. No. instars with spines; no. with crest; no. neck spines	Kruskal-Wallis	*
	2. Clutch size; carapace length	1-way ANOVA	*
Williamson (1980)	Starvation time on predatory behavior of *Mesocyclops*	None	Ei
Williamson (1983)	Prey type and predator source on *Mesocyclops* predation behavior	Kruskal-Wallis	*?
Williamson (1984)	*Mesocyclops* on prey density (ingestion rate)	*F*-test	*

Table 1. Continued

Article (author, year)	Nature of experiment (effect of...)	Statistical procedure	Errors detected
Williamson (1987)	Phytoplankton, sex, prey type on *Diaptomus* predation behavior	Kruskal-Wallis	*
Williamson and Butler (1986)	1. Prey density on *Diaptomus* feeding behavior	None	*
	2. Algal density on *Diaptomus* predation rates on rotifers	None	*
	3. Rotifer species on *Diaptomus* predation rate	$\chi 2$-test	Sa
	4. Diaptomus predation on rotifer survival	Kruskal-Wallis	*
	5. Food type on *Diaptomus* survival and reproduction	None	Ei
Wong (1981)	*1. Bosmina* size on *Epischura* attack	1. Wilcoxon paired ranks	*
		2. χ^2-test	Sa
	2. Ratio of size classes on *Epischura* preference	χ^2-test	Sa
	3. Previous diet and *Bosmina* percent on preference for *Bosmina*	ANCOVA, Regression	Di
	4. Algal concentration on *Epischura* predation rate	Regression	*
Wong et al. (1986)	Presence of predaceous copepods on *Diaptomus* swimming behavior	Wilcoxon and Kruskal-Wallis	Si

experimental unit. When there is only one experimental unit per treatment, it is *not* legitimate to treat the evaluation units the evaluation units as surrogate experimental units. When this is done and when a significant difference is detected, all that has been done, strictly speaking, is to demonstrate that the two experimental units are probably not identical; there are no statistical grounds for attributing the difference to an effect of the experimental variable.

The simple pseudoreplication in Figure 2 is a more unusual sort and is complicated by the presence of additional problems in the analysis. Pseudoreplication is evident in that the ANOVA apparently used 178 error degrees of freedom in testing for treatment effects, while the experiment involved a total of only 36 experimental units (cups), one under each block-treatment combination. Clearly this analysis treated each individual *Daphnia* (evaluation unit) as if it represented a separate experimental unit, i.e., as if each *Daphnia* measured was treated and maintained in its own individual cup. This is not how the experiment was conducted.

Though not explicit in the paper, the spatial arrangement of cups on the laboratory bench in this experiment actually corresponded to a randomized block split-unit design, though without randomized assignment of levels of the sub-unit factor *(=Chaoborus)* within each whole unit (=a pair of adjacent cups each containing individuals from the same clone) (P. Hebert, pers. comm.). A block consisted of a row of 12 cups, a pair for each of the six clones.

The conventional analysis for such a design, ignoring the lack of randomization within whole units, is given at the bottom of Figure 2. Note that the error degrees of freedom available for testing for effects of the two experimental variables are 10 and 20, respectively.

This example is useful for discussing the various consequences of non-concordance between design and analysis. Hebert and Grewe (1985) could have ignored the split-unit

Table 2. Frequency of statistical errors in experimental studies on freshwater invertebrate zooplanktivores, 1966-1990

A. Percentage of papers having errors of different types (N = 95)

Percentage	Tyoe if error
41%	Pseudoreplication (of any sort)
31%	Sacrificial pseudoreplication *(Sa)*
7%	Temporal pseudoreplication *(I')*
11%	Simple pseudoreplication *(Si)*
5%	Pseudofactorialism *(Pf)*
9%	Metric-interaction mismatch *(M)*
4%	Misuse of one-tailed test *(t1)*
51%	Any of the above
36%	Information adequate (no Ei or Di evaluations) and no statistical errors detected

B. Temporal trends in frequency of major categories of error.

	Time period			Total
Number of	1966-1975	1976-1985	1986-1990	No. papers
papers examined	15	49	31	95
Percentage of papers containing				
1. Pseudoreplication	40%	49%	32%	39
2. Other types of error *(Pf M, t1)*	7%	20%	6%	17
3. Adequate information and no errors	40%	29%	48%	34

aspect, and analyzed the experiment as one with a simple randomized block design, with 10 $(=a(c - l)(n - 1))$ error degrees of freedom, as perhaps was their intention. Or they might have ignored both the split-unit aspect *and* the blocking and analyzed it as a completely randomized design, with 24 $(=ac(n - 1))$ error degrees of freedom. Either of these "rogue" analyses could have been criticized on the general ground that the analysis would not have corresponded to the design. Both could have been defended, however, on the more specific and relevant ground that this non-concordance of design and analysis would have been expected to cause a) either no change or a very slight decrease in the probability of a type I error if experimental variables had been without effect, and b) a decrease in the probability of a type II error, if there had been no effect of cup position (block effect or whole-unit effect) but had been effects of the experimental variables. That is, the analysis would have been both conservative with respect to type I error and more powerful for detecting real effects.

On the other hand, if the effects of cup position on helmet size had been large, either "rogue" analysis might have had less power to detect real effects. In that situation one would have needed only to be cautious in deciding how much weight was given to any finding of statistical non-significance.

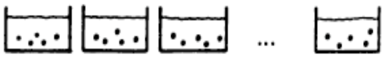
So, had Hebert and Grewe (1985) carried out one of the above-mentioned rogue analyses, we would have awarded it an asterisk in Table 1. Though on formal grounds some persons might classify these analyses as invalid, we would regard them as representing not even minor error. If cup position had no or a negligible effect, which seems the most likely situation, the analysis appropriate to a completely randomized design would, indeed, have been the ideal one.

As it was, the analysis actually carried out by Hebert and Grewe (1985) was classified as representing major error on the grounds that the actual probability of a type I error was increased to an unknown extent over the nominal one (α) by treatment of the *Daphnia*

EFFECTS OF T°C, TURBULENCE & NOTONECTIDS ON CREST DEVELOPMENT IN DAPHNIA CARINATA

(Grant & Bayly, 1981, L & 0)

DESIGN



Treatment No.	1	2	3	...	7
Turbulence	+	++	0	...	++
Notonectid	-	-	-	...	-
T℃	25	25	25	...	25

- 7 treatment
- I aquarium I treatment
- many Daphnia in each aquarium

ANALYSIS

Several pairwise comparisons, each testing a different hypothesis, e.g. 2" vs. "3" to test effect of turbulence.

All comparisons treat each individual Daphnia (= evaluation unit) as if it were a separate experimental unit. This =

SIMPLE PSEUDOREPLICATION

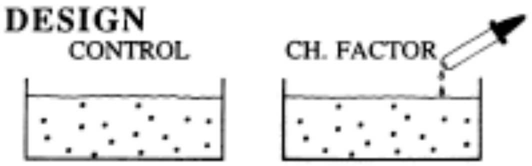
Figure 1. An example of simple pseudoreplication.

individuals as the experimental units. Implicit in their analysis was an assumption that, aside from differences due to block or treatment effects, the cups were *identical* with respect to properties that might influence helmet size. These would include intrinsic properties of the cups, properties determined by the experimenter (e.g., number and condition of *Daphnia* added, exact quantity of *Chaoborus* extract added, quality of water added, etc.) and properties due to chance events affecting only particular cups (contamination events, etc.). Even in the laboratory environment, it is unrealistic to assume lack of such "cup effects"; when adequate data are available to test for them, differences between "identical" experimental units are almost invariably detected (Hurlbert, 1984). This is a very different matter from that of cup *position* effects discussed earlier.

TEMPORAL PSEUDOREPLICATION. This was found in papers by Barry and Bayly (1985), Brandl and Fernando (1974), Dodson (1988b), Hebert and Grewe (1985), Hewett (1980b), Murdoch and Scott (1984), Peacock (1982) and Sprules (1972). It may be associated with any of several statistical procedures. When there is only one experimental unit per treatment, commonly a *t*-test or *U*-test is applied to the successive measurements in time (Peacock, 1982). When treatments are replicated it is common to find time invalidly treated as a blocking factor in a multi-

EFFECTS OF CHAOBORUS FACTOR ON HELMET DEVELOPMENT IN 6 DAPHNIA CLONES

(Hebert & Grewe, 1985, L & 0)



- 2 Chaoborus treatments (a)
- 6 Daphnia clones (c)
- 3 blocks (n), each containing 1 cup for each of the (ac) treatment combinations
- Several Daphnia measured in each cup (d_i, a variable)

ANALYSIS as reported

3-way ANOVA of Helmet size

Source	df	ms	F
Ch factor	(a-1) = 1	ms 1	msl/ms4
Clone	(c-l) = 5	ms 2	ms2/ms4
Replicate	(n-1) = 2	ms 3	ms3/ms4
Error	$(\sum d_i\text{-can}) = 178$	ms 4	

= SIMPLE PSEUDOREPLICATION AND??

ANALYSIS appropriate to design

ANOVA for randomized block split-unit design

Source	df	ms	F
Clone	(c-l) = 5	ms l	msl/ms3
Block(=replicate)	(n- 1) = 2	ms 2	ms2/ms3
Error (1)	(c-1)(n-1) =10	ms 3	
Ch. factor	(a-1) = 1	ms 4	ms4/ms6
Ch. f. x Clone	(a-1)(c-1) = 5	ms 5	ms5/ms6
Error (2)	(a)(c-1)(n-1)=20	ms 6	

Figure 2. An example of simple pseudoreplication in a complex design.

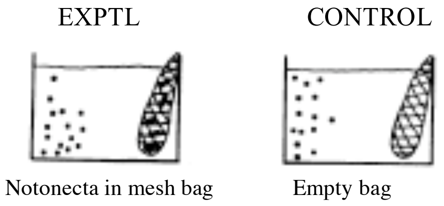
way ANOVA (Dodson, 1988b; Fig. 3). Time can function as a valid blocking factor, but this usually requires that measurements are made on a different set of experimental units at each successive monitoring time; laboratory experiments are occasionally designed this way, but field experiments rarely are.

Repeated measures designs with replicated treatments can be analyzed with a repeated measures ANOVA (Fig. 3) or by carrying out a separate ANOVA on each date. Repeated measures ANOVA is usually not the best way to analyze such data (Mead, 1988). If it is used, however, then the degrees of freedom for testing the time and time x treatment effects must be

EFFECTS OF CHAOBORUS AND NOTONECTA ON VERTICAL POSITION OF DAPHNIA

(Dodson, 1988, L & 0)

DESIGN



- 2 treatments (a)
- 3 aquaria / treatment (n)
- Determined mean depth of Daphnia in each aquarium
- at 3 times during 48h period (t)

ANALYSIS as reported

2-way ANOVA of Mean Depth

Source	df	ms	F
Predator	(a-1) = 1	ms l	msl/ms4
Time	(t-1) = 2	ms 2	ms2/ms4
TxP	(a-1)(t-1) = 2	ms 3	ms3/ms4
Error	at(n-1) = 12	ms 4	

ANALYSIS appropriate to design

Repeated Measures ANOVA of Mean Depth

Source	df	ms	F
Predator	(a-1)= 1	ms l	msl/ms2
Error(1)	a(n-1)=4	ms 2	
Time	ε(t-1)=?	ms 3	ms3/ms5
TxP	ε(a-1)(t-l)=?	ms 4	ms4/rns5
Error(2)	εa(n-1)(t-l)=?	ms 5	

Figure 3. An example of temporal pseudoreplication. e is a correction factor between 1 and 1/(*t* - 1) (see text).

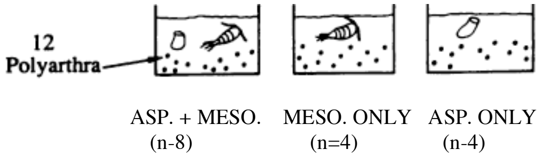
adjusted downward (from those appropriate to a split-unit design) by multiplying by the factor ε (Fig. 3). This factor can range from 1 to l/(t - 1) (Crowder and Hand, 1990; Milliken and Johnson, 1984).

SACRIFICIAL PSEUDOREPLICATION. This was by far the commonest error, being found in 31% of the papers (Tables 1, 2). By definition this error is possible only when a response variable is measured on two or more evaluation units in each experimental unit. Its frequency based solely on papers reporting such experiments would be about 60-80%; our records do not permit its exact calculation.

EFFECT OF ASPLANCHNA AND MESOCYCLOPS ON ROTIFER SURVIVAL

(Gilbert & Williamson, 1978, Oecologia)

DESIGN



RESULTS (sample)

Treatment	ASP. + MESO.	MESO.
Repl.No.	1 2 3 4 5 6 7 8	1 2 3 4
# Surviving Polyanhra after 12 h	1 6 2 0 2 6 3 0	0 1 2 0

ANALYSIS (typical) as reported

	ASP.+MESO.	MESO.	
# Surviving	20	3	**G-test, P=.016**
# Not Surviving	76	45	

= SACRIFICIAL PSEUDOREPLICATION

REANALYSIS of all comparisons

Comparison	P values G-test[a] (invalid)	P values t-test[b] (correct)
Polyartisra survival		
1. Asp. v. Anp.+Mcso.	<.000000001	.0014
2. Asp. v. Meso.	<.000000001	.006
3. Mono v. Asp.+Meso.	.016	.22
Keratella survival		
4. Asp. v. Mcuo.	.000019	.005
5. Meso. v. Aps.+ Meso.	.335	.70

[a] As reported in original paper.

[b] As carried out after arcsin transfomation of percent survival data presented in original paper. Analyses carried out with the software package Stat View 512+, version 1.1 (Brain Power 1986).

MORAL: Pseudoreplication of any sort **often** results in a P value one or more orders of magnitude **lower** than the 'true' P value.

Figure 4. An example of sacrificial pseudoreplication, with demonstration of how pseudoreplication leads to underestimation of P values.

The frequency of sacrificial pseudoreplication is surprising as it usually requires two conceptual errors: 1) the treatment of evaluation units as experimental units, and 2) the treatment of a nested data set as an unnested one. It is notable that the majority of cases of sacrificial pseudoreplication involved the χ^2-test or the *G*-test, and that almost every application of those tests to experimental data in the 95 papers examined represented either simple or sacrificial pseudoreplication (Table 1). Many of these were behavioral studies, and the misuse of such tests in animal behavior is indeed widespread (Machlis et al., 1985; Hurlbert and Lombardi, in prep.). In our experience, pseudoreplication is present in the majority of cases where these and other tests (e.g., Fisher's exact, Kolmogorov-Smirnov) for so-called categorical or enumeration data are applied to experimental biological data.

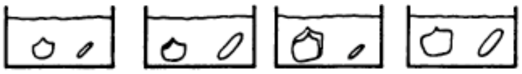
Sacrificial pseudoreplication, like other types of pseudoreplication, generally leads to *P* values that are severe underestimates of the *P* values yielded by valid statistical procedures. This can be demonstrated with an experiment (Gilbert and Williamson, 1979) where reporting of the data for individual experimental units (aquaria) has allowed us to reanalyze the data (Fig. 4). The original analysis employed the *G*-test and the sacrificial pooling of data for all replicates in each treatment. This would be a correct approach only if each individual *Polyarthra* or *Keratella* represented, in the physical conduct of the experiment, a separate experimental unit rather than simply one of 12 evaluation units *within an* experimental unit.

Correct analysis of these data might entail calculating a single datum, percent survival, for each experimental unit, applying the arcsin transformation to it, and carrying out t-tests for the comparisons

EFFECT OF DIDINIUM SIZE AND PARAMECIUM SIZE ON DIDINIUM CAPTURE SUCCESS, ETC.

(Hewett, 1988, Ecology)

DESIGN



- 2 x 2 factorial design: prey size (b: 1g. sm)
  vs. predator size (a: 1g. sm)
- 4 dishes per treatment combination (n)
- recorded for each dish two measures of
  capture time (C): 1) time to 1st capture (X_1),
  and 2) mean time between subsequent captures (X_2)

ANALYSIS as reported

3-way ANOVA of Capture Time

Source	df	ms	F
Predator size	(a-1) = I	msl	msl/ms4
Prey size	(b-l) = 1	ms2	ms2/ms4
Sequence (1st v. later)	(c-l)= 1	ms3	ms3/ms4
Error	abc(n-1) = 24	ms4	

= PSEUDOFACTORIALISM

ANALYSIS appropriate to design

2-way ANOVA of Capture Time

Source	df	ms	F
Predator size	(a-1)= I	ms l	msl/ms4
Prey size	(b-l) = 1	ms 2	ms2/ms4
A x B	(a-1)(b-l) = 1	ms 3	ms3/ms4
Error	ab(n-1) = 12	ms 4	

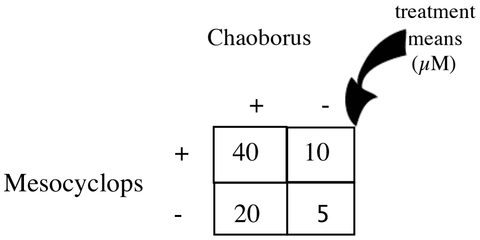
Figure 5. An example of pseudofactorialism.

of interest. Since there were three treatments, a slightly more conventional approach would be to first carry out a 1-way ANOVA and then make pair-wise comparisons via some procedure such as the least significant difference (LSD) method. This more conventional approach is in no way obligatory, however.

We have carried out *t*-tests on the data of Gilbert and Williamson (1979) as this allows comparison of the *P* values so obtained with the *P* values yielded by their use of the *G*-test. The *t*-tests yielded *P* values that were in all cases much greater, sometimes by more than 5 orders of magnitude, than the *P* values yielded by the *G*-test (Fig. 4). This reflects the fact that the *G*-tests presumed a much more powerful experiment, one involving 48 or 96 experimental units per treatment, whereas the *t*-tests presume, correctly, that there were only 4 or 8 experimental units per treatment.

LOG TRANSFORMATION
OBLIGATORY (usually) IN
MULTI-WAY ANOVAS

Hypothetical example: 2x2 factorial experiment to test for effects of *Chaoborus* and *Mesocyclops* on *Keratella* spine length



IS THERE FACTOR INTERACTION??

ANOVA on untransformed data: YES

ANOVA on log-transformed data: NO!

In biological terms: NO!
-because the percent increase caused by each predator is independent of whether the other is present or absent

MORAL: Wherever ‘effect’ is most appropriately or meaningfully measured as **percent** change rather than as **absolute** change and where a multi-way ANOVA is used, data should be log-transformed prior to analysis.

Figure 6. Hypothetical example showing the importance of log-transformation to the clear interpretation of tests for factor interaction.

The five comparisons are useful for demonstrating when pseudoreplication is most likely to affect biological conclusions. If the *P* value yielded by an analysis representing pseudoreplication is "clearly non-significant" (e.g., >0.10), as in comparison 5 (Fig. 4), then carrying out a correct analysis usually will increase the *P* value, and the biological conclusion of "no clear effect" will be unchanged. Likewise, if the erroneously calculated *P* value is extremely small (e.g., <<0.001), as in comparisons 1, 2 and 4, then the *P* value yielded by a correct analysis will usually be higher than the erroneous value but still lower than the maximum value (e.g., 0.05) we often are willing to accept as "definitive evidence of an effect"; so again our conclusions would be unaffected. But if our miscalculated *P* values indicate "intermediate" levels of significance (e.g., $0.001 < P < 0.05$), then their recalculation in a correct manner will often completely alter our conclusions (e.g., comparison 3 in Fig. 4).

The consequences of pseudoreplication for *P* values can only be stated in general terms, as above, because the consequences are of course dependent on the properties of the particular data set and on the degree to which error degrees of freedom are inflated. Almost always pseudoreplication leads to underestimation of *P*, though there are rare circumstances in which pseudoreplication can also lead to overestimation of P (Hurlbert and Lombardi, in prep.).

Pseudofactorialism. -This problem was present in studies by Barry and Bayly (1985), Dodson and Havel (1988), Hebert and Loaring (1980), and Hewett (1980a, 1988) (Table 1). Though these represent only 5% of the papers reviewed (Table 2), they represent 28% of the 18 papers using multi-way ANOVAs, the only type of paper for which the error is possible.

In the study by Hewett (1988) some analyses were of the sort shown in Figure 5. Two successive measurements were made on each experimental unit (= aquarium), and these were treated statistically as if they represented two levels ("first," "later") of a third experimental variable ("position in sequence"). In fact they represent two different response variables and nothing more. That the 3-way ANOVA was inappropriate can be appreciated by observing that the error degrees of freedom that Hewett used in testing for main effects exceed (24 vs. 16) the total number of experimental units used in the experiment.

In the analysis carried out, position in sequence may be regarded as a "pseudofactor" in the sense that it is a "false" treatment factor or experimental variable. An alternative and equally valid view is that "position in sequence" represents the factor "time." In that case time is the pseudofactor in the sense that it is a false blocking factor. In the experiment, time is a repeated measures "factor," not a blocking factor. Inconsistent and conflicting usage of the terms "factor" and "factorial" by statisticians seems to be a major source of confusion in this area (Hurlbert, in prep.).

Two sorts of correct analyses for the data of Hewett (1988) would be possible. If there is interest in testing not only for effects of the two experimental variables but also for a difference between the two "times" or response variables (X_1, X_2), then a repeated measures 2-way ANOVA would be suitable. If there is no interest in testing H_0: $X_1 = X_2$, then separate, simple 2-way ANOVAs for X_1, X_2, and, possibly, X_1/X_2 or $X_1 - X_2$ would be sufficient (Fig. 5, bottom). In either case there would be 12 error degrees of freedom for testing for main effects.

This example is useful in demonstrating that temporal pseudoreplication and pseudofactorialism are so defined as to be partly overlapping categories of error. To the extent that position in sequence in Hewett's example is equated with time, the error made also corresponds to a type of temporal pseudoreplication.

Many more examples of pseudofactorialism are discussed elsewhere (Hurlbert, in prep.). The frequency of this error, rare in the older literature, seems clearly a consequence of the ease with which multi-way ANOVAs can be carried out by canned programs at little cost in time or mental effort to the investigator.

Metric-interaction Mismatch. — This problem was detected in Dodson and Havel (1988), Von Ende and Dempsey (1981), Hebert and Grewe (1985), Hewett (1980b, 1988), Luecke and O'Brien (1983), Murdoch and Scott (1984), Nero and Sprules (1986), and Salt (1974). These studies applied multi-way ANOVAs to data that had not been, but should have been, log-transformed. As a consequence their tests for factor interaction may be flawed.

Though these represent only 9% of the papers reviewed, they represent 50% of the 18 papers that used multi-way ANOVAs, the only sort of paper for which the error is possible.

None of the cited studies provide enough information to allow determination of the degree to which interpretations of factor interaction may be faulty. So we present a hypothetical example of a 2 x 2 factorial experiment (Fig. 6) to demonstrate the general nature of the problem. In this example we presume that each treatment mean is based on values from several experimental units (e.g., aquaria). We presume that both *Chaoborus* and *Mesocyclops* do affect spine length in *Keratella*. We further presume that the variation among experimental units (within a treatment) is very small, so that the real effects are detectable statistically and so that the treatment means will be essentially the same whether calculated as arithmetic means (i.e., without transformation of data) or as geometric means (i.e., with log transformation of data).

In this example an ANOVA applied to untransformed data will lead to the conclusion that there is strong factor interaction. This is because *Mesocyclops* caused a 20 µm increase in spine length in the presence of *Chaoborus* but only a 5 µm increase in the absence of *Chaoborus*. To test for the effects of a *Chaoborus* x *Mesocyclops* interaction with untransformed data is to test whether these two absolute increments (5 µm, 20 µm) are significantly different from each other, and we find this to be the case (since we have allowed ourselves to specify standard errors to be very low).

On biological grounds, the above conclusion is unlikely to be acceptable because usually we regard the magnitude of a treatment effect to be most meaningfully and appropriately measured as percent change, not as absolute change. Because *Mesocyclops* causes a 50 *percent* increase in spine length regardless of whether *Chaoborus is* present or absent, there is, in a biological sense, *no* interaction between the effects of the two predators. And this is exactly what would be suggested by a 2-way ANOVA applied to log-transformed data. Wherever effect is best measured as percent change and there is interest in testing for factor interaction, log transformation of data would seem to be obligatory.

Biologist perpetrators of metric-interaction mismatch can be forgiven the error, for there appears to be no single clear and complete analysis of this problem in the statistical literature. We are attempting to fill the gap (Hurlbert and White, in prep.).

One-tailed Tests. — One-tailed tests were used in only four studies: Gilbert (1976b), Havel and Dodson (1984, 1987) and Murtaugh (1981) (Table 1). In none of these was any explanation offered for the use of such tests. In two cases some response variables in a given experiment were analyzed with one-tailed tests while other response variables in the same experiment were analyzed with two-tailed tests (Havel and Dodson, 1987; Murtaugh, 1981).

Possibly these authors were following advice widely prescribed in statistics books (Lombardi and Hurlbert, in prep.) to the effect that one may use or should use a 1-tailed test if one can predict or is fairly certain that the effect, if one occurs, will be in a particular direction. That is simply bad advice. Fortunately it is ignored by most zooplanktologists and other biologists most of the time. Our suspicions about the direction of potential effects are irrelevant. As stated earlier, the criterion for choosing between 1-tailed and 2-tailed tests must be whether the investigator will or is likely to simply ignore an effect if its direction is the *opposite* of that "suspected" or "predicted" beforehand (Lombardi and Hurlbert, in prep.).

Again, we cannot fault biologists when the statistical literature itself is so poor on this topic.

Other Errors. — Aside from those already discussed, we found only one other major error worth noting. Li and Li (1979) used 8 different prey species and carried out several different experiments each testing *Acanthocyclops'* relative preference for two prey species. One analysis applied to these data involved application of a χ^2- test to an 8 (prey species) x 2 (eaten/not eaten) contingency table wherein the data from the various separate experiments were lumped. The general inappropriateness of the procedure is aggravated by the fact that the number of replicate trials per prey-pair was variable but not reported. Problems also appear to exist with the other χ^2-tests presented. For example, the test of *Acanthocyclops*' preference for *Asplanchna* over all other prey combined seems to use data from trials where *Asplanchna* was not one of the prey offered.

CONCLUSIONS

The extent of serious statistical error in this particular area of experimental planktology is very great, much greater than would have been inferred from our earlier review (Hurlbert, 1984). There is no reason to suspect that this research area is an outlier in this regard. Much the same situation obtains across the full spectrum of ecology. There are other fields, such as animal behavior, where the situation is worse (Machlis et al., 1985; Kroodsma, 1989, 1990; Hurlbert and Lombardi, in prep.; Lombardi and Hurlbert, in prep.), but that should be of little consolation.

We have made no attempt to assess for individual studies, for any general hypotheses, or for the invertebrate zooplanktivore literature collectively, whether the errors detected have led to erroneous biological conclusions. In some specific instances, they certainly have done so. In many more instances they have simply exaggerated the true strength of the evidence. We suspect that the principal conclusions in this area of research, those concerning phenomena that have been studied by several investigators, have been unaffected.

The problem, then, is not that we have been led into numerous cases of serious biological error. The principal problem rather is that tremendous burdens have been created. First, there is the burden for the reader, who can no longer assume that a reported statistical analysis and *P* value are even approximately correct. Every paper must be regarded skeptically. The reader must *meticulously* examine the study design and statistical analyses and make his own judgment. If the details of these are omitted or if the reader is not capable of judging the statistical procedures then he is simply up a creek. The reported *P* values and related conclusions cannot be accepted simply because the author, journal, editor, and reviewers all seem "reputable" and the author's interpretations "reasonable."

Secondly, there is a great burden created for editors, reviewers, statisticians, and thesis advisors. The abundant incorrect analyses in the literature inevitably are taken by many scientists as models of acceptable practice. This results in the production of still more flawed manuscripts and an ever increasing workload for those charged with the largely thankless task of quality control.

Improvement of this situation requires that every investigator take responsibility for fully understanding the statistical methodologies he uses. This is central to his competence as a scientist; these are not peripheral "details" to be relegated to a statistician.

Attaining moderate statistical self-sufficiency, however, can be difficult, even for the highly motivated. Many statistics books and courses provide no information or incorrect information on important topics. For example, sampling design and experimental design are completely ignored in many statistics books. Chapters on analysis of categorical data typically fail to distinguish the concepts of evaluation unit and experimental unit. The general significance of log transformation to multi-way ANOVAs is *never* clearly treated. The list is long-and it, not incompetence or lack of interest on the part of biologists, is the primary cause of the problematic situation here documented.

ACKNOWLEDGMENTS

We thank M. Arts, I. Bayly, M. Barry, S. Dodson, J. Gilbert, J. Havel, P. Hebert, H. Li, J. Li, G. Quinn, P. Schulze, G. Sprules, C. Williamson and two anonymous reviewers for critical comments on a first draft of this article that greatly improved its accuracy. We also thank D. Soto for some early detective work, C. Lombardi for discussion of statistical problems in animal behavior, J. Rinks and the Ogden Graphics Department for assistance with illustrations, D. Hamilton for excellent typing, and P. Donaghay and G. Paffenhöffer for editorial patience. Work on this review was supported in part by a grant from the National Science Foundation (BSR-8509535).

LITERATURE CITED

Addicott, J. F. 1974. Predation and prey community structure: an experimental study of the effect of mosquito larvae on protozoan communities of pitcher plants. Ecology 55: 475-492.

Arts, M. T., E. J. Maly and M. Pasitschniak. 1981. The influence of *Acilius* (Dytiscidae) predation on *Daphnia* in a small pond. Limnol. Oceanogr. 26: 1172-1175.

Barry, M. J. and I. A. E. Bayly. 1985. Further studies on predator induction of crests in Australian *Daphnia* and the effects of crests on predation. Aust. J. Mar. Freshw. Res. 36: 519-535.

BrainPower, Inc. 1986. StatView 512+, Version 1.1. Abacus Concepts, Berkeley, California.

Brandl, Z. and C. H. Fernando. 1974. Feeding of the copepod *Acanthocyclops vernalis* on the cladoceran *Ceriodaphnia reticulata* under laboratory conditions. Can. J. Zool. 52: 99-105.

——— and ———. 1975. Food consumption and utilization by two freshwater cyclopoid copepods *(Mesocyclops edax* and *Cyclops vicinus).* Int. Rev. Ges. Hydrobiol. 60: 471-494.

——— and ———. 1978. Prey selection by the cyclopoid copepods *Mesocyclops edax* and Cyclops *vicinus*. Verh. Int. Verein. Limnol.20:2505-2510.

——— and ———. 1981. The impact of predation by cyclopoid copepods on zooplankton. Verh. Int. Verein. Limnol. 21: 1573-1577.

Confer, J. L. 1971. Intrazooplankton predation by *Mesocyclops edax* at natural prey densities. Limnol. Oceanogr. 16: 663-666.

Cooper, S. D. 1983. Selective predation on cladocerans by common pond insects. Can. J. Zoo!. 61: 879-886.

——— and C. R. Goldman. 1980. Opossum shrimp *(Mysis relicta)* predation on zooplankton. Can. J. Fish. Aquat. Sci. 37:909-919.

——— and D. W. Smith. 1982. Competition, predation and the relative abundance of two species of *Daphnia*. J. Plankt. Res. 4: 859-879.

Crowder, M. J. and D. J. Hand. 1990. Analysis of repeated measures. Chapman and Hall, New York. 257 pp.

Dodson, S. I. 1974. Zooplankton competition and predation: an experimental test of the size-efficiency hypothesis. Ecology 55: 605-613.

———. 1988a. Cyclomorphosis in *Daphnia galeata mendotae* Birge and *D. retrocurva* Forbes as a predator-induced response. Freshw. Biol. 19: 109-114.

———. 1988b. The ecological role of chemical stimuli for the zooplankton: predator avoidance behavior in *Daphnia*. Limnol. Oceanogr. 33: 1431-1439.

———. 1989. The ecological role of chemical stimuli for the zooplankton: predator-induced morphology in *Daphnia*. Oecologia 78: 361-367.

——— and S. D. Cooper. 1983. Trophic relationships of the freshwater jellyfish *Craspedacusta sowerbyi* Lankester 1880. Limnol. Oceanogr. 28: 345-351.

——— and J. E. Havel. 1988. Indirect prey effects: some morphological and life history responses of *Daphnia pulex* exposed to *Nolonecta undulata*. Limnol. Oceanogr. 33(6, part 1): 1274-1285.

Fedorenko, A. Y. 1975. Feeding characteristics and predation impact of *Chaoborus* (Diptera, Chaoboridae) larvae in a small lake. Limnol. Oceanogr. 20: 250-258.

Folt, C. L. 1987. An experimental analysis of costs and benefits of zooplankton aggregation. Pages 300-314 *in* W. C. Kerfoot and A. Sih, eds. Predation: direct and indirect impacts on aquatic communities. Univ. Press of New England, Hanover, New Hampshire.

———. J. T. Ryback and C. R. Goldman. 1982. The effect of prey composition and abundance on the predation rate and selectivity of *Mysis relicta*. Hydrobiologia 93: 133-144.

Gilbert, J. J. 1966. Rotifer ecology and embryological induction. Science 151: 1234-1237.

———. 1967. *Asplanchna* and postero-lateral spine production in *Brachionus calyciflorus*. Arch. Hydrobiol. 64: 1-62.

———. 1973a. The adaptive significance of polymorphism in the rotifer *Asplanchna*. Humps in males and females. Oecologia 13: 135-146.

———. 1973b. Induction and ecological significance of gigantism in the rotifer *Asplanchna sieboldi*. Science 181: 63-66.

———. 1975. Polymorphism and sexuality in the rotifer *Asp/anchna*, with special reference to the effects of prey type and clonal variation. Arch. Hydrobiol. 75: 442-483.

1976a. Selective cannibalism in the rotifer *Asplanchna sieboldi:* contact recognition of morphotype and clone. Proc. Natl. Acad. Sci. USA 73: 3233-3237.

———. 1976b. Polymorphism in the rotifer *Asplanchna sieboldi:* biomass, growth, and reproductive rate of the saccate and companulate morphotypes. Ecology 57: 542-551.

———. 1976c. Sex-specific cannibalism in the rotifer *Asplanchna sieboldi*. Science 194: 730-732.

———. 1977a. Defenses of males against cannibalism in the rotifer *Asplanchna:* size, shape, and failure to elicit tactile feeding responses. Ecology 58: 1128-1135.

———. 1977b. Control of feeding behaviour and selective cannibalism in the rotifer *Asplanchna*. Freshw. Biol. 7: 337-341.

———. 1988. Susceptibilities of ten rotifer species to interference from *Daphnia pulex*. Ecology 69: 1826-1838.

———. 1989. The effect of *Daphnia* interference on a natural rotifer and ciliate community: short-term bottle experiments. Limnol. Oceanogr. 34: 606-617.

——— and K. L. Kirk. 1988. Escape response of the rotifer *Keratella:* description, stimulation, fluid dynamics, and ecological significance. Limnol. Oceanogr. 33: 1440-1450.

——— and R. S. Stemberger. 1984. *Asplanchna*-induced polymorphism in the rotifer *Keratella slacki*. Limnol. Oceanogr. 29: 1309-1316.

——— and C. E. Williamson. 1979. Predator-prey behavior and its effect on rotifer survival in associations of *Mesocyclops edax, Asplanchna girodi*, *Polyarthra vulgaris*, and *Keratella cochlearis.* Oecologia 37: 13-22.

Glantz, S. A. 1980. Biostatistics: how to detect, correct and prevent errors in the medical literature. Circulation 61: 1-7.

Grant, J. W. G. and I. A. E. Bayly. 1981. Predator induction of crests in morphs of the *Daphnia carinata* King complex. Limnol. Oceanogr. 26: 201-218.

Hanazato, T. 1990. Induction of helmet development by a *Chaoborus factor* in *Daphnia ambigua* during juvenile stages. J. Plankt. Res. 12: 1287-1294.

——— and M. Yasuno. 1989. Zooplankton community structure driven by vertebrate and invertebrate predators. Oecologia 81: 450-458.

——— and ———. 1990. Influence of *Chaoborus* density on the effects of an insecticide on zooplankton communities in ponds. Hydrobiologia 194: 183-197.

Havel, J. E. 1985a. Cyclomorphosis of *Daphnia pulex* spined morphs. Limnol. Oceanogr. 30: 853861.

———. 1985b. Predation of common invertebrate predators on long- and short-featured *Daphnia retrocurva*. Hydrobiologia 124: 141-149.

——— and S. I. Dodson. 1984. *Chaoborus predation* on typical and spined morphs of *Daphnia pu/ex:* behavioral observations. Limnol. Oceanogr. 29: 487-494.

——— and ———. 1987. Reproductive costs of *Chaoborus*-induced polymorphism in Daphnia *pulex*. Hydrobiologia 150: 273-281.

Havens, K. E. 1990. *Chaoborus* predation and zooplankton community structure in a rotifer-dominated lake. Hydrobiologia 198: 215-226.

Hebert, P. D. N. and P. M. Grewe. 1985. *Chaoborus*-induced shifts in the morphology of *Daphnia ambigua*. Limnol. Oceanogr. 30: 1291-1297.

——— and J. M. Loaring. 1980. Selective predation and the species composition of Arctic ponds. Can. J. Zool. 58: 422-426.

Hewett, S. W. 1980a. Prey-dependent cell size in a protozoan predator. J. Protozool. 27: 311-313.

———. 1980b. The effect of prey size on the functional and numerical responses of a protozoan predator to its prey. Ecology 61: 1075-1081.

———. 1988. Predation by *Didinium nasutum*: Effects of predator and prey size. Ecology 69: 135-45.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54:187-211.

Janicki, A. and J. DeCosta. 1990. An analysis of prey selection by *Mesocyc/ops edax*. Hydrobiologia 198:133-139.

Kerfoot, W. C. 1975. The divergence of adjacent populations. Ecology 56: 1298-1313.

———. 1977. Implications of copepod predation. Limnol. Oceanogr. 22: 316-325.

———. 1987. Translocation experiments: *Bosmina* responses to copepod predation. Ecology 68: 596-610.

——— and C. Peterson. 1980. Predatory copepods and *Bosmina:* replacement cycles and further influences of predation upon prey reproduction. Ecology 61: 417-431.

Kirk, K. L. and J. J. Gilbert. 1988. Escape behavior of *Polyarthra* in response to artificial flow stimuli. Bull. Mar. Sci. 43: 551-560.

Kroodsma, D. E. 1989. Suggested experimental designs for song playbacks. An. Behav. 37: 600609.

———. 1990. How the mismatch between experimental design and the intended hypothesis limits confidence in knowledge, as illustrated by an example from bird-song dialects. Pages 226-245 *in* M. Bekoff and D. Jamieson, eds. Interpretation and explanation in the study of animal behavior Vol. II. Westview Press, Boulder, Colorado.

Krueger, D. A. and S. I. Dodson. 1981. Embryological induction and predation ecology in *Daphnia* pulex. Limnol. Oceanogr. 26: 219-223.

Li, I. L. and H. W. Li. 1979. Species-specific factors affecting predator-prey interactions of the copepod *Acanthocyclops vernalis* with its natural prey. Limnol. Oceanogr. 24: 613-626.

Luecke, C. and W. J. O'Brien. 1983. The effect of *Heterocope* predation on zooplankton communities in arctic ponds. Limnol. Oceanogr. 28: 367-377.

Lunte, C. C. and C. Luecke. 1990. Trophic interactions of *Leptodora* in Lake Mendota. Limnol. Oceanogr. 35: 1091-1 100.

Machlis, L., P. W. D. Dodd and J. C. Fentress. 1985. The pooling fallacy: problems arising when individuals contribute more than one observation to the data set. Z. Tierpsychol. 68: 201-214.

McQueen, D. J. 1969. Reduction of zooplankton standing stocks by predaceous *Cyclops bicuspidatus thomasi* in Marion Lake, British Columbia. J. Fish. Res. Bd. Can. 26: 1605-1618.

Mead, R. 1988. The design of experiments. Cambridge Univ. Press, New York. 620 pp.

Milliken, G. A. and D. E. Johnson. 1984. Analysis of messy data, Vol. 1: designed experiments. Van Nostrand Reinhold, New York. 473 pp.

Moore, M. V. 1988. Density-dependent predation of early instar *Chaoborus* feeding on multispecies prey assemblages. Limnol. Oceanogr. 33: 256-268.

Murdoch, W. W. and M. A. Scott. 1984. Stability and extinction of laboratory populations of zooplankton preyed on by the backswimmer *Notonecta*. Ecology 65: 1231-1248.

———,——— and P. Ebsworth. 1984. Effects of the general predator, *Notonecta* (Hemiptera), upon a freshwater community. J. An. Ecol. 53: 791-808.

Murtaugh, P. A. 1981. Selective predation by *Neomysis mercedis* in Lake Washington. Limnol. Oceanogr. 26: 445-453.

Neill, W. E. 1981. Impact of *Chaoborus* predation upon the structure and dynamics of a crustacean zooplankton community. Oecologia 48: 164-177.

———. 1984. Regulation of rotifer densities by crustacean zooplankton in an oligotrophic montane lake in British Columbia. Oecologia 61: 175-181.

Nero, R. W. and W. G. Sprules. 1986. Predation by three glacial opportunists on natural zooplankton communities. Can. J. Zool. 64: 57-64.

O'Brien, W. J. and D. Schmidt. 1979. Arctic *Bosmina* morphology and copepod predation. Limnol. Oceanogr. 24: 564-568.

——— and G. L. Vinyard. 1978. Polymorphism and predation: the effect of invertebrate predation on the distribution of two varieties of *Daphnia carinata* in South India ponds. Limnol. Oceanogr. 23:452-460.

———, D. Kettle and H. Riessen. 1979. Helmets and invisible armor: structures reducing predation from tactile and visual planktivores. Ecology 60: 287-294.

O'Fallon, J. R., S. D. Dubey, D. S. Salsburg, J. H. Edmondson, A. Soffer and T. Colton. 1978. Should there be statistical guidelines for medical research papers? Biometrics 34: 687-695.

Pastorok, R. A. 1980. The effects of predator hunger and food abundance on prey selection by *Chaoborus* larvae. Limnol. Oceanogr. 25: 910-921.

Peacock, A. 1982. Responses of Cyclops *bicuspidatus thomasi* to alterations in food and predators. Can. J. Zool. 60: 1446-1462.

——— and W. J. P. Smyly. 1983. Experimental studies on the factors limiting *Tropocyclopsprasinus* (Fischer) 1860 in an oligotrophic lake. Can. J. Zool. 61: 250-265.

Riessen, J. P., J. W. Sommerville, C. Chiappari and D. Gustafson. 1988. *Chaoborus* predation, prey vulnerability, and their effect in zooplankton communities. Can. J. Fish. Aq. Sci. 45: 1912-1920.

Salt, G. W. 1974. Predator and prey densities as controls of the rate of capture by the predator *Didinium nasutum*. Ecology 55: 434-439.

Schor, S. and I. Karten. 1966. Statistical evaluations of medical journal manuscripts. Amer. Med. Assoc. 195: 1123-1128.

Schulze, P. C. and C. L. Folt. 1989. Effects of conspecifics and phytoplankton on predation rates of the omnivorous copepods *Epischura lacustris* and *Epischura nordenskioldi*. Limnol. Oceanogr. 34:444-450.

——— and ———. 1990. Food resources, survivorship, and reproduction of the omnivorous copepod *Epischura lacustris*. Ecology 71: 2224-2240.

Scott, M. A. and W. W. Murdoch. 1983. Selective predation by the backswimmer, *Notonecta*. Limnol. Oceanogr. 28: 352-366.

Soto, D. 1985. Experimental evaluation of copepod interactions. Verh. Int. Verein. Limnol. 22: 3199-3204.

Sprules, W. G. 1972. Effects of size-selective predation and food competition on high altitude zooplankton communities. Ecology 53: 375-386.

Stemberger, R. S. 1985. Prey selection by the copepod *Diacyclops thomasi*. Oecologia 65: 492-497.

——— and J. J. Gilbert. 1984. Spine development in the rotifer *Keratella cochlearis:* induction by cyclopoid copepods and *Asplanchna*. Freshw. Biol. 14: 639-647.

——— and ———. 1987. Multiple-species inductions of morphological defenses in the rotifer *Keratella testudo*. Ecology 68: 370-372.

Stenson, J. A. E. 1987. Variation in capsule size of *Holopedium gibberum* (Zaddach): a response to invertebratepredation.Ecology68:928-4.

Underwood, A. J. 1981. Techniques of analysis of variance in experimental marine biology and ecology.Oceanogr.Mar.Biol.Annu.Rev.19:513-65.

Urquhart, N. S. 1981. The anatomy of a study. HortScience 16: 621-627.

Vanni, M. J. 1988. Freshwater zooplankton community structure: introduction of large invertebrate predators and large herbivores to a small-species community. Can. J. Fish. Aquat. Sci. 45: 1758-1770.

Vinyard, G. L. and R. A. Menger. 1980. *Chaoborus americanus* predation on various zooplankters: functional response and behavioral observations. Oecologia 45:90-93.

Von Ende, C. M. and D. 0. Dempsey. 1981. Apparent exclusion of the cladoceran *Bosmina lonigrostris* by invertebrate predator *Chaoborus americanus*. Amer. MidI. Nat. 105: 240-248.

Vuorinen, I., M. Ketola and M. Walls. 1989. Defensive spine formation in *Daphnia pulex* Leydig and induction by *Chaoborus crystallinus* Dc Geer. Limnol. Oceanogr. 34: 245-248.

Walls, M. and M. Ketola. 1989. Effects of predator-induced spines on individual fitness in *Daphnia* pulex.Limnol.Oceanogr.34:390-36.

Williamson, C. E. 1980. The predatory behavior of *Mesocyclops edax:* predator preferences, prey defenses and starvation induced changes. Limnol. Oceanogr. 25: 903-909.

———. 1983. Behavioral interactions between a cyclopoid copepod predator and its prey. J. Plankt. Res. 5:701-711.

———. 1984. Laboratory and field experiments on the feeding ecology of the cyclopoid copepod, *Mesocyclops edax*. Freshw. Biol. 14: 575-585.

———. 1987. Predatory-prey interactions between omnivorous diaptomid copepods and rotifers: the role of prey morphology and behavior. Limnol. Oceanogr. 32: 167-177.

———. and N. M. Butler. 1986. Predation on rotifers by a suspension-feeding calanoid copepod *Diaptomus pallidus* .Limnol. Oceanogr.31:393-402.

Wong, C. K. 1981. Predatory feeding behavior of *Epischura lacustris* (Copepoda, Calanoida) and prey defense. Can. J. Fish. Aquat. Sd. 38: 275-279.

———. C. W. Ramcharan and W. G. Sprules. 1986. Behavioral responses of a herbivorous calanoid copepod to the presence of other zooplankton. Can. J. Zool. 64: 1422-1425.

DATE ACCEPTED: March 3. 1993.

ADDRESSES: *(S.H.H.) Department of Biology, San Diego State University, San Diego, California 92182; (M.D.W.) Ogden Environmental and Energy Services Company, 5510 Morehouse Drive, San Diego, California 92121.*