Multispecies Toxicity Tests Using Indigenous Organisms: Predicting the Effects of Hazardous Materials in Streams

by

Kurt Walter Pontasch

Dissertation submitted to the Graduate Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Zoology

APPROVED:

John Cairns, Jr., Chairman

E. Fred Benfield

Robert A. Paterson

Eric P. Smith

Bruce Wallace

May, 1988

Blacksburg, Virginia

MULTISPECIES TOXICITY TESTS USING INDIGENOUS ORGANISMS: PREDICTING THE EFFECTS OF HAZARDOUS MATERIALS IN STREAMS

by

Kurt W. Pontasch

John Cairns, Jr., Chairman

Biology

(ABSTRACT)

The purpose of the investigation presented in chapter 1 was to determine which of the following artificial stream designs would be most logistically simple yet effective in maintaining riffle insects during a 30-d bioassay: 1) static and no current (S-NC); 2) flow-through and no current (FT-NC); 3) static with current (S-C); or 4) flow-through with current (FT-C). Flow-through and current, when provided, were 12 ml min^{-1} and 30 cm sec^{-1}, respectively. Streams were covered by emergence traps, and daylight equivalent lights provided a natural photoperiod. The four stream designs were evaluated in triplicate based on changes in insect species-abundances after 30 d. Test organisms were transferred to the artificial streams in rock-filled containers previously colonized for 30 d in a third-order mountain stream riffle. Relative to benthic samples taken directly from the source riffle, the artificial substrates selected for collector-filterers and against collector-gatherers. The FT-C and S-C stream designs maintained most taxa at or above initial densities. Emergent adults comprised a large proportion of mayfly and chironomid densities and must be monitored during bioassays with aquatic insects.

The investigation reported in chapter 2 was conducted to determine if contaminant-induced changes in macroinvertebrate and periphyton communities in laboratory stream microcosms could be used to predict macroinvertebrate and periphyton responses in a natural stream receiving the same contaminant. The microcosms were dosed in quadruplicate with four (0.0, 0.1, 1.0 and 10.0%) concentrations of a complex effluent; these concentrations reflected those in the field. Mayfly densities in the microcosms were significantly ($P \leq 0.05$) reduced at 1.0 or 10.0% effluent depending on species. Hydropsychids were not affected by the effluent, and chironomids and

periphyton were stimulated. Overall, the stream microcosms accurately predicted the macroinvertebrate and periphyton response observed in the field.

Chapter 3 compared responses to a complex effluent from microcosms of indigenous macroinvertebrates and protozoans to responses observed in acute tests with *Daphnia magna*, *Ceriodaphnia dubia* and *Pimephales promelas* and chronic survival and reproductive tests with *C. dubia* The predictive utility of these various tests was then evaluated against observed effects in the receiving stream. The LC_{50}s (% effluent) from the acute tests were 63.09 for *Pimephales promelas*, 18.8 to 31.3 for *Daphnia magna* and 54.7 for *Ceriodaphnia dubia*. Results from 7-day chronic tests indicated that *C. dubia* survival was significantly ($P \leq 0.05$) affected at 30% effluent and reproduction was affected at concentrations $\geq 3.0\%$ effluent. In the protozoan microcosms, community composition was significantly ($P \leq 0.05$) changed at 1.0%; while protozoan species richness was significantly reduced at 3.0% effluent. The microcosms not only were the most sensitive indicators of effluent toxicity, they also correctly predicted which indigenous organisms would be lost and which would be stimulated at various ambient concentrations of the effluent.

In the fourth chapter canonical discriminant analysis, 2 diversity indices, and 7 community comparison indices were evaluated to determine their utility in quantifying macroinvertebrate response to a complex effluent in laboratory microcosms. A permutation and randomization procedure was used to test the hypothesis of no treatment effect based on the community comparison indices. The Bray-Curtis index provided the most meaningful condensation of the data.

Acknowledgements

This reasearch was supported in part by the E. I. du Pont Educational Foundation, the Mobil Foundation Incorporated, and the Virginia Academy of Science. Without the financial support of these organizations the research reported in this dissertation could not have been accomplished.

My deepest appreciation is extended to Dr. John Cairns, Jr., my major professor. In addition to invaluable scientific guidance, Dr. Cairns has provided great insights into some of the more political aspects of academe, regulatory agencies, and industry. For a critical review of the dissertation, I would like to thank Dr. Fred Benfield, Dr. Robert Paterson, Dr. Eric Smith, and Dr. Bruce Wallace. Dr. Smith's expertise was especially useful in the investigation reported in chapter 4.

B. R. Niederlehner collaborated with me in conducting the research reported in chapter 3. She conducted the protozoan microcosm tests, and all tests with *Ceriodaphnia dubia*. My appreciation is also extended to Scott Miller and Carol Beth Kniskern who provided invaluable assistance in the field and laboratory. Both of these individuals were willing to work long hours during critical research periods.

Lastly, but most importantly, I would like to thank all my family members whose encouragement and support allowed me to complete this degree. I particularly owe a special thanks to my wife and children, and , for their unwavering love, support and understanding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PROLOGUE

Ideology

Managing the hazards of stressors in aquatic ecosystems requires predictive techniques that are cost effective and accurate. Accurate predictive tools will allow us to avoid the costs of underprotection of the environment (i.e., damage to environmental resources) and the costs of overprotection (i.e., construction of waste treatment facilities that provide no biological benefit), while inaccurate predictions will engender these costs. An obvious way to determine if our predictions of hazard are sufficient to the task of cost-effective management is by means of validation studies which compare predictions of environmental safety derived for management purposes to observed effects in natural systems. We can then identify the margins of error of our predictive tools and judge whether the accuracy is sufficient in light of costs of errors.

Although toxicity testing has been advocated as a tool for predicting environmental safety and harm since at least the 1940's (Hart *et al.* 1945), management strategies for many years ignored biological response as an assessment tool. In more recent years, toxicity test data finally has been incorporated into the regulatory framework, and this has resulted in an improved control of toxics.

Short-term, single-species toxicity tests were initially chosen to identify, regulate, and monitor environmental problems, primarily due to familiarity with these tests and reports of their ease and low cost (USEPA 1982ab; 1985). However, in the years between the development of single-species tests and their formal application to management problems, appreciation for the complexity of intact ecosystems has grown, and exclusive reliance on single-species toxicity tests for predictions of environmental safety and harm has been criticized (e.g., NAS 1981; Cairns 1983; Odum 1984; Kimball and Levin 1985). Although single-species tests can efficiently examine relative toxicity, they may not be the most accurate or efficient method for predicting responses in contaminated ecosystems because single-species tests: 1) do not take into account interactions among species; 2) utilize genetically homogeneous laboratory-stock test populations

which may lack the adaptive capability of heterogenous natural populations; 3) utilize species of unknown relative sensitivity; 4) are conducted under physical and chemical conditions which lack similarity to natural habitats; and 5) utilize species which are usually not indigenous to the receiving ecosystem, thus complicating field-validation of the test.

Studies of the effectiveness of short-term, single-species biological testing of effluents under the EPA's Complex Effluent Toxicity Testing Program (as summarized by Wall and Hanmer 1987) have found correlations between response of single species in laboratory tests and certain indices of field response. However, single-species tests are often overprotective when large application factors are employed, and accurate predictions of how much and what kind of damage to ecosystem structure and function to expect are not possible. Greater predictive accuracy is necessary in many management applications such as premanufacture licensing where a confirmatory field survey is not possible and/or where costs of inaccurate predictions would be unacceptable.

Another problem with the field-validation studies conducted is that attributes necessary for ecosystem function (e.g., metabolism, colonization, energy flow) were not examined and impairments in these functions go undetected. Functional measures are necessary to satisfy an expectation that whole ecosystems are being protected. For example, Cairns *et al.* (1986) found that the colonization of protozoans was impaired by concentrations of cadmium that had no effect on the source pool of species acting as colonizers. Because colonization must occur continuously in aquatic systems (e.g., each time a leaf falls in or high water scours the substrate) most ecologists would consider colonization to be an important ecological process and would expect impairments to parlay into impaired succession, energy flow, etc. While single-species tests might indicate no toxicity, impaired colonization might result in a delayed effect on species richness. Single-species tests might also indicate no toxicity, whereas contaminant-induced macroinvertebrate drift might result in a decrease in macroinvertebrate species richness below an effluent outfall.

Thus, whether current assessments of single-species toxicity tests constitute adequate validation of their use as predictors of ecosystem effects, as suggested by Wall and Hanmer (1987), depends on hotly disputed judgements of what is sufficient accuracy in predictions of environmental safety, and what ecosystem properties and behaviors are the essential objects of environmental protection. It is unlikely that a majority of ecologists would be satisfied that entire ecosystems were, in fact, being protected based on 7 day tests with three species.

In view of the inadequacies of single-species toxicity tests, many ecologists believe that predictions of environmental safety and harm could be substantially improved by incorporating tests that directly examine those properties of ecosystems that are the object of protective legislation (e.g., species richness). Such multispecies toxicity tests are being developed (see Giesy 1980; Hammons 1981; Taub 1980; Cairns 1985; Cairns 1986). Early work suggests that they are not only more realistic models of exposure than single-species toxicity tests but also no more expensive than (Perez and Morrison 1985, Niederlehner *et al.* 1986) and equally replicable as (Geisy and Allred 1985) as conventional tests. Nevertheless, multispecies toxicity tests have not been generally accepted in the regulatory framework.

Approach

The overall goal of my research was to develop a multispecies toxicity testing protocol using microcosms of macroinvertebrates and periphyton. Periphyton was used because: 1) it is involved in both primary production and decomposition, and should therefore reflect toxicant effects on these important ecosystem functions; and 2) periphyton is an important source of food for many macroinvertebrates. Macroinvertebrates were used because: 1) they serve as intermediates in the trophic chain between microbial species and important vertebrate species such as fish and waterfowl; and 2) changes in macroinvertebrate community structure are widely used in environmental monitoring and surveillance.

The primary focus of my research was on changes in aquatic insect species-abundances in response to stressors. In most streams and rivers aquatic insects constitute the majority of both macroinvertebrate species and individuals. Aquatic insects exhibit a wide range of responses to toxicants, possibly because of their morphological, physiological, and ecological diversity. However, aquatic insects are used relatively infrequently as bioassay organisms, because of logistical problems in collecting and maintaining many species.

Care must be taken when developing toxicity testing protocols which use aquatic macroinvertebrates as test organisms. The sensitivity of aquatic invertebrates to toxicants varies during the life cycle (Rosenberg 1972; Czyzewska 1976; and Armstrong *et al.* 1976), and invertebrates are often most sensitive during molting (Emery 1970; Swedmark *et al.* 1971; and Armstrong *et al.*; 1976). In addition, aquatic insects tend to exhibit increased sensitivity with longer exposures (Clubb *et al.* 1975; Maki *et al.* 1975; and Federle and Collins 1976), Therefore, a 30 day exposure period was used in my research to ensure that most species would undergo one or more molts during testing, and that some species would have undergone entire life cycles during the period of exposure. However, long-term testing with aquatic insects requires monitoring of adult emergence so that losses due to emergence are not attributed to toxicant-induced mortality. These and other factors must be accounted for in developing toxicity testing protocols using aquatic macroinvertebrates as test organisms.

To evaluate the predictive accuracy of toxicity testing protocols, validation studies which compare predictions to observed effects in natural systems must be conducted simultaneously with the laboratory tests. A major hurdle in the development of the toxicity testing protocol presented in this dissertation was finding a suitable receiving system in which to validate the protocol. Several receiving systems for possible use in validating the protocol were evaluated. Factors considered in evaluating receiving systems included: 1) presence of similar morphometric characteristics (e.g., substrate, flow, depth) at all (upstream and downstream) intensive study sites; 2) presence of a relatively unperturbed reference reach upstream from the point of impact; 3) presence of a con-

sistent impact; and 4) cooperation from the discharging company. Although several promising sites were identified, cooperation from only one discharging company could be secured. That company cooperated with the understanding that they would not be identified. The process of finding and negotiating access to a suitable site was damn educational experience.

Most toxic compounds are discharged into lotic ecosystems. Water quality below effluent outfalls is often assessed by censusing the number and kinds of organisms present. A major limitation of this procedure is an inability to control the experimental conditions of exposure. Outdoor stream mesocosms have been used in multispecies toxicity testing for some time in an attempt to better control the dosage received by the communities being tested (e.g., Burks and Wilhm 1977; Giesy *et al* 1979; Stout and Cooper 1983; and Allard and Moreau 1987). However, such mesocosms are often costly to construct, impossible to transport, and are susceptible to meteorologic and biological perturbations during testing. In contrast, laboratory stream microcosms can maintain more constant conditions of exposure (e.g., dosage, illumination, temperature, and current velocity), are not as prone to vandalism, and reduce variability in results caused by adult insect emergence from experimental channels, and adult insect oviposition into experimental channels. However, because of logistical problems, development of laboratory-based toxicity tests utilizing stream microcosms has lagged behind multispecies tests which utilize outdoor stream mesocosms, outdoor lacustrine mesocosms, or laboratory-based lacustrine microcosms. Therefore, time, money, and manpower were initially devoted to developing an inexpensive yet environmentally realistic and replicable system of portable artificial stream microcosms.

Chapter 1 addresses some of the requirements for establishing and maintaining evironmentally realistic laboratory-based stream microcosms, and the degree to which such microcosms reflect the natural communities from which they were derived.

> OBJECTIVE I. *Determine which artificial stream design (static without current, flow-through without current, static with current, or flow-through with current) is most logistically simple yet effective in maintaining microcosms of stream macroinvertebrates.*

OBJECTIVE II. *Determine if macroinvertebrate communities maintained in artificial stream microcosms for 30 days reflect macroinvertebrate communities in the natural stream system from which the microcosms were derived.*

In designing and testing the laboratory stream microcosms, I attempted to infuse some "ecology" into ecotoxicology. Too often in the past, workers have not considered the ecology of test organisms in designing, conducting, and interpreting ecotoxicological research. In some cases, stream-dwelling organisms have been tested in static aquaria, adult insect emergence has not been monitored, behavioral drift responses to contaminants have not been considered, changes in tolerance during a life cycle have not been accounted for etc. Chapter 1 considers the importance of some of these ecological parameters, and the feasibility of including these parameters in ecotoxicological testing.

The goal of the research presented in Chapter 2 was to compare environmental safety assessments from multispecies toxicity tests on laboratory microcosms of periphyton and macroinvertebrates to the response of a natural system receiving the same waste.

OBJECTIVE III. *Determine if contaminant-induced changes in macroinvertebrate and periphyton community composition in laboratry stream microcosms can be used to predict macroinvertebrate and periphyton responses in a natural stream system receiving the same contaminant.*

A major component of the research reported in Chapter 2 was finding a discharging industry that would cooperate with the investigaton. Although several suitable sites were found, only one industry would cooperate, and they did so with the understanding that they would not be identified. Communities present in the microcosms were very similar to those present in the natural receiving system. Therefore, field validation of the multispecies tests was more meaningful than past validations of single-species tests.

Standard single-species tests were conducted simultaneously with the multispecies tests, thus allowing comparisons of multispecies microcosm tests to other more conventional tests (Chapter 3).

OBJECTIVE IV. *Compare single-species, microcosm, and field responses to a complex effluent*

The comparisons reported in Chapter 3 are rarely made, but are important to the incorporation of multispecies toxicity tests in the regulatory framework and the development

of a hierarchy of toxicty testing protocols. In addition to comparisons of predictive accuracy, costs, etc., Chapter 3 also suggests applications where single-species testing may be more appropriate than multispecies testing, and vice-versa.

One of the principle arguments against the use of multispecies toxicity tests is that they are harder to interpret than single-species tests (Loewengart and Maki 1985). Chapter 4 evaluates some mathematical and statistical techniques for handling the large data sets generated during the course of multispecies toxicity testing.

> OBJECTIVE V. *Evaluate the utility of selected diversity indices, community comparison indices and multivariate statistical techniques in quantifying field and microcosm responses of macroinvertebrates to a complex effluent.*

Relatively simple techniques which provide unambiguous interpretations of results must be developed before multispecies tests will be accepted in the regulatory framework.

References

Allard, M. and G. Moreau. 1987. Effects of experimental acidification on a lotic macroinvertebrate community. Hydrobiologia 144:37-49.

Armstrong, D.A., D.V. Buchanon, M.A. Mallon, R.S. Caldwell, and R.E. Milleman. 1976. Toxicity of the insecticide methoxychlor to the dungeness crab *Cancer magister*. Marine Biology 38:239-252.

Burks, S.L. and J.L. Wilhm. 1977. Bioassays with a natural assemblage of benthic macroinvertebrates. In. [F.L. Mayer and J.L. Hamelink eds.] Aquatic Toxicology and Hazard Evaluation, ASTM STP 634, American Society for Testing and Materials.

Cairns, J., Jr. (editor) 1986. Community toxicity testing. ASTM STP 920, American Society for Testing and Materials. Philadelphia. 350 p.

Cairns, J., Jr., J.R. Pratt, B.R. Niederlehner, and P.V. McCormick. 1986. A simple cost-effective multispecies toxicity test using organisms with a cosmopolitan distribution. Environmental Monitoring and Assessment 6:207-220.

Cairns, J., Jr. 1985. Multispecies toxicity testing. Pergamon Press, 261 p.

Cairns, J., Jr. 1983. Are single species toxicity tests alone adequate for estimating environmental hazard? Hydrobiologia 100:47-57.

Clubb, R.W., A.R. Gaufin, and J.L. Lord. 1975. Acute cadmium toxicity studies upon nine species of aquatic insects. Environmental Research 9:332-341.

Czyzewska, K., 1976. The effect of detergents on larval development of a crab. Marine Pollution Bulletin 7:108-112.

Emery, R.M. 1970. The comparative acute toxicity of cresol to two benthic crustaceans. Water Research 4:485-491.

Federle, P.J. and W.J. Collins 1976. Insecticide toxicity to three insects from Ohio ponds. Ohio Journal of Science 76:19-24.

Giesy, J.P. 1980. Microcosms in ecological research. U.S. Department of Energy. 1110 p.

Giesy, J.P., Jr. H.J. Kania, J.W. Bowling, R.I. Knight, S. Mashburn, and S. Clarkin. 1979. Fate and biological effects of cadmium introduced into channel microcosms. EPA-6002/3-79-039. USEPA, Athens, GA.

Giesy, J.P. and P.M. Allred. 1985. Replicability of aquatic multispecies test systems. In. [J. Cairns, Jr. Ed.] Multispecies toxicity testing. Pergamon Press. 261 p.

Hammons, A.S. 1981. Methods for ecological toxicology: a critical review of laboratory multispecies tests. Ann Arbor Science Publ.. 307 p.

Hart, W.B., P. Doudoroff, and J. Greenback. 1945. The evolution of the toxicity of industrial wastes, chemicals and other substances to freshwater fishes. Waste Cont. Lab., Atlantic Refining Co., Philadelphia, PA., 349 p.

Kimball, K.D. and S.A. Levin. 1985. Limitations of laboratory bioassays: the need for ecosystem level testing. Bioscience 35:165-171.

Loewengart G., and A.W. Maki. 1985. Multispecies toxicity tests in the safety assessment of chemicals: necessity or curiosity?. In. [J. Cairns, Jr. Ed.] Multispecies toxicity testing. Pergamon Press. 261 p.

Maki, A.W., L. Geissel and H.E. Johnson. 1975. Comparative toxicity of larval lampricide (TFM: 3-triflouromethyl-4-nitrophenol) to selected benthic macroinvertebrates. Journal of the Fisheries Research Board of Canada 32:1455-1459.

National Academy of Sciences. 1981. Testing for effects of chemicals on ecosystems. National Academy Press, Washington, DC, 98 p.

Niederlehner, B.R., J.R. Pratt, A.L. Buikema, Jr. and J. Cairns, Jr. 1986. Comparisons of estimates of hazard derived at three levels of complexity. In. [J. Cairns, Jr. Ed.] Community toxicity testing. ASTM STP 920, American Society for Testing and Materials. Philadelphia. 350 p.

Odum, E.P. 1984. The mesocosm. Bioscience 34:558-562.

Perez, K.T. and G.E. Morrison. 1985. Environmental assessments from simple test systems and a microcosm: Comparison of monetary costs. In. [J. Cairns, Jr. Ed.] Multispecies toxicity testing. Pergamon Press. 261 p.

Rosenberg, R. 1972. Effects of chlorinated aliphatic hydrocarbons on larval and juvenile *Balanus balanoides* (L.). Environmental Pollution 3:313-318.

Stout, R.J., and W.E. Cooper. 1983. Effect of p-cresol on leaf decomposition and invertebrate colonization in experimental outdoor streams. Can J. Fish. Aquat. Sci. 40:1647-1657.

Swedmark, M., B. Braaten, E. Emanualson, and A. Granmo. 1971. Biological effects of surface active agents on marine mammals. Marine Biology 9:183-201.

Taub, F.B., and M.E. Crow. 1980. Synthesizing aquatic microcosms. In [J.P. Giesy, Jr., Ed.] Microcosms in ecological research. U.S. Dept. of Energy. 1110 p.

U.S. Environmental Protection Agency. 1982a. Chemical fate test guidelines. Accession no. PB82-233008. Technical Information Service, Springfield, VA.

U.S. Environmental Protection Agency. 1982b. Environmental effects test guidelines. Accession no. PB82-232992. Technical Information Service, Springfield, VA.

U.S. Environmental Protection Agency. 1985. Methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. Draft Document. Environmental Monitoring and Support Laboratory, Cincinnati, OH. EPA/60 0/4-85/014, 129 p.

Wall, T.M. and R.W. Hanmer. 1987. Biological testing to control toxic water pollutants. Journal Water Pollution Control Federation 59:7-12.

CHAPTER ONE

ESTABLISHING AND MAINTAINING LABORATORY-BASED MICROCOSMS OF RIFFLE INSECT COMMUNITIES: THEIR POTENTIAL FOR MULITSPECIES TOXICITY TESTS

Abstract

The purpose of this study was to determine which of the following artificial stream designs would be most logistically simple yet effective in maintaining riffle insects during a 30-d bioassay: 1) static and no current (S-NC); 2) flow-through and no current (FT-NC); 3) static with current (S-C); or 4) flow-through with current (FT-C). Flow-through and current, when provided, were 12 ml min^{-1} and 30 cm sec^{-1}, respectively. Streams were covered by emergence traps, and daylight equivalent lights provided a natural photoperiod. The four stream designs were evaluated in triplicate based on changes in insect species-abundances after 30 d. Test organisms were transferred to the artificial streams in rock-filled containers previously colonized for 30 d in a third-order mountain stream riffle. Additional colonized substrates were sampled immediately to provide an estimate of initial densities placed in the artificial streams. Hess samples were taken directly from the source riffle to evaluate how well the artificial substrates reflected species-abundances on the natural substrate. Adults were collected from the artificial streams every 48-72 h to determine the percentage of initial densities that emerged. After 30 d (7Aug-6 Sept 1986), all organisms remaining in the streams were censused. Designs were evaluated using combined densities of adults and young.

Relative to benthic samples taken directly from the source riffle, the artificial substrates selected for collector-filterers and against collector-gatherers. The FT-C and S-C stream designs maintained most taxa at or above initial densities, and even in the FT-NC and S-NC streams densities of some taxa were not significantly different ($P \leq 0.05$) from initial densities. Emergent adults comprised a large proportion of mayfly and chironomid densities and must be monitored during bioassays with aquatic insects.

These results indicate that microcosms of riffle insect communities can be maintained for at least 30 d with moderate current and minimal flow-through.

Introduction

Microcosms have been used for some time in studying the fundamental properties and processes of ecosystems (see Giesy, 1980). More recently, research on new protocols for assessing chemical hazards in ecosystems has moved toward the development of multispecies microcosm toxicity tests that take into account interspecific interactions, test a broad range of indigenous species, and are conducted under more environmentally realistic conditions (Giesy, 1980; Cairns, 1985, 1986). Effects of anthropogenic stressors observed in the microcosms may be extrapolated to predict the effect of the same stressors in the natural system. However, in order to make such extrapolations the microcosm should reflect the abiotic and biotic properties and processes of the source ecosystem.

Most industrial wastes are discharged into lotic ecosystems, and water quality below effluent outfalls is often assessed by censusing the number and kinds of organisms present. A major limitation of this technique is the inability to control the experimental conditions of exposure. Outdoor stream mesocosms (*c.f.*, Odum, 1984) have been used in multispecies toxicity testing for some time in an attempt to better control the dosage received by the communities being tested (*e.g.*, Burks & Wilhm, 1977; Stout & Cooper, 1983; Allard & Moreau, 1987). However, such mesocosms often are costly to construct, are difficult or impossible to transport (therefore site-specific), and are susceptible to meteorologic, anthropogenic, and biological perturbations during testing. In contrast, laboratory stream microcosms can maintain more constant conditions of exposure, are not as prone to vandalism, and reduce variability in results caused by insect drift, emergence, and oviposition. However, development of laboratory-based toxicity tests utilizing stream microcosms has lagged behind multispecies tests that utilize outdoor stream mesocosms, outdoor lacustrine mesocosms, or laboratory-based lacustrine microcosms.

In most streams and rivers, aquatic insects constitute the majority of both macroinvertebrate species and individuals, and, because of their morphological, physiological, and ecological diversity, they exhibit a wide range of responses to stressors. Short-term bioassays with aquatic insects often indicate that they are more tolerant than fish or other invertebrates to environmental stressors such as heavy metals (Rehwoldt *et al.*, 1973; Clubb *et al.*, 1975; Sloof, 1983), ammonia (Williams *et al.*, 1986a), and various hydrocarbons (Sloof, 1983; Millemann *et al.*, 1984; Green *et al.*, 1986). However, aquatic insects, possibly because of lower resistance during molting, tend to exhibit increased sensitivity with longer exposures (Clubb *et al.*, 1975; Maki *et al.*, 1975; Spehar *et al.*, 1978). In addition, short-term bioassays often use late-instar larvae because they are more easily collected in the field and manipulated in the laboratory. Several studies have suggested that earlier instars are more sensitive than later ones to pesticides (Jensen & Gaufin, 1964; Sanders & Cope, 1968; Maki *et al.*, 1975;) and heavy metals (Clubb *et al.*, 1975; Gauss *et al.*, 1985; Williams *et al.* 1986b). Therefore, in order to conduct ecologically meaningful bioassays with aquatic insects, a relatively long exposure period is necessary to ensure that most species undergo one or more molts during testing and that some species have complete life cycles during the period of exposure.

Bioassays using aquatic insects require monitoring of adult emergence so that losses due to emergence are not attributed to toxicant-induced mortality and so that any toxicant-induced inhibition of adult emergence (*c.f.*, Clubb *et al.*, 1975) can be detected. In addition to toxicant effects, changes in photoperiod, current velocity, nutrition, and especially temperature can affect the life history and, therefore, the results of bioassays with these organisms. These and other ecological factors must be accounted for in developing toxicity testing protocols using aquatic insects as test organisms. Although it is a common practice to test insects collected from cool mountain streams in beakers at 20°C in ambient laboratory lighting, it is my opinion that test conditions should reflect those in the source ecosystem at the time of testing.

During the present study, microcosm temperature and photoperiod were nearly identical to ambient conditions in the source ecosystem. The purpose of the study was

to determine the relative importance of current and flow-through in maintaining microcosms of riffle insects during a 30-d bioassay and the degree to which such microcosms reflect the natural communities from which they were derived.

Study Area

Test organisms were derived from a relatively unimpacted riffle of the North Fork of the Roanoke River (NFRR), a third-order mountain stream located in the Ridge and Valley physiographic region of southwest Virginia (Fenneman, 1938). The upper reaches of the drainage are in the Jefferson National Forest, but private residences and pastures are interspersed along most of the drainage. The geology of the drainage consists primarily of limestone and shale. The altitude at the study site was ~445 m above sea level; stream gradient was ~0.8%.

The source riffle was 10 m in length and relatively shaded by riparian vegetation. A large pool was located immediately above the riffle. The source-riffle substrate was cobbles (10-20 cm), ~25% embedded in pebbles (2-5 cm) and sand (0.5-3.0 mm). During the study period, the average riffle width was 5 m, and the average depth was 0.3 m. Current velocity, temperature, and various water chemistry parameters measured in the source riffle are reported in Table 1.

Materials and Methods

Artificial stream designs

Oval artificial streams (1.7 X 0.24 X 0.13 m channel) were constructed of molded fiberglass. Four designs were evaluated to determine their ability to maintain riffle insect communities. The first design was static (no flow-through), and current was not provided. The second design was also static, but current (30 cm sec^{-1}) was provided by a paddlewheel. The third design was flow-through (12 ml min^{-1} dechlorinated New River water) without current, and the fourth design was flow-through with current. The four stream designs were constructed and evaluated in triplicate. Durotest vita-lites® suspended above each stream provided daylight-equivalent lighting and a photoperiod that

corresponded to 0.5 h after sunrise and 0.5 h before sunset on day 15 of the experiment. Each stream was covered by a 1.00 X 0.75 X 0.30 m emergence trap (mesh size ≃1.0 mm).

Derivation of test organisms

Macroinvertebrate communities were developed in rock-filled plastic containers (10.6 X 10.6 X 8.3 cm) that had six circular holes (12 mm dia) in each side and solid bottoms. The substrate placed in the containers was collected from a gravel bar adjacent to the NFRR riffle used to colonize the substrates. One small cobble (7-9 cm) was placed in each container; pebbles (3-5 cm) were then added until the containers were full. The artificial substrates were secured on wooden frames that had previously been anchored to the stream bottom with iron rods and cement blocks.

The artificial substrates were allowed to colonize for 30 d in the NFRR riffle. A review by Rosenberg & Resh (1982) that examined 20 studies containing time series data on macroinvertebrate colonization of artificial substrates suggests that a 30-d colonization period should be sufficient to ensure that macroinvertebrate species equilibrium has been attained. However, maximum densities of some taxa may not be reached during this period (Shaw & Minshall, 1980). Colonized substrates were removed from the riffle by placing a dip net (350 μ mesh) behind the substrate as it was transferred to a cooler filled with stream water for transportation back to the laboratory (transportation time ~0.5 h). Organisms captured in the dip net during the transfer were gently released into the coolers.

Periphyton communities, for use as a source of macroinvertebrate nutrition, were developed on polyurethane foam (PF) artificial substrates (5 X 6.5 X 7.5 cm). The PF substrates were secured to the wooden frames used to hold the macroinvertebrate artificial substrates. The PF substrates were allowed to colonize for 7 d; periphyton species equilibrium on PF substrates is usually attained in 3-7 d in lotic systems (Cairns & Henebry, 1982).

Experimental chronology

On July 31, two colonized PF substrates were placed in each artificial stream and squeezed several times in order to initiate growth of a macroinvertebrate food source on the channel walls. Current was supplied to all streams during the one week period prior to the introduction of macroinvertebrates. On August 7, three colonized macroinvertebrate substrates were randomly assigned to each of the 12 artificial streams; an additional nine substrates were sampled immediately to provide an estimate of macroinvertebrate species-abundances initially placed into each of the artificial stream designs. Also on August 7, nine 0.093 m^2 Hess samples (350 μ mesh) were taken directly from the source riffle to determine how well the artificial substrates reflected species abundances on the natural substrate. During sampling from the natural substrate, the sediments were disturbed to a depth of 7-10 cm. On September 6, all organisms remaining in the artificial streams were sampled, and an additional nine artificial substrates were collected from the NFRR riffle (total colonization time = 60 d) to determine the degree of similarity between the artificial stream microcosms and the NFRR community at the end of the experiment.

Temperature, current velocity, dissolved oxygen, pH, and conductivity were measured in the NFRR and the various artificial streams on August 8 and 24 and on September 6. Alkalinity and hardness were also measured on August 8 and September 6.

Sampling

Adult insects were collected from each artificial stream with an aspirator every 48-72 h throughout the experiment. Artificial substrates were removed from the various streams in a dip net and then washed through a 500 μ sieve. Additional macroinvertebrates and fish were collected by draining the artificial streams through the sieve. The organisms and debris collected were preserved in labeled jars containing 70% ETOH. Sorting was done by hand, and macroinvertebrates, with the exception of midge larvae (Chironomidae:Diptera), Gastropoda, Decapoda, and Oligochaeta, were identified to the lowest possible taxonomic unit using appropriate references. Each taxa was enumerated, and the abundance per stream determined.

Data analysis

The data were generated by determining the total density (adults and young combined) per taxon in each artificial stream. Individual substrates could not be used as subsamples because: 1) the organisms colonizing the internal surfaces of the artificial streams and the adults could not be attributed to a given substrate; and 2) for several taxa, densities in substrates closer to the current source were greater than densities in substrates that received less current. Mean macroinvertebrate species-abundances initially placed in each artificial stream (NFRR30) were determined by randomly combining the nine substrates sampled on August 7 into three, three-substrate groups. Similarly, the nine substrates sampled from the NFRR on September 6 (NFRR60) were randomly combined into three, three-substrate groups in order to determine the degree of similarity between the artificial stream microcosms and the NFRR community at the end of the experiment. These composited three-substrate groups were considered as replicates in statistical comparisons with the various artificial stream designs.

Macroinvertebrate taxa with mean densities greater than three per replicate for any experimental group were considered a core taxon. The density of individuals in each core taxon compared over all experimental groups was used to determine treatment effects. The data were analyzed by a one-way ANOVA in conjunction with the Least Significant Difference criterion for the separation of means.

Species abundances from the August 7 Hess samples (HESS30) could not be directly compared to those from the NFRR30 samples because it was impossible to define a suitable correction factor for the differences in surface area and volume of substrate sampled. Therefore, the analysis was based on the proportional contribution each core taxon made to total numbers in each sample type. In addition, differences in percent composition by trophic relationships using the classification of Merritt & Cummins (1984) were examined.

Results and Discussion

Comparison of artificial and natural streams

Environmental parameters in the artificial streams varied depending on the design. Current velocities in streams supplied with current did not approach velocities measured in the NFRR (Table 1-1). The increase in evaporation due to the paddlewheels caused water levels in the static-current (S-C) streams to drop rapidly, resulting in a loss of current over the substrates by day 10. Temperature in all artificial streams was very similar to the NFRR throughout the experiment. Stream designs with current tended to be slightly cooler than those without current, probably because of an evaporative cooling effect. Dissolved oxygen was lower in the static-no current (S-NC) treatment, and a surface film formed in these streams during the latter portion of the experiment. The flow-through-no current (FT-NC) streams showed an initial drop in dissolved oxygen but were similar to streams supplied with current and the NFRR by the end of the experiment. The initial dissolved oxygen depression in the FT-NC streams may have been because of to the decompositon of macroinvertebrates (primarily Hydropsychidae) which could not survive without current. Conductivity, alkalinity, and hardness followed a similar trend during the course of the experiment and were higher (and closer to NFRR values) in the static streams regardless of the presence of current. The addition of dechlorinated New River water to the flow-through streams apparently lowered these three parameters by dilution.

Fifty-six taxa were collected during the course of the experiment. Of these, 17 taxa were considered to be core taxa (mean densities$\geq$3 per replicate). However, because of the inability to identify the adults of some taxa accurately to species (*e.g., Stenonema* and *Hydropsyche* species), some core taxa were lumped under a higher taxonomic grouping.

The mayfly (Ephemeroptera) genus *Isonychia* (Oligoneuriidae) was represented by two species, *I. bicolor* and *I. tusculalensis*, but *I. bicolor* comprised over 90 percent of total numbers. Densities found in the artificial streams supplied with current were nearly identical to the initial density (Fig. 1-1). Adults collected from the FT-C streams during

Table 1-1. Mean current velocity, temperature, dissolved oxygen, pH, conductivity, alkalinity, and hardness measured in the various treatments. NFRR = North Fork Roanoke River; C-FT = current-flow-through; C-S = current-static; NC-FT = no current-flow-through; NC-S = no current-static.

Treatment	Velocity cm sec^{-1}	Temp °C	DO mg ℓ^{-1}	pH	Cond $\mu S\ cm^{-1}$	Alk mg ℓ^{-1}	Hard mg ℓ^{-1} $CaCO_3$
			8 August 1986				
NFRR	50.3	21.0	7.2	8.2	400	183.2	250.00
C-FT	30.0	20.5	8.8	8.1	190	66.5	93.00
C-S	30.0	20.5	8.7	8.3	195	88.2	110.00
NC-FT	00.0	20.5	7.2	7.9	190	65.6	87.00
NC-S	00.0	20.5	6.5	8.2	193	85.7	108.00
			24 August 1986				
NFRR	98.5	22.0	9.2	8.1	400	NA**	NA
C-FT	30.0	19.0	9.0	8.1	143	NA	NA
C-S	*	19.0	9.0	8.4	258	NA	NA
NC-FT	00.0	20.5	7.2	7.7	170	NA	NA
NC-S	00.0	20.5	6.9	7.7	237	NA	NA
			6 September 1986				
NFRR	73.4	18.0	9.2	8.1	345	151.2	210.00
C-FT	30.0	19.0	8.6	8.0	165	62.5	70.00
C-S	*	19.0	8.5	8.4	350	157.9	195.00
NC-FT	00.0	21.0	8.1	8.3	163	59.2	75.00
NC-S	00.0	21.0	6.8	8.1	253	108.1	121.00

* due to low water levels paddlewheels produced waves instead of a current
** NA—parameter not measured on that date

the course of the experiment comprised ~40 percent of total numbers. Significantly fewer *Isonychia* survived in the FT-NC design, and only two adults were recovered early in the experiment from the S-NC design. In the FT-C design, densities of *Isonychia* were much greater on substrates closer to the current source. It is apparent that *Isonychia*, a filter-feeder, requires current for long-term maintenance in laboratory streams.

The mayfly genus *Stenonema* (Heptageniidae) was represented by three core species: *S. mediopunctatum, S. modestum,* and *S. luteum*. The FT-C, S-C, and FT-NC stream designs maintained the *Stenonema* species at or above initial densities, and it is interesting to note that a number of individuals, primarily *S. luteum*, were able to survive in the S-NC design (Fig. 1-2). *Stenonema luteum* was not apparent in the NFRR30 samples, but many early and middle instar naiads were found in the artificial streams and the

NFRR60 samples at the end of the experiment. Apparently, *S. luteum* entered the stream microcosms in the egg stage (or as nymphs < 500 μ), and therefore was not recorded in the NFRR30 samples. This suggests that the artificial stream microcosms were able to track changes in the species composition of the natural stream community, a necessary component for field validation of future toxicity tests. The presence of *S. luteum* may also explain the higher densities found in the artificial streams supplied with current relative to initial (NFRR30) densities.

The mayfly genus *Baetis* (Baetidae), primarily *Baetis intercalaris*, was well maintained in streams provided with current but did not survive in streams without current (Fig. 1-3). Adults accounted for over 95 percent of the *Baetis* collected from the FT-C and S-C streams, and the lower numbers of *Baetis* nymphs collected from the source riffle at the end of the experiment (NFRR60) suggests that an emergence also occurred in the field. The number of adults collected in the FT-C and S-C streams was much greater than the number of nymphs initially put into the streams (range of NRRR30 replicates = 28-39 individuals) , suggesting that some of the adults collected may have been introduced as eggs (or nymphs < 500 μ), and completed their entire life cycle in the artificial streams. *Baetis* species are known to exhibit a fast-seasonal life cycle (Merritt & Cummins, 1984), but completion of the entire life cycle in <30 d has not previously been reported. The very low densities of *Baetis* nymphs present in the artificial streams at the end of the experiment indicate the necessity of collecting adults during the course of ecological or toxicological experimentation with this widespread genus.

No significant differences in treatment effects were found for the mayfly genus *Caenis* sp. (Caenidae) (Fig. 1-4). However, densities in the S-NC streams were considerably lower than initial densities. The observation that *Caenis* sp. was able to survive in all stream designs is not surprising because nymphs of this genus are known to inhabit quiet or even stagnant water (Edmunds *et al.*, 1976). No adult *Caenis* were collected during the experiment.

Artificial streams supplied with current were able to maintain all mayfly taxa at or above initial levels for the entire 30-d experiment. The successful maintenance of

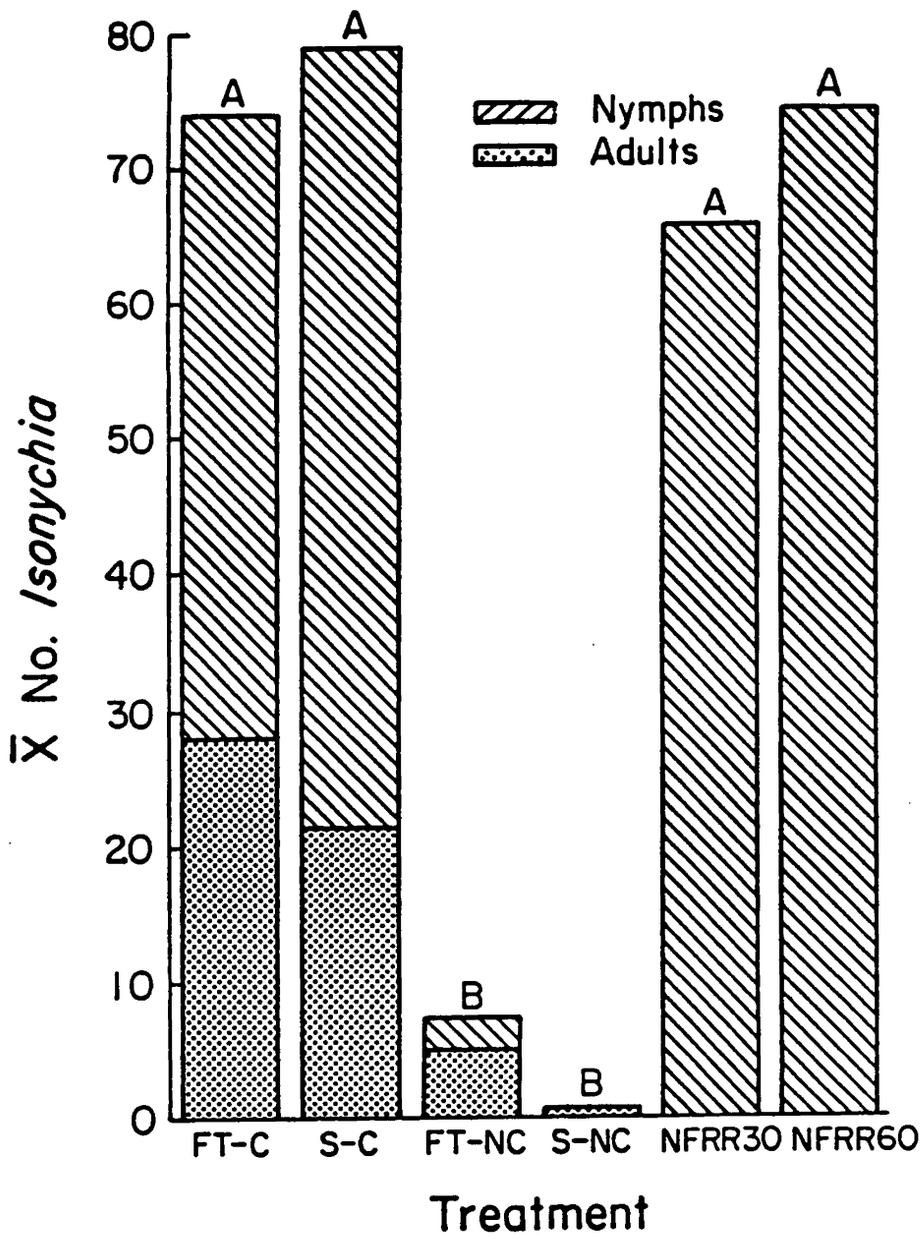

Figure 1-1. Mean number of *Isonychia* (Ephemeroptera:Oligoneuriidae) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

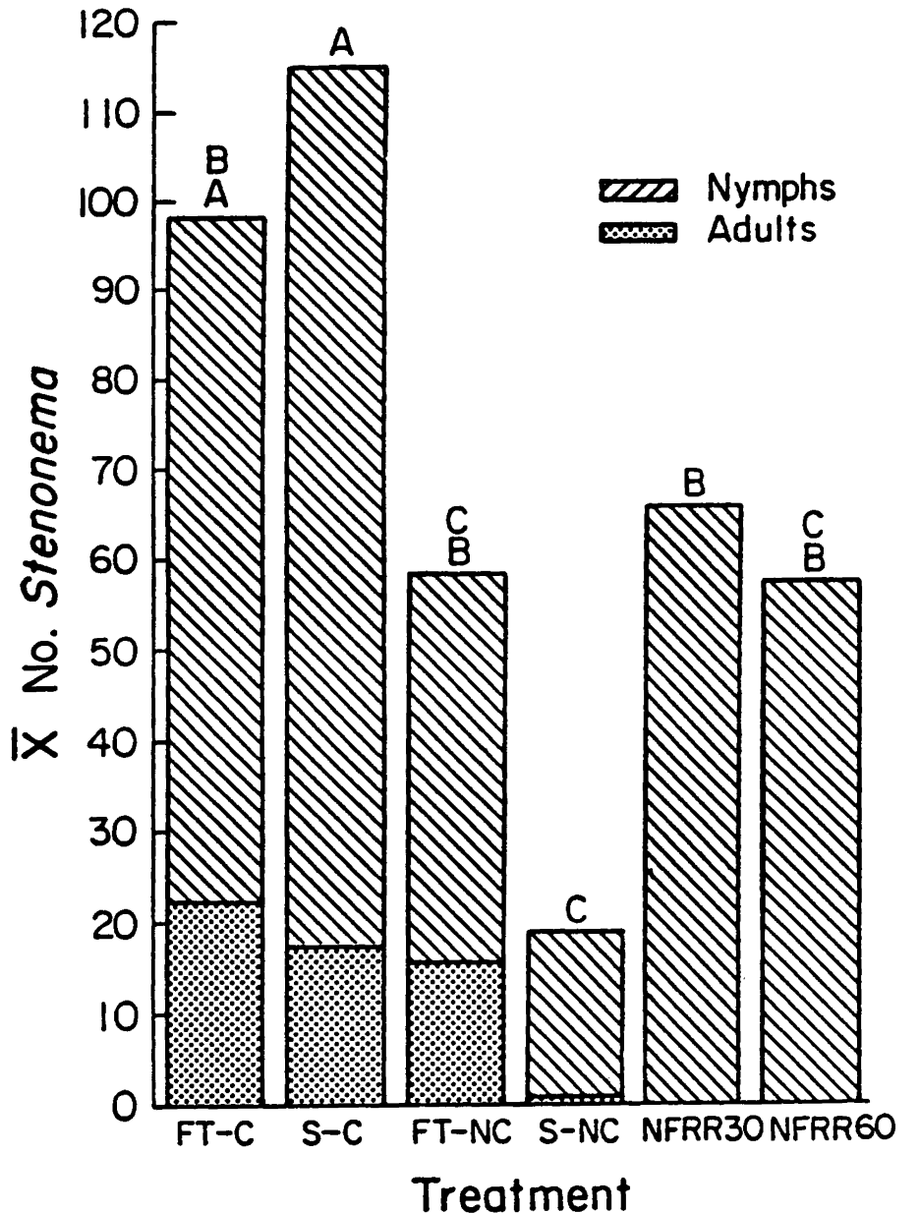

Figure 1-2. Mean number of *Stenonema* (Ephemeroptera:Heptageniidae) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

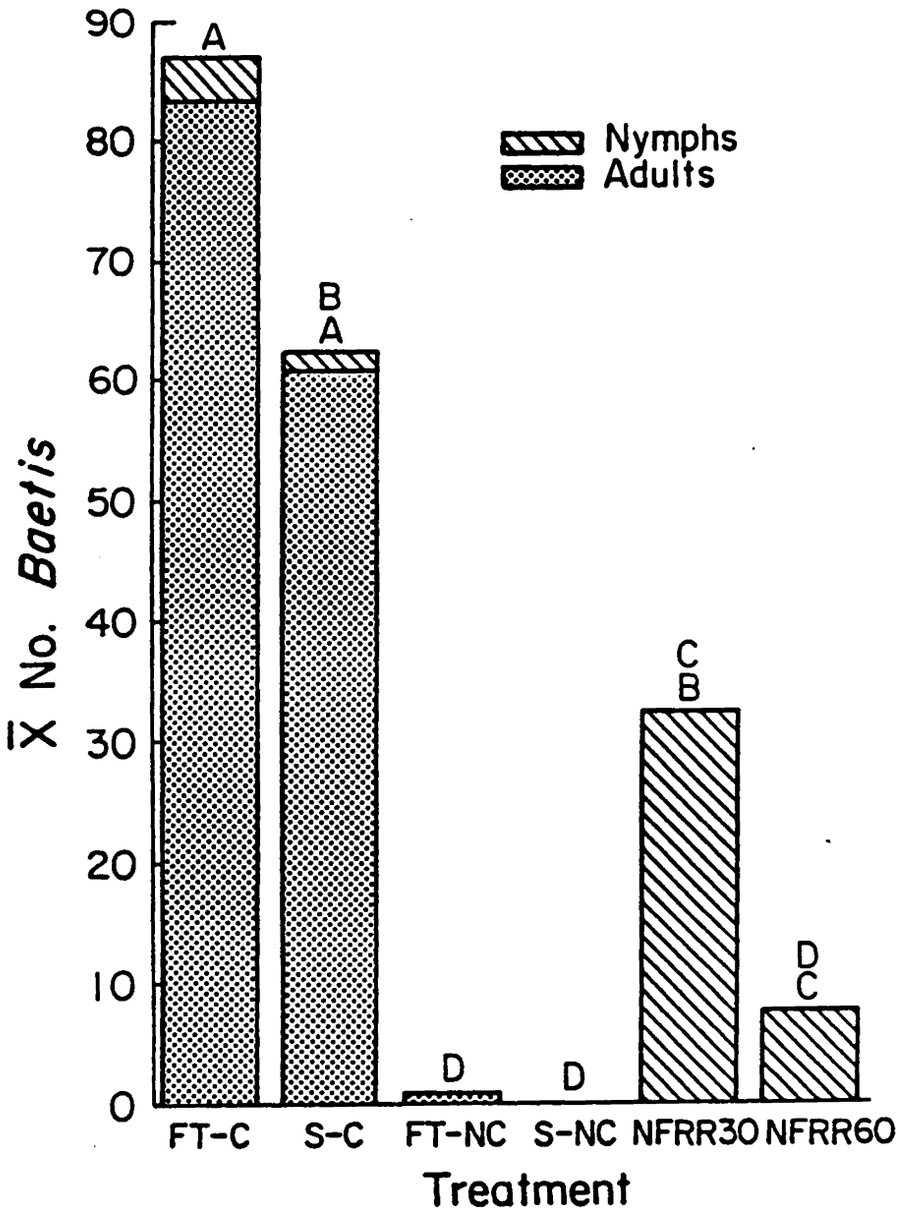

Figure 1-3. Mean number of *Baetis* (Ephemeroptera:Baetidae) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

mayflies is important for toxicological work because mayflies as a group are considered to be relatively sensitive to environmental stressors. The data indicate that long-term toxicity tests (possibly extending over entire life cycles) can be run with this group of insects. However, adult emergence must be monitored in order to interpret test results accurately.

The riffle beetles (Coleoptera: Elmidae) were represented by five species: *Stenelmis crenata, Dubiraphia quadrinota, Optioservus trivittatus, Microcylloepus pusillus,* and *Promoresia elegans.* However, the only elmid with densities in the artificial substrates sufficient to permit an accurate analysis was *S. crenata. Stenelmis crenata* was maintained at initial (NFRR30) densities in all artificial stream designs (Fig. 1-5). The artificial streams that included current had the largest populations, primarily because of higher larval densities. The survival of this beetle in the two artificial stream designs without current corroborates other reports of elmid tolerance to environmental stressors (Cairns *et al.*, 1971; Barton & Wallace, 1979: Pontasch & Brusven, in press). The significantly higher densities of *S. crenata* in the NFRR60 samples probably resulted from a substantial increase (relative to NFRR30 substrates) in the amount of sand and small pebbles present in those substrates. This material was deposited during two periods of high flow in August and made the NFRR60 substrates more similar, in terms of substrate composition and interstitial pore size, to the hyporheic zone of the natural substrate. *Stenelmis crenata* has been reported to occur in larger numbers in the hyporheic zone relative to surficial areas (Godbout & Hynes, 1982).

The Hydropsychidae (Trichoptera) of the genera *Hydropsyche* and *Cheumatopsyche* were the dominant caddisflies and colonized the artificial substrates in large numbers (Figs. 1-6 and 1-7). The major *Hydropsyche* species collected included: *H. bronta, H. betteni, H. frisoni,* and *H. venularis.* Two other caddisflies collected in low numbers were larval *Chimarra* sp. (Philopotamidae) and adult Hydroptilidae. The FT-C stream provided the best environment for the hydropsychids but was only able to maintain ~35 percent of initial densities. Apparently, the quality or quantity of food available in the seston of the artificial streams was insufficient to maintain the very large numbers of

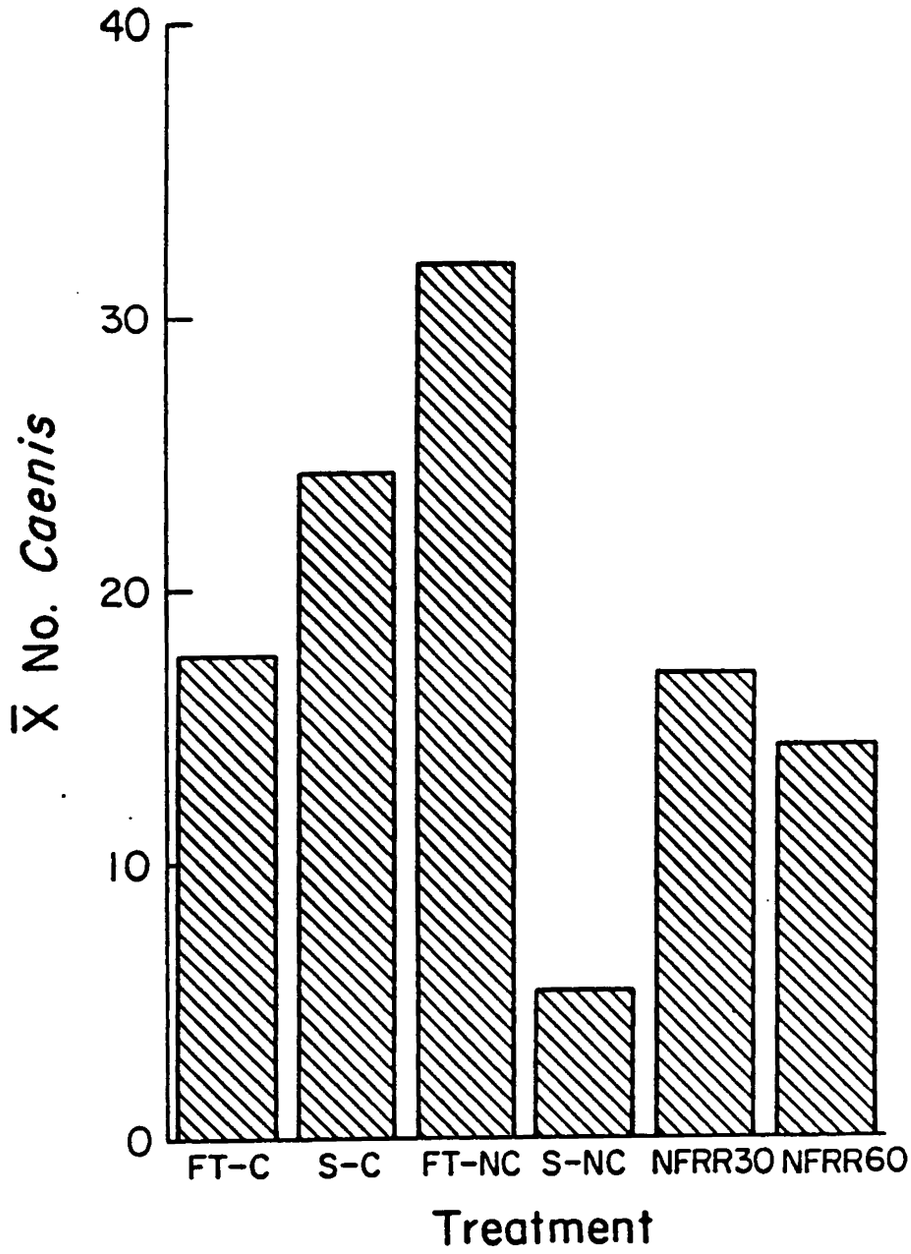

Figure 1-4. Mean number of *Caenis* (Ephemeroptera:Caenidae) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Their were no adults collected and no significant differences among treatments.

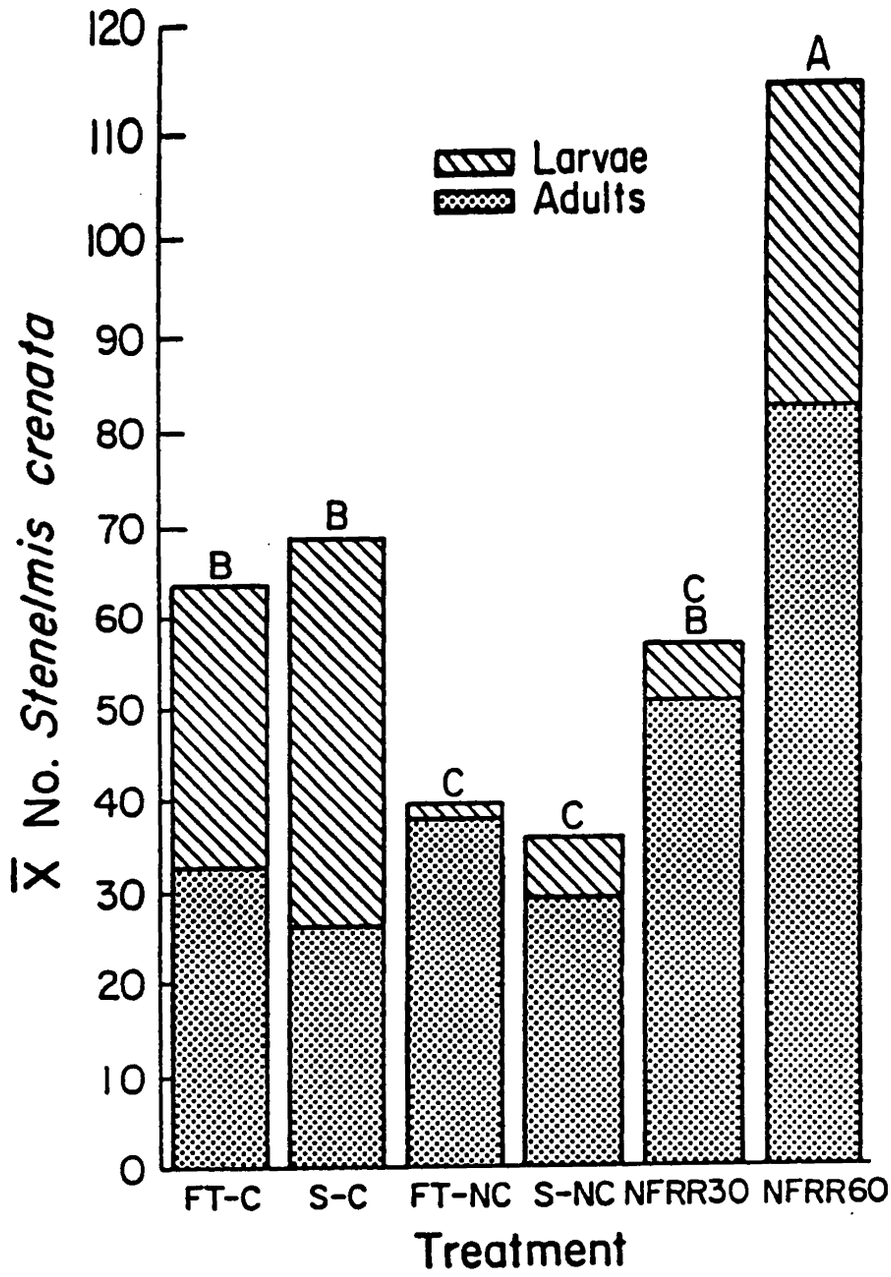

Figure 1-5. Mean number of *Stenelmis Crenata* (Coleoptera:Elmidae) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

hydropsychids initially present. A supplemental source of nutrition may be required during long-term experiments with large numbers of hydropsychids.

The Chironomidae (Diptera) exhibited the most dramatic response to the treatments. Three subfamilies, Chironominae, Orthocladiinae, and Tanypodinae, were represented. The initial number placed in each stream was ~100 individuals, and this density was maintained or exceeded in all stream designs (Fig. 1-8). However, by the end of the experiment, the mean number collected from the FT-NC streams approached 1700. Most of these chironomids were *Paratanytarsus boiemicus* complex; a group that can reproduce parthenogenetically (S.S. Roback, personal communication). It is apparent that many individuals had complete life cycles during the 30-d test, and once again, the necessity of monitoring adult emergence is indicated.

The Corydalidae (Megaloptera), represented by two species *Corydalus cornutus* and *Nigronia serricornis*, were able to survive in all of the artificial stream designs (Fig. 1-9). Artificial streams supplied with current maintained these two species slightly better than streams without current. No stream design maintained the corydalids at initial densities, but streams supplied with current were not significantly different from the NFRR60 samples taken at the end of the experiment. No adults were collected.

The monitoring of adult emergence was most important for *Baetis*, *Isonychia*, and the chironomids. The emergence of *Isonychia* was fairly uniform in the FT-C design throughout the experiment, while *Baetis* and the chironomids had higher emergence rates during the final 20 d of the experiment (Fig. 1-10). By day 10, ~35 percent of initial chironomid and *Baetis* densities, and 15 percent of *Isonychia* densities had emerged from the FT-C streams, indicating that adult emergence should be monitored even during short-term bioassays with these taxa.

Comparison of artificial and natural substrates

Some taxa failed to colonize the artificial substrates in numbers that reflected their natural benthic densities. One species of elmid beetle, *Optioservus trivittatus* , comprised over 10 percent of total numbers in the Hess samples but was rarely found in the artificial substrate samples (Table 1-2). Similarly, larvae of another elmid, *Stenelmis crenata*, and

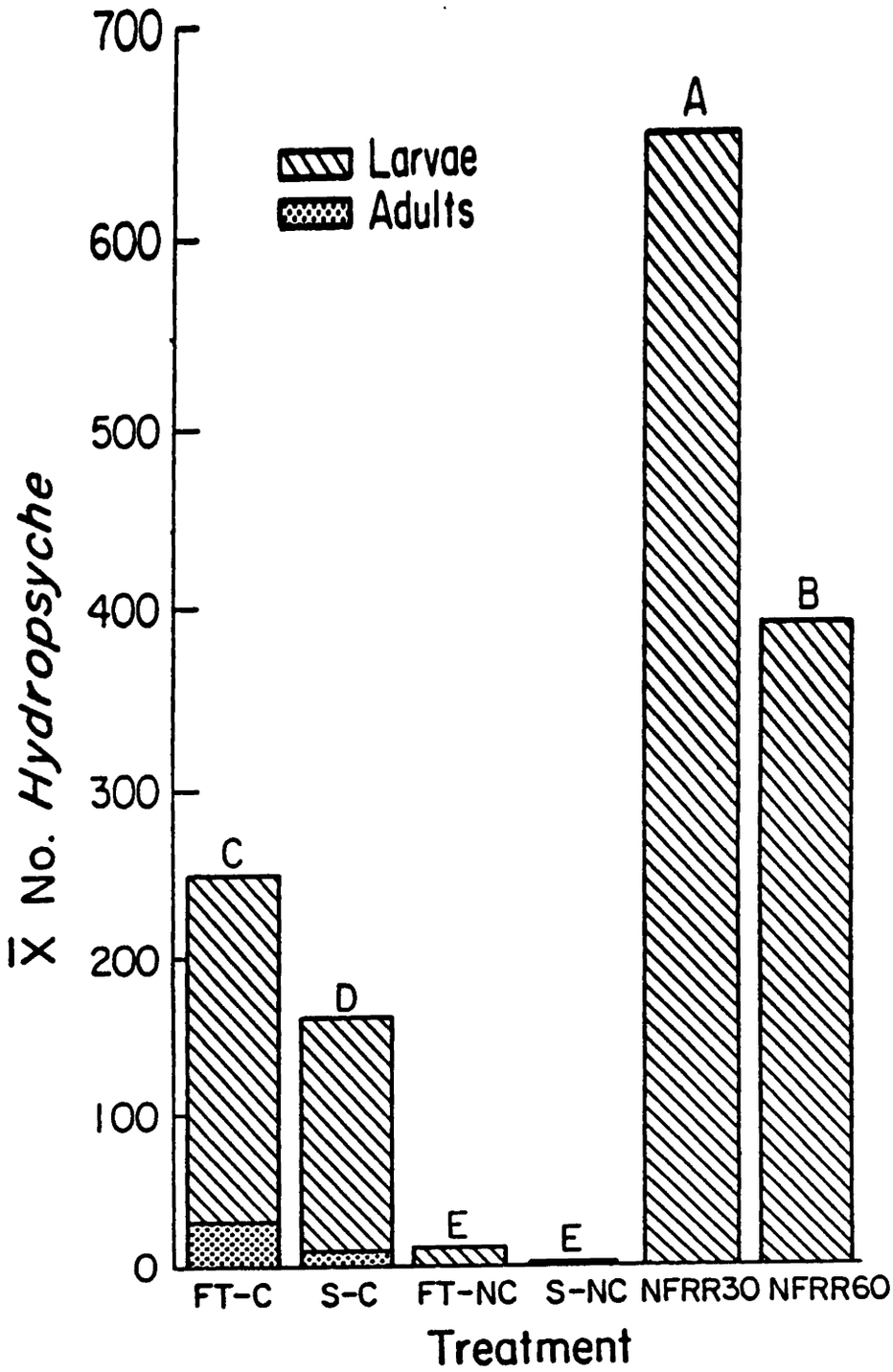

Figure 1-6. Mean number of *Hydropsyche* (Trichoptera:Hydropsychidae) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

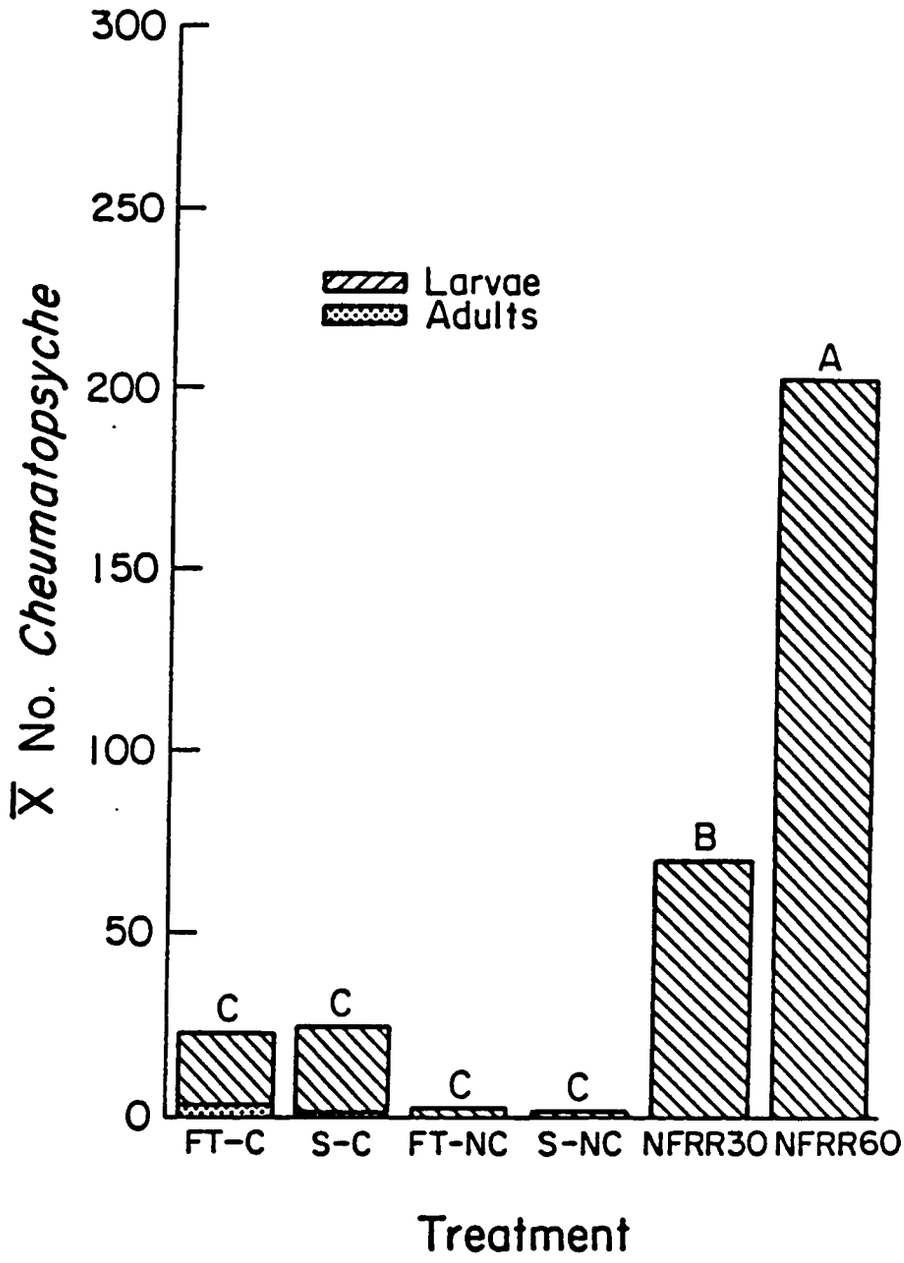

Figure 1-7. Mean number of *Cheumatopsyche* (Trichoptera:Hydropsychidae) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

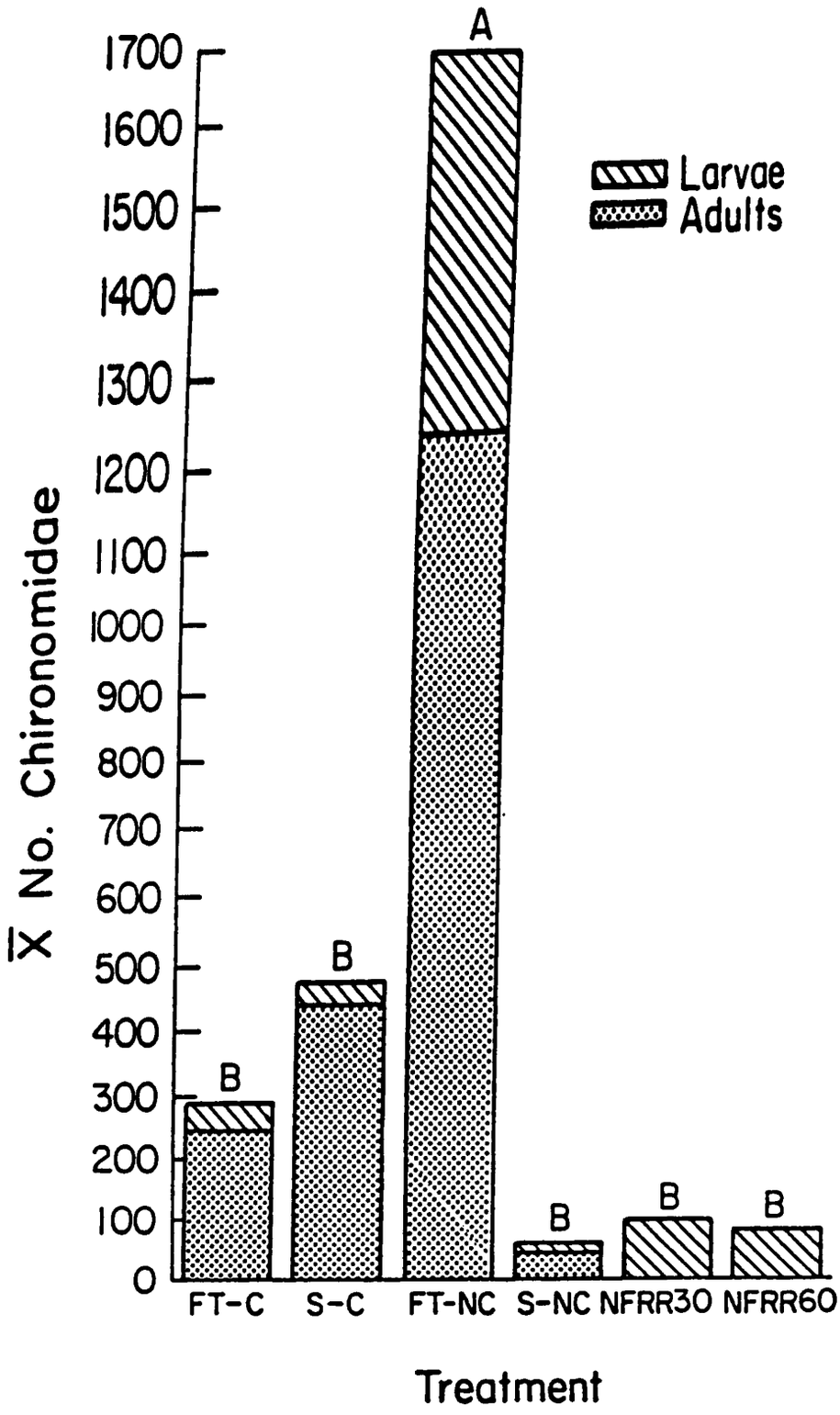

Figure 1-8. Mean number of Chironomidae (Diptera) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

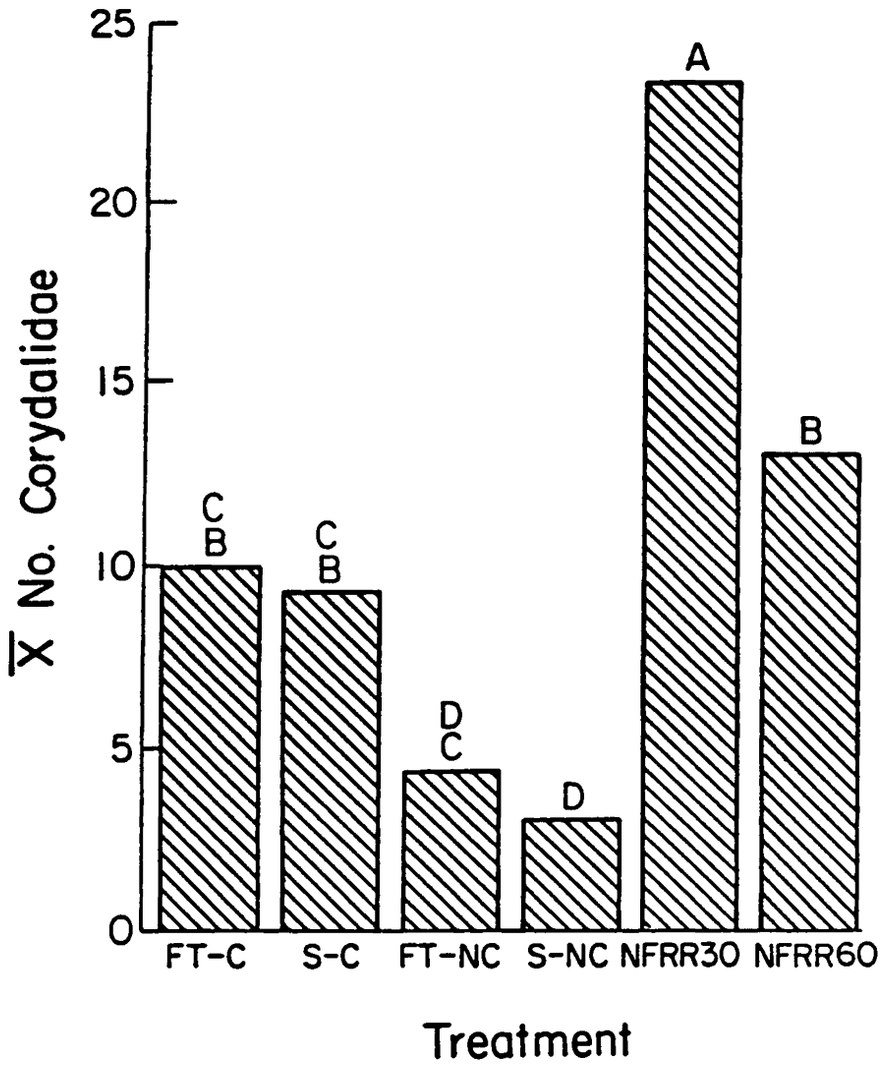

Figure 1-9. Mean number of Corydalidae (Megaloptera) after 30 d in four artificial stream designs: FT-C = flow-through with current; S-C = static with current; FT-NC = flow-through without current; and S-NC = static without current. NFRR30 is the estimated number of organisms placed in each stream at the beginning of the experiment. NFRR60 is the number of organisms present on substrates sampled from the source riffle at the end of the experiment. Statistical analysis is based on combined densities of adults and nymphs. Bars with the same letter are not significantly different ($P \leq 0.05$)—analyzed by ANOVA/LSD.

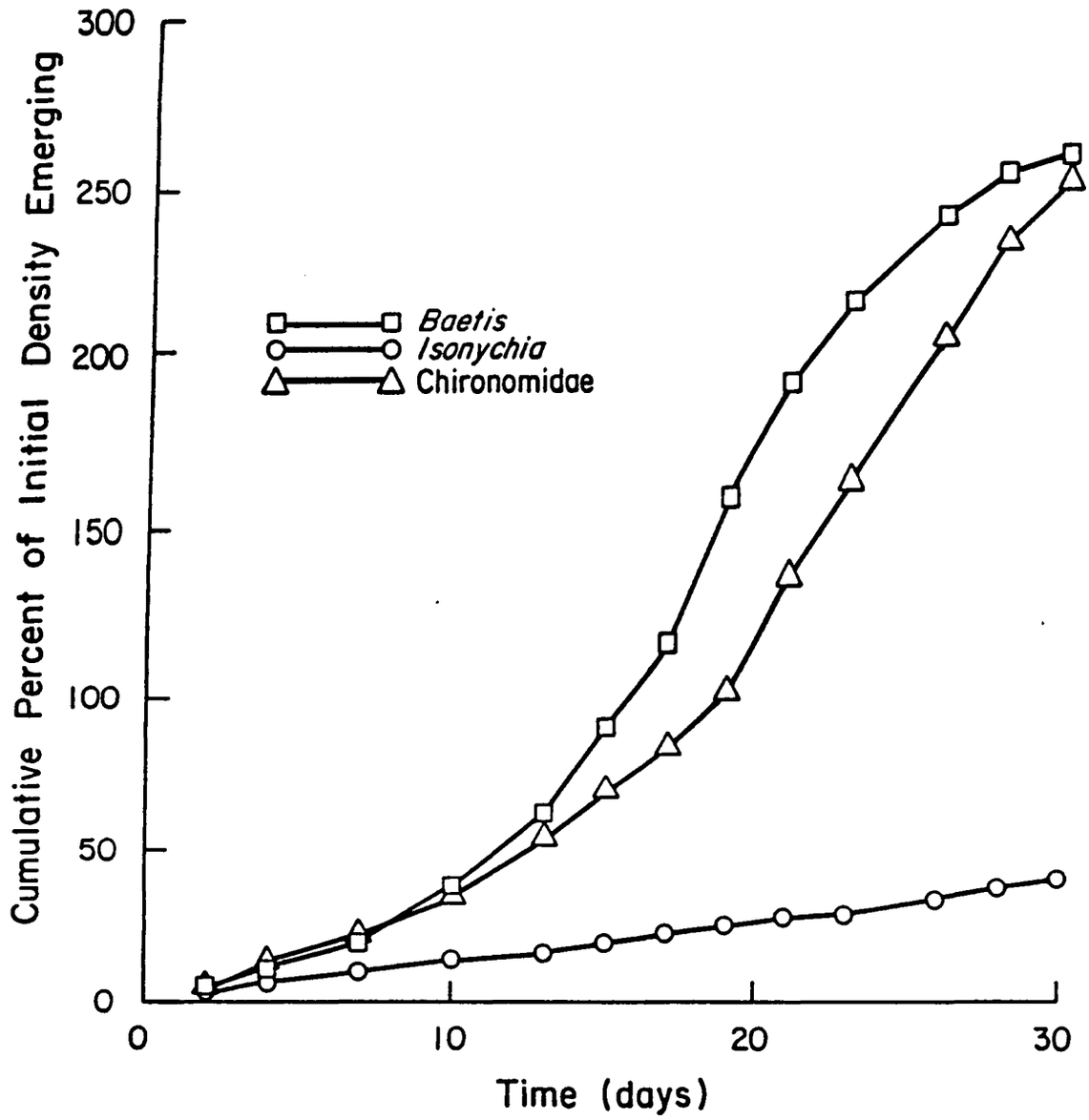

Figure 1-10. Cumulative percent of initial densities (NFRR30) of *Baetis*, *Isonychia*, and chironomids emerging as adults from an artificial stream design with flow-through and current during a 30-d experiment.

Caenis sp. nymphs were much more abundant in the Hess samples than in the artificial substrates. Elmid beetles and *Caenis* are both reported to be more abundant in the hyporheic zone relative to surficial sediments (Williams & Hynes, 1974; Godbout & Hynes, 1982). The depth to which the substrate was disturbed during sampling (7-10 cm) and the relatively large amount of sand and small pebbles in the samples suggest that the Hess samples included the upper hyporheos. It is apparent that the artificial substrates used in this study were not good models of a hyporheic habitat and were, therefore, not colonized as well by the hyporheos. The taxa that failed to colonize the artificial substrates adequately were all collector-gatherers, suggesting that a lack of detrital deposits may have made the artificial substrates a less suitable habitat for these organisms. In a study by Culp *et al.* (1983) in which substrates were embedded in a riffle, the presence of a detrital food source was shown to be more important than substrate size or heterogeneity in controlling detritivore colonization of artificial substrates. Detrital amendments to the type of artificial substrates used in this experiment may have increased representation of the hyporheos, but a thigmotactic or negative phototactic response may still have inhibited colonization.

In contrast, the Hydropsychidae and to a lesser degree *Isonychia* were more abundant in the artificial substrates than in the Hess samples (Table 1-2). These organisms prefer larger cobbles and boulders that protrude into the current in order to maximize filter-feeding efficiency. The artificial substrates used in this study were apparently good analogues for these larger cobbles and boulders. Selectivity of artificial substrates for hydropsychids has previously been reported (*e.g.*, Benfield *et al.*, 1974) and could probably be reduced by embedding the substrates farther into the riffle.

The artificial substrates did not collect a fauna representative of the riffle community in terms of the relative or absolute abundances of some taxa, but the number and kinds of species collected in the two sample types were nearly identical. In addition, the densities of two of the three taxa that colonized the artificial substrates in lower numbers relative to their benthic densities would still have been high enough to permit a determination of treatment effects during a 30-d bioassay. Although, quantitative data gener-

Table 1-2. Percent contribution of core taxa and functional groups to pooled total numbers in the nine Hess (total number = 4,702) and nine artificial substrate (total number = 3,236) samples taken on 7 August 1987. C-G = collector-gatherers; C-F = collector filterers; SC = scrapers; P = predators.

Taxon / Functional Group	Artificial Substrates Percent	Hess Samples Percent
Isonychia (C-F)	**6.08**	**2.08**
Stenonema (SC)	6.12	6.65
Baetis (C-G)	3.00	1.17
Caenis (C-G)	**1.57**	**18.08**
Stenelmis crenata adults (C-G)	4.70	7.95
S. crenata larvae (C-G)	**0.60**	**25.44**
Optioservus trivittatus adults (C-G)	0.00	1.27
O. trivittatus larvae (C-G)	**0.06**	**10.14**
Hydropsyche (C-F)	**60.11**	**15.46**
Cheumatopsyche (C-F)	6.52	3.76
Tanypodinae (P)	1.08	1.18
Other Chironomidae (C-G)	8.00	4.51
Corydalidae (P)	2.16	2.36
Collector-gatherers	**18.0**	**68.6**
Collector-filterers	**72.7**	**21.3**
Scrapers	6.1	6.6
Predators	3.2	3.5

ated during a bioassay may not be directly relatable to benthic data collected from the natural substrate of a receiving stream, my results indicate that such data could be field validated through the use of artificial substrates colonized downstream from the point of impact.

Conclusions and Recommendations

My results indicate that microcosms of riffle insect communities can be used in long-term multispecies toxicity testing. Current and flow-through should be provided, and temperature and photoperiod should be similar to ambient conditions in the source ecosystem. Instead of the one week period used in this study, the artificial streams should be inoculated with a periphytic food source three weeks prior to the introduction of

macroinvertebrates (unpublished data). Adult emergence must be monitored in order to interpret test results accurately.

The artificial substrates used in this study were selectively colonized by some taxa, but most taxa present in the source riffle colonized the substrates in numbers sufficient to permit a determination of treatment effects during a bioassay. Substrate selectivity may be reduced by providing detrital amendments and embedding the substrates in the source riffle.

Further research in other laboratories and geographic regions should be undertaken in an effort to develop standardized protocols for conducting multispecies toxicity tests with lotic macroinvertebrates. These protocols would provide environmentally realistic and scientifically justifiable tools for reducing uncertainty in environmental risk assessment.

References

Allard, M. & G. Moreau, 1987. Effects of experimental acidification on a lotic macroinvertebrate community. Hydrobiologia 144:37-49.

Barton, D. R. & R. R. Wallace, 1979. The effects of an experimental spillage of oil sands tailings on benthic invertebrates. Envir. Pollut. 18:305-312.

Benfield, E. F., A. C. Hendricks & J. Cairns, Jr., 1974. Proficiencies of two artificial substrates in collecting stream macroinvertebrates. Hydrobiologia 45:431-440.

Burks, S. L. & J. L. Wilhm, 1977. Bioassays with a natural assemblage of benthic macroinvertebrates. In F. L. Mayer & J. L. Hamelink (eds.), Aquatic Toxicology and Hazard Evaluation, STP 634. American Society for Testing and Materials, Philadelphia, Pennsylvania:127-136.

Cairns, J., Jr. (ed.), 1985. Multispecies Toxicity Testing. Pergamon Press, London, 261 pp.

Cairns, J., Jr. (ed.), 1986. Community Toxicity Testing, STP 920. American Society for Testing and Materials, Philadelphia, Pennsylvania, 350 pp.

Cairns, J., Jr. & M. S. Henebry, 1982. Interactive and noninteractive protozoan colonization processes. In J. Cairns, Jr. (ed.), Artificial Substrates. Ann Arbor Science Publishers, Inc., Ann Arbor, Michigan:23-70.

Cairns J., Jr., J. S. Crossman, K. L. Dickson & E. E. Herricks, 1971. The recovery of damaged streams. Assoc. Southeastern Biol. Bull. 18:79-106.

Clubb, R. W., A. R. Gaufin & J. L. Lord, 1975. Acute cadmium toxicity tests upon nine species of aquatic insects. Envir. Res. 9:332-341.

Culp, J. M., S. J. Walde & R. W. Davies, 1983. Relative importance of substrate particle size and detritus to stream benthic macroinvertebrate distribution. Can. J. Fish. Aquat. Sci. 40:1568-1574.

Edmunds, G. F., S. L. Jensen & L. Berner, 1976. The mayflies of North and Central America. University of Minnesota Press, Minneapolis, 330 pp.

Fenneman, N. M., 1938. Physiography of the eastern United States. McGraw-Hill, N.Y., 714 pp.

Gauss, J. D., P. E. Woods, R. W. Winner & J. H. Skillings, 1985. Acute toxicity of copper to three life stages of *Chironomus tentans* as affected by water hardness-alkalinity. Envir. Pollut. 37:149-157.

Giesy, J. P., Jr., 1980. Microcosms in Ecological Research. National Technical Information Center, Springfield, Virginia, 1110 pp.

Godbout, L. & H. B. N. Hynes, 1982. The three dimensional distribution of the fauna in a single riffle in a stream in Ontario. Hydrobiologia 97:87-96

Green, D. W. J., K. A. Williams & D. Pascoe, 1986. Studies on the acute toxicity of pollutants to freshwater macroinvertebrates 4. Lindane (γ—Hexachlorocyclohexane). Arch. Hydrobiol. 106:263-273.

Jensen, L. D. & A. R. Gaufin, 1964. Effects of ten organic insecticides on two species of stonefly naiads. Trans. am. Fish. Soc. 93:27-34.

Maki, A. W., L. Geissel & H. E. Johnson, 1975. Comparative toxicity of the larval lampricide TFM (3 triflouromethyl-4-nitrophenol) to selected benthic macroinvertebrates. J. Fish. Res. Bd Can. 32:1455-1459.

Merritt, R. W. & K. W. Cummins, 1984. An Introduction to the Aquatic Insects of North America, 2nd ed., Kendall/Hunt, Dubuque, Iowa, 722 pp.

Millemann, R. E., W. J. Birge, J. A. Black, R. M. Cushman, K. L. Daniels, P. J. Franco, J. M. Giddings, J. F. McCarthy & A. J. Stewart, 1984. Comparative acute toxicity to aquatic organisms of components of coal-derived synthetic fuels. Trans. am. Fish. Soc. 113:74-85.

Odum, E. P., 1984. The mesocosm. Bioscience 34:558-562.

Pontasch, K. W. & M. A. Brusven, in press. Macroinvertebrate response to a gasoline spill in Wolf Lodge Creek, Idaho, USA. Archiv. Hydrobiol.

Rehwoldt, R., L. Lasko, C. Shaw & E. Wirhowski, 1973. The acute toxicity of some heavy metal ions toward benthic organisms. Bull. Envir. Contam. Toxicol. 10:291-294.

Rosenberg, D. M. & V. H. Resh, 1982. The use of artificial substrates in the study of freshwater benthic macroinvertebrates. In J. Cairns, Jr. (ed.), Artificial Substrates. Ann Arbor Science Publishers, Inc., Ann Arbor, Michigan:175-236.

Sanders, H. O. & O. B. Cope, 1968. The relative toxicities of several pesticides to naiads of three species of stonefly. Limnol. Oceanogr. 33:112-117.

Shaw, D. W. & G. W. Minshall, 1980. Colonization of an introduced substrate by stream macroinvertebrates. Oikos 34:259-271.

Slooff, W., 1983. Benthic macroinvertebrates and water quality assessment: some toxicological considerations. Aquat. Toxicol. 4:73-82.

Spehar, R. L., R. L. Anderson & J. T. Fiandt, 1978. Toxicity and bioaccumulation of cadmium and lead in aquatic invertebrates. Envir. Pollut. 15:195-208.

Stout, R. J., & W. E. Cooper, 1983. Effect of p-cresol on leaf decomposition and invertebrate colonization in experimental outdoor streams. Can J. Fish. Aquat. Sci. 40:1647-1657.

Williams, D. D. & H. B. N. Hynes, 1974. The occurence of benthos deep in the substratum of a stream. Freshwat. Biol. 4:233-256.

Williams, K. A., D. W. J. Green & D. Pascoe, 1986a. Studies on the acute toxicity of pollutants to freshwater macroinvertebrates, 3. Ammonia. Arch. Hydrobiol. 106:61-70.

Williams, K. A., D. W. J. Green, D. Pascoe & D. E. Gower, 1986b. The acute toxicity of cadmium to different larval stages of *Chironomus riparius* (Diptera: Chironomidae) and its ecological significance for pollution regulation. Oecologia 70:362-366.

CHAPTER TWO

MULTISPECIES TOXICITY TESTS USING INDIGENOUS ORGANISMS: PREDICTING THE EFFECTS OF COMPLEX EFFLUENTS IN STREAMS

Abstract

A study was conducted to determine if contaminant-induced changes in macroinvertebrate and periphyton communities in laboratory stream microcosms could be used to predict macroinvertebrate and periphyton responses in a natural stream receiving the same contaminant. The laboratory streams were covered by emergence traps, and daylight equivalent lights provided a natural photoperiod. Test organisms were transferred to the laboratory streams in rock-filled containers previously colonized for 30 d upstream of the effluent outfall. The microcosms were dosed in quadruplicate with four (0.0, 0.1, 1.0 and 10.0%) concentrations of a complex effluent; these concentrations reflected those in the field. Macroinvertebrate and periphyton samples were taken at field sites to evaluate how well responses in the laboratory microcosms corresponded to responses in the receiving stream. Adult insects were collected every 48-72 h from each microcosm, and drift was monitored 1 h after initial placement and 1 h after dosing. After 20 d (5 August-25 August 1987), all remaining macroinvertebrates were censused and periphyton samples were taken. Treatment effects were evaluated using combined densities of adult and young insects. Results were compared to field samples to determine the predictive utility of the laboratory tests.

The stream microcosms were able to maintain most taxa at or above initial densities. Mayfly densities in the microcosms were significantly ($P \leq 0.05$) reduced at 1.0 or 10.0% effluent depending on species. Mayflies responded at much lower concentrations than acute tests with *Daphnia magna* would have predicted. Hydropsychids were not affected by the effluent, and chironomids and periphyton were stimulated. Overall, the stream microcosms accurately predicted the macroinvertebrate and periphyton response observed in the field.

Introduction

Managing the hazards of anthropogenic stressors in aquatic ecosystems requires predictive techniques that are cost-effective and accurate. Studies of the effectiveness of short-term, single species biological testing of effluents under the U.S. Environmental Protection Agency's (USEPA) Complex Effluent Toxicity Testing Program (as summarized by Wall and Hanmer 1987) have found correlations between response of single species in laboratory tests and certain indices of field response. However, single species tests are often overprotective when large application factors are employed, and accurate predictions of how much and what kind of damage to ecosystem structure and function to expect are not possible. Accurate predictive tools will avoid the costs of underprotection of the environment (i.e., damage to environmental resources) and the costs of overprotection (i.e., construction of waste treatment facilities that provide no biological benefit), while inaccurate predictions engender these costs.

An obvious way to reduce uncertainty in environmental risk assessments is by incorporating tests that directly examine those properties of ecosystems that are the object of protective legislation (e.g., species richness). Such multispecies toxicity tests are being developed (see Giesy 1980; Cairns 1985, 1986). Early work suggests that multispecies tests are not only more realistic models of exposure than single species toxicity tests, but they may also be no more expensive (Perez and Morrison 1985; Niederlehner et al. 1986; Cairns and Pratt 1987) and just as replicable (Geisy and Allred 1985) as conventional tests. Nevertheless, multispecies toxicity tests have not been generally accepted in the regulatory framework.

The overall goal of the research reported here was to develop and field-validate a laboratory-based, multispecies toxicity testing protocol using artificial stream microcosms containing macroinvertebrates and periphyton. In order to field-validate the protocol, this study included a site-specific, complex effluent. However, the protocol could be standardized and used to predict adverse effects of individual chemicals by using regional reference streams as sources of macroinvertebrates and periphyton.

Periphyton was used because: 1) it is involved in both primary production and decomposition and should, therefore, reflect toxicant effects on these important ecosystem functions; and 2) it is an important source of food for many macroinvertebrates. Macroinvertebrates were used because: 1) they serve as direct intermediates in the trophic chain between microbial species and important vertebrate species such as fish and waterfowl; and 2) they are widely used in environmental monitoring and surveillance.

Primary emphasis was placed on aquatic insects because: 1) they constitute the majority of macroinvertebrate species and individuals in most streams and rivers; and 2) they exhibit a wide range of tolerances to stressors because of their morphological, physiological and ecological diversity. One of the earliest reported bioassays employing aquatic insects was conducted by Surber and Thatcher (1963). Burks and Wilhm (1977) demonstrated that natural assemblages of benthic insects and other macroinvertebrates for use in bioassays could be transported from a source ecosystem to outdoor artificial streams on previously colonized artificial substrates. More recently, a number of investigations have employed outdoor stream mesocosms (c.f., Odum 1984) of various sizes in conducting bioassays with macroinvertebrate communities (e.g., Perkins 1983; Stout and Cooper 1983; Allard and Moreau 1987). However, laboratory-based bioassays utilizing lotic macroinvertebrates have primarily been single species tests on laboratory-tolerant taxa.

Investigators attempting to monitor either outdoor mesocosm or laboratory microcosm responses of lotic insects encounter logistical problems with: 1) long-term maintenance of less tolerant taxa; 2) increased sensitivity to toxicants during molts; 3) escape of emergent adults during testing; and 4) contaminant-induced avoidance behavior (drift). Initial experiments (Pontasch and Cairns, in press) utilizing the techniques described here demonstrated that: 1) relatively sensitive taxa such as mayflies can be maintained for at least 30 d, thus ensuring that most taxa will molt and some taxa (e.g., *Paratanytarsus* sp.:Chironomidae) will have complete life cycles during exposure; 2) adult emergence can and must be monitored during testing to interpret test results accurately; 3) drift response to a contaminant can be quantified; and 4) microcosm com-

munities can reflect natural changes due to recruitment and emergence in the source ecosystem, thus allowing field validation of the protocol.

The techniques employed in the present research directly or indirectly measured the impact of a site-specific stressor on the survival, drift behavior, emergence, growth, reproduction, and interactions of indigenous macroinvertebrates. Other ecotoxicological end points such as bioconcentration, bioaccumulation, carcinogenesis, teratogenesis, and histological, physiological, and biochemical responses were beyond the scope of the present investigation. However, these alternative end points could easily be incorporated into the testing protocol. Similarly, the nontaxonomic periphyton analysis employed here could be expanded to include a more detailed evaluation of periphytic response, macroinvertebrate-periphyton interactions etc. The wealth of information that can be gleaned during testing is only limited by the time, expertise, and imagination of the investigators.

Methods

Study area

Several receiving systems for possible use in field validating the protocol were evaluated. Factors considered in evaluating receiving systems included: 1) presence of similar morphometric characteristics (e.g., substrate, flow, depth) at all (upstream and downstream) study sites; 2) presence of a relatively unperturbed reference reach upstream of the point of impact; 3) presence of a consistent impact; and 4) cooperation from the discharging industry. Although several promising sites were identified, cooperation from only one discharging industry could be secured. That company cooperated with the understanding that they would not be identified.

The study was conducted in August, 1987 on a third-order mountain stream receiving a complex effluent. Components of the effluent thought to be major contributors to its in-stream effects included chlorides, ammonia, phenols, and heavy metals including lead and copper. The surrounding watershed is largely forested and sparsely populated. There are no other point source discharges for >20 km upstream.

An upstream control site (UP) was located immediately above the effluent outfall. The first downstream site (DN1) was located ~5 m below the outfall in the effluent plume. The second downstream site (DN2) was located ~0.25 km below the discharge after thorough mixing. A third downstream site (DN3) was established ~1.4 km below the outfall. The fourth downstream site (DN4) was located ~6.4 km downstream, below a convergence with another third-order stream that provides substantial dilution and a new source of colonists. All study sites were located on riffle habitats and were relatively unshaded by riparian vegetation. The substrate at all sites was primarily cobbles (6-25 cm) with a few boulders (> 25 cm) ~25% embedded in pebbles (1-5 cm) and sand (0.5-3.0 mm). During the study period, the average riffle width was ~7 m, except at site DN4 where the width was ~12 m. The average depth was 20 cm.

Effluent for laboratory tests was collected weekly as a grab sample at a point immediately before discharge into the stream. Samples were pumped into 210-L drums and transported to the laboratory for testing.

Stream microcosms

The experimental system consisted of 16 oval artificial streams (1.7 X 0.24 m channel) constructed of molded fiberglass. A standpipe in each stream maintained depth at 13 cm; a 500-μ mesh screen prevented the loss of macroinvertebrates down the standpipe. Streams were dosed in quadruplicate at 0.0, 0.1, 1.0, and 10.0% effluent. Two of the four streams from each treatment were randomly assigned to each of two benches. The benches were identical so they were not used as statistical blocks. Each stream received its dose from a 121.1-L plastic can assigned to that stream. The dosage (10 ml/min) was delivered by peristaltic pumps through pvc tubing. Dilution water (90 ml/min) was supplied to each artificial stream from a headbox system. Current (25 cm/sec) was provided by paddle wheels attached to a 1-cm iron rod that was turned at 30 rpm by a 0.25 hp electric motor. Two 120-cm Durotest Vita-Lites® over each stream provided daylight equivalent light. The photoperiod corresponded to sunrise and sunset on day 15 of the test. Each stream was covered by a 1.00 X 0.75 X 0.30 m emergence trap (mesh size ~1.0 mm). Each stream contained four 60-cm glass tubes (10 mm id) for use

as heat exchangers in controlling stream temperature. The glass tubes for each bench were connected by insulated garden hose (12 mm id) to a 121.1-L reservoir equipped with a chilling unit; submersible pumps allowed circulation through the heat exchangers. Another chilling unit was located in the dilution water headbox. Due to relatively high temperatures in the source stream during this experiment, these cooling systems were not operated.

Derivation of test organisms

Macroinvertebrate communities (175) were developed in rock-filled plastic containers (10.6 X 10.6 X 8.3 cm) with six circular holes (12 mm dia) in each side. The river rock placed in the containers was obtained commercially to ensure a uniform lack of periphytic growth prior to colonization. One small cobble (7-9 cm) was placed in each container; pebbles (3-5 cm) were then added until the containers were full. The artificial substrates were secured on wooden frames that had previously been anchored to the stream bottom with iron rods and cement blocks.

The artificial substrates were allowed to colonize for 30 days in a riffle area ~ 100 m above the point of impact. A review by Rosenberg and Resh (1982) that examined 20 studies containing time series data on macroinvertebrate colonization of artificial substrates suggests that a 30-day colonization period should be sufficient to ensure that macroinvertebrate species equilibrium has been attained. However, maximum densities of some taxa may not be reached during this period (Shaw and Minshall 1980).

Periphyton communities were developed on 10.6 X 10.6 cm high-density, plastic container lids (80) and on polyurethane foam sponges (5 X 6.5 X 7.5 cm). The plastic lids were attached to wooden frames that had previously been anchored to the substrate. The foam sponges were randomly interspersed among the macroinvertebrate artificial substrates. The plastic lid substrates were allowed to colonize for the entire 30-day macroinvertebrate colonization period; while the foam sponges were colonized for 14 days. No sloughing of periphytic growth during handling was noticed for either substrate type.

The periphyton communities developed on the foam sponges were used to seed the artificial streams with a macroinvertebrate food source 3 weeks prior to starting the bioassay. The contents of three sponges were filtered through a 350-μ mesh screen into each stream; the filtering prevented introduction of macroinvertebrates.

After colonization, macroinvertebrates were removed from the stream by placing a dip net (mesh size $\sim$ 350 μ) behind the container as it was transferred to one of 32 coolers (7-L capacity), filled with river water, for transportation back to the laboratory. Two substrates were placed in each cooler and two coolers were used for each stream microcosm. The contents of a cooler could not be divided between two microcosms because some organisms become disassociated from the substrates during transport. Periphyton substrates were also transferred to the stream microcosms in small ice chests filled with stream water. The transportation time back to the laboratory was 4 h on a day when the air temperature reached 35°C. Temperatures in the coolers were maintained at the stream temperature of 20°C by pumping ambient air through a small radiator placed in a cooler of ice; the cooled air was then shunted to each of the coolers and out through an airstone in the cooler.

Experimental chronology/design

On 17 July 1987, the artificial streams were turned on and seeded with periphyton (see above). On 2 August, 16 macroinvertebrate and 8 periphyton substrates were sampled (see sampling below) to provide an estimate of periphyton composition and macroinvertebrate species-abundances initially placed in the stream microcosms. An additional 20 macroinvertebrate and 16 periphyton substrates, after being colonized in the unimpacted reach, were randomly selected and transferred to each of the four downstream field sites for use in field validation of the laboratory results. In addition, a 0.25-m^2 kick-sample was taken from the natural substrate at all five field validation sites, and eight 20.27 cm^2 periphyton samples were taken from the natural substrate at sites UP and DN2. On 3 August, 64 macroinvertebrate and 32 periphyton substrates were randomly selected and transported to the 16 artificial streams (four macroinvertebrate and two periphyton substrates in each). On 5 August, 400 L of effluent was returned to

the laboratory and the stream microcosms were dosed at 0.0, 0.1, 1.0, and 10.0% effluent by pouring an appropriate amount of 100% effluent into each stream. A new batch of effluent was collected on 8 and 15 August. On 17 August, discharge of the effluent was discontinued because of low flow conditions; a fact that was not discovered until 23 August. The lack of discharge forced termination of the experiment 10 days early and resulted in a loss of most of the field-validation component. Therefore, on 25 August, after 20 days of exposure, the macroinvertebrates and periphyton were sampled from the stream microcosms.

Sampling

Macroinvertebrate drift was measured in each of the artificial stream microcosms 1 h after placement of the substrates by inserting a 15 X 12-cm^2 dip net (mesh size ~ 350 μ) into the stream for 1 min. Previous studies indicated that nearly all material drifting in the artificial streams was captured using this technique. Drifting insects were placed in 3 cm of water in an enamel pan, identified to family, counted, and returned to the artificial stream. Drift was also measured 1 h after the streams were dosed. Drift data were used to determine: 1) mortalities during transport; and 2) if a drift response to the stressor is present. Adult insects were collected from each artificial stream with an aspirator every 48-72 h throughout the course of the experiment.

After 20 days, the four macroinvertebrate artificial substrates were removed from the stream microcosms in a dip net and washed through a 500-μ sieve. Additional macroinvertebrates and fish were collected by draining the artificial streams through the dip net. The organisms and debris collected were preserved in labeled jars containing 70% ETOH. Sorting was done by hand, and insects, with the exception of some midges (Chironomidae:Diptera), were identified to the lowest possible taxonomic unit using appropriate references. During identification, each organism was size-classed in a cursory fashion. Hemimetabolous insects were classed, using wing pad development as a criterion, as small, medium, large, or adult. Holometabolous insects were classed, using a previous knowledge of approximate size prior to pupation, as small, medium, large,

pupae, or adult. Each taxa was enumerated, and estimates of the number per stream microcosm were calculated.

A 20.27-cm^2 sample was taken from each of the periphyton substrates by scraping the delimited surface with a bristle brush and filtering the scrapings onto Whatman® GFC filter paper. Determination of chlorophyll *a* concentration was made with one half of the filter paper; chlorophyll *a* was used to evaluate the impact of the contaminant on the standing crop of autotrophs. Biomass (ash-free dry weight) was determined from the other half of the filter paper (American Public Health Association et al. 1985). The Autotrophic Index (Weber 1973) was then calculated:

$$AI = \frac{\text{biomass mg/m}^2}{\text{chlorophyll } a \text{ mg/m}^2}$$

The Autotrophic Index was used to determine the relative response of heterotrophs versus autotrophs to the contaminant. The index value increases as the heterotrophic proportion increases or the autotrophic proportion decreases.

Temperature, dissolved oxygen, pH, hardness, conductivity, current velocity, and other appropriate parameters were monitored in the artificial streams and the natural receiving system. Effluent concentrations were estimated by: 1) regressing conductivity values on known concentrations from serial dilutions of effluent for each effluent batch; and then 2) using the regression equation to predict concentration from conductivity measured in the laboratory microcosms and at the field sites. Both laboratory and field diluent water were used to determine separate regression equations for the laboratory and field estimates.

Degradation of the effluent (i.e., volatilization of phenols) could not be quantified using conductivity measurements, so acute (48 h) toxicity tests (USEPA 1985) with *Daphnia magna* were conducted at the beginning and end of use of each effluent batch. These acute, static tests were conducted using both laboratory and stream water as the diluent to determine if the source of diluent had an effect on the toxicity of the effluent.

Data analysis

The data were generated by determining total insect density (adults and young combined) per taxon in each stream microcosm. Individual substrates could not be used as subsamples because: 1) the organisms colonizing the internal surfaces of the streams and the adults could not be attributed to a given substrate; and 2) for several taxa, densities in substrates closer to the current source were greater than densities in substrates that recieved less current. Macroinvertebrate taxa with mean densities ≥four per stream microcosm for any experimental group were considered a core taxon. The density of individuals in each core taxon compared over all experimental groups was used to determine treatment effects. The data were log-transformed and then analyzed by a one-way ANOVA in conjunction with Duncan's Multiple Range Test for the separation of means. Mean macroinvertebrate densities placed in each artificial stream (INIT) were determined by randomly combining the 16 substrates sampled on 2 August into four, four-substrate groups. These composited substrates were not included in the ANOVA, but the mean and standard error for each core taxon are reported to give an estimate of initial densities.

The quantitative macroinvertebrate data from the *in situ* bioassays were lost when the industry discontinued discharge without notice. Therefore, field validation of the laboratory microcosm results is based on results from the kick-samples taken at each field site on 2 August. However, these samples do accurately reflect the gross differences among the sites during the study period.

Chlorophyll *a*, total biomass, and the Autotrophic Index values from each stream microcosm were also analyzed by a one-way ANOVA in conjunction with Duncan's criterion for the separation of means. Field data were analyzed separately. The data were not transformed prior to analysis.

Results and Discussion

Water/effluent quality

Because of the large quantities needed and the 3.5-h travel time to the field site, dilution water for use in the laboratory stream microcosms could not be transported from the receiving stream, and effluent was usualy collected at weekly intervals. Characteristics of the laboratory diluent, river water at UP and DN2, and the effluent are reported in Table 2-1. Average hardness of the laboratory diluent was twice that of the upstream river water. However, hardness of the laboratory diluent and the upstream river water was variable, overlapped in time, and the differences apparently had no effect on organisms in the control stream microcosms.

The LC_{50}s from the acute toxicity tests with *D. magna* indicated no measureable degradation of the effluent's toxicity over the 1 week storage period, and only a slight difference between the two types of diluent (LC_{50} = 18.3 for laboratory diluent and 22.2 for upstream diluent). Therefore, using laboratory diluent water and storing the effluent should have had little effect on the results of the multispecies toxicity test. However, the acute LC_{50}s for *D. magna*, (using laboratory diluent), ranged from 18.3 to 37.5% concentration for different effluent batches, indicating some between sample variability in effluent toxicity.

Comparison of artificial and natural streams

Environmental parameters measured in the artificial and natural streams are reported in Table 2-2. The experiment was conducted during a period of extreme low flow in the receiving stream, as evidenced by the relatively high temperatures and low current velocities. The differences in field site temperatures reported in Table 2-2 are due to differences in the time of day the site was sampled. Temperature at site UP on 1 August ranged from a low of 15°C at night to 25°C in the afternoon. Microcosm temperatures were set at 18°C and current velocity at 25 cm/sec to better approximate August temperature and flow conditions reported in previous years (unpublished industry data). The drop in dissolved oxygen at field sites DN1-DN3 is probably because of decomposition of organic matter present in the effluent (August mean, whole effluent: suspended solids

Table 2-1. Water quality characteristics of river water (UP = upstream reference reach; DN2 = downstream site 2), laboratory diluent, and whole effluent. Means for the month of August, 1987. Data (except lab diluent) were provided by the discharging industry.

Characteristic	UP	DN2	Diluent	Effluent
pH (units)	7.09	7.11	8.00	6.80
Hardness (mg/L $CaCO_3$)	37.70	44.3	80.00	860.00
Chlorides (mg/L)	3.43	76.2	8.67	3,131.70
NH_4—N (mg/L)	0.39	20.82	0.06	121.30
Total lead (μg/L)	3	3	<1	106

= 217.1 mg/L; BOD_5 = 65.1 mg/L). Dissolved oxygen also decreased slightly in the treated stream microcosms. The pH in the stream microcosms was higher than that in the natural stream, but the difference apparently had no adverse effect on test organisms. The pH of the effluent probably did not contribute to its toxicity. Conductivity and hardness were good indicators of effluent concentrations in both the laboratory microcosms and at the field sites.

Because of slightly increased flows on 17 August, the estimated effluent concentrations at the field sites on that date (Table 2-3) may be conservative relative to concentrations present on 7 August, but relative to normal flows probably represent the "worst case" for this particular discharge. The estimated effluent concentrations in the laboratory microcosms were very close to the nominal concentration targeted for each stream and reflected the ambient concentrations in the receiving stream.

Table 2-2. Mean current velocity, temperature, dissolved oxygen, pH, conductivity, and hardness measured in the stream microcosms on 7 August and at the field sites on 5 August 1987. UP = upstream reference reach; DN1-DN4 = sites downstream of the discharge; EFF = whole effluent; 0.0 = control microcosms; 0.1-10.0 = nominal % effluent concentration in dosed microcosms.

Site or Treatment	Velocity cm/sec	Temp °C	DO mg/L	pH	Cond µS/cm	Hardness mg/L $CaCO_3$
UP	13.0	25.0	7.4	6.5	160	20
DN1	11.4	24.0	5.2	6.7	1600	125
DN2	19.6	23.0	4.3	6.9	710	55
DN3	15.1	23.0	4.3	6.7	710	60
DN4	18.6	23.0	7.0	7.2	410	45
EFF	*	24.0	0.3	6.6	6000	510
0.0	25.0	18.0	8.6	8.2	123	45
0.1	25.0	18.0	8.5	8.2	130	50
1.0	25.0	18.0	8.2	8.1	193	55
10.0	25.0	18.0	8.2	8.1	688	85

Periphyton

The benthic microorganisms present downstream of the outfall were typical of a stream receiving large amounts of organic wastes. Thick mats of "sewage fungus" (i.e., *Sphaerotilus natans*) covered the substrate at site DN1. Within 1 week, these growths also covered the periphyton and macroinvertebrate substrates transferred to DN1 from UP at the beginning of the test. This station also supported large numbers of a few heterotrophic protozoan species including *Paramecium caudatum*, *Epistylis* sp., and *Loxocephalus pagius*. Site DN2 also contained these protozoans and traces of "sewage fungus" were present, but this site was primarily characterized by thick mats of filamentous green algae (possibly *Cladophora* sp.). Most of the planned analyses on the periphytic communities above and below the outfall were lost when the industry halted discharge on 17 August. By 23 August, there was no visible evidence of the "sewage fungus" community present at DN1.

Periphyton communities sampled from artificial substrates and the natural substrate on 2 August were not significantly different in terms of chlorophyll *a*, ash-free dry weight, or the Autotrophic index (Table 2-4), indicating that the artificial substrates

Table 2-3. Effluent concentrations (% effluent) on 15 August estimated from conductivity measurements on serial dilutions of effluent. Field site estimates ($r^2 > 0.99$) were determined using diluent from the upstream reference reach; stream microcosm estimates ($r^2 > 0.99$) were from dilutions with laboratory diluent. UP = upstream reference site; DN1-DN4 = sites downstream of the discharge; 0.0 = control microcosms; 0.1-10.0 = nominal % effluent concentration in dosed microcosms; Effluent = 3rd batch, 100%.

Site	Conductivity µS/cm	Estimated Concent	Treatment	Conductivity µS/cm	Esti-mated Concent.
UP	80	0.0	0.0	131	0.00
DN1	2690	14.1	0.0	132	0.00
DN2	820	4.1	0.0	135	0.00
DN3	710	3.5	0.0	134	0.00
DN4	260	1.1	0.1	150	0.10
Effluent	19000	*	0.1	149	0.10
			0.1	155	0.13
			0.1	160	0.16
			1.0	340	1.09
			1.0	330	1.05
			1.0	360	1.20
			1.0	350	1.15
			10.0	1910	9.32
			10.0	1900	9.27
			10.0	2120	10.40
			10.0	2020	9.9

are suitable analogues for use in the stream microcosms. The significantly higher chlorophyll *a* and ash-free dry weight values found on the natural substrate at DN2 relative to UP reflect the obvious visible differences between these two sites. The Autotrophic Index values were significantly lower at site DN2 relative to the UP values because of the large autotrophic community present at DN2. Due to the large amounts of "sewage fungus" at DN1, the Autotrophic Index would probably have been much higher at that site.

Results from the stream microcosms indicate that overall biomass and chlorophyll *a* were elevated in all treated streams and were significantly higher in the 1.0 microcosms (Table 2-4). Therefore, periphytic response in the stream microcosms would have predicted an increase in chlorophll *a* and total biomass at effluent concentrations

Table 2-4. Mean chlorophyll *a*, biomass (ash-free dry weight) and Autotrophic Index values from field sites and laboratory microcosms. A separate ANOVA was run for field and laboratory data. Treatments within an ANOVA with the same letter are not significantly different (Duncan's $P \leq 0.05$). UP = upstream reference site—natural substrate; DN2 = downstream site two— natural substrate; AS = upstream reference site—artificial substrate; 0.0 = control microcosms; 0.1-10.0 = nominal % effluent concentration in dosed microcosms.

Site or Treatment	Chlorophyll *a* mg/m²	AFDW mg/m²	Autotrophic Index
Field Sites			
UP	29.3 b	9951.0 b	341.3 a
AS	18.5 b	6960.8 b	380.3 a
DN2	124.2 a	18137.3 a	147.0 b
P value	<0.0001	<0.0001	<0.0001
Laboratory Microcosms			
0.0	71.8 b	8774 c	124.0
0.1	109.4 ab	10637 bc	107.6
1.0	142.8 a	16667 a	130.2
10.0	104.7 ab	13186 ab	130.1
P value	0.0234	0.0021	0.6884

near 1.0%. Not only was this trend observed in the field, but in fact, the chlorophyll *a* and ash-free dry weight values found on the natural substrate at DN2 compare quite favorably with those obtained from the artificial substrates in the 1.0 stream microcosms. However, "sewage fungus" was not observed even in the 10.0 stream microcosms, so the test would not have predicted the thick growths of "sewage fungus" that occurred at site DN1.

Macroinvertebrates

Other than insects, the only macroscopic organisms regularly observed in either field or laboratory samples were Decapoda and a small darter (*Etheostoma* sp.:Percidae). These organisms colonized the macroinvertebrate substrates in numbers insufficient to allow a statistical analysis of treatment effects. However, these crayfish

and darters were present in three of four samples from the high effluent microcosms and at all field sites, suggesting that they are relatively tolerant of the effluent.

No organisms were found in any of the drift samples taken 1 h after placing the colonized substrates in the artificial streams, suggesting that the transfer from the natural to the artificial streams was accomplished without measurable mortality. Surprisingly, only three chironomids were found in the drift samples taken 1 h after dosing. These chironomids were all collected from one microcosm that received the high effluent treatment.

Thirty-five insect taxa were collected from the stream microcosms at the end of the experiment. However, because of the lack of either larval (e.g., *Caenis* spp.) or adult (e.g., *Baetis* spp. and female Heptageniidae) keys to species, some taxa were analyzed at the generic level. Fifteen of these taxa were considered to be core taxa (mean density$\geq$4 per microcosm), and the following analysis is based on their response to the effluent.

The mayfly (Ephemeroptera) *Paraleptophlebia swannanoa* (Leptophlebiidae) colonized the artificial substrates in relatively low numbers (Fig. 2-1). These nymphs were all early instars and no adults were collected during testing. This species was present in the receiving stream from March to October (unpublished data) but was not found in the kick-samples from 2 August. Densities in the control and 0.1% effluent treatment were maintained at or above initial densities. However, density in the 1.0% effluent treatment was significantly lower than that in the control, and no *P. swannanoa* survived the 10.0% effluent treatment.

Based on differences in color pattern, the mayfly genus *Pseudocloeon* (Baetidae) appeared to be represented by two species. However, color pattern is variable within species in this genus, and a reliable key to species was not available. No adults were collected, but some organisms were nearing emergence. The estimated density initially placed in each stream microcosm was very low (Fig. 2-2) as were densities on the natural substrate, suggesting that recruitment from eggs (or nymphs<500 μ) may have oc-

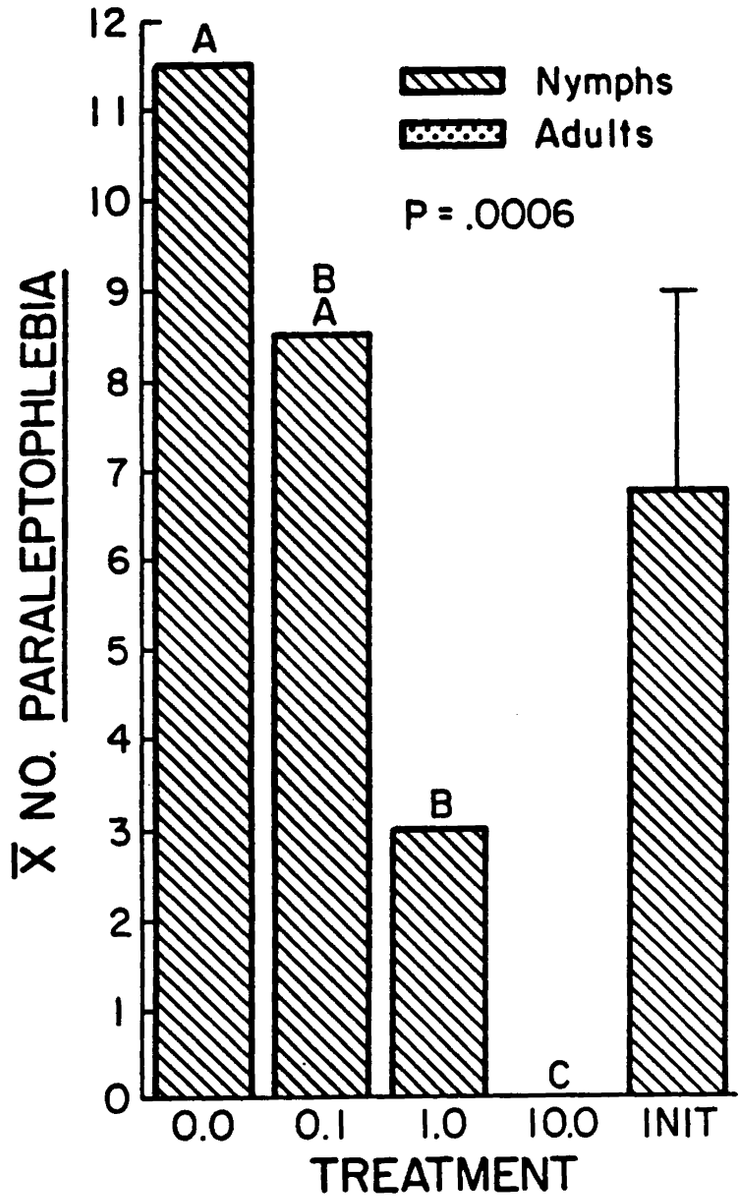

Figure 2-1. Mean number of *Paraleptophlebia* (Ephemeroptera:Leptophlebiidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

curred in the control and 0.1 microcosms. *Pseudocloeon* spp. were unable to survive in the stream microcosms receiving 1.0 or 10.0% effluent.

The mayfly genus *Stenonema* (Heptageniidae) exhibited the highest mayfly densities in the field and the stream microcosms (Fig. 2-3). This genus was represented by at least two core species, *S. modestum* and *S. luteum*. However, because of the small size of most individuals and the lack of keys for adult females, results are reported at the generic level. The mean number of organisms collected in the control and 0.1 microcosms was nearly identical and much higher than initial densities, again suggesting that recruitment from eggs (or nymphs<500 μ) occurred during the course of the experiment. Significantly fewer organisms survived the 1.0 and 10.0 treatments. The reduction in numbers in the 1.0 treatment occurred over all size classes for both *S. modestum* and *S. luteum*, so survival was apparently dependant on individual fitness, not on species or size. However, early instar insects are often more sensitive to stressors than later instars of the same species (Maki *et al.* 1975; Clubb *et al.* 1975), so it is important to size class individuals during identification. Both *S. modestum* and *S. luteum* were present in relatively high numbers in the kick-sample from site UP, were absent from sites DN1-DN3, and were present in reduced numbers at site DN4 (Table 2-5). Given the estimated effluent concentrations at the field sites, this distribution is exactly what would have been predicted from the microcosm test results.

The mayfly genera *Leucrocuta* sp. and *Stenacron* sp. (Heptageniidae) both exhibited a graded response to the treatments similar to that of the *Stenonema* species, although the lower density of *Stenacron* in the 1.0 treatment is not statistically significant (Fig. 2-4 and 2-5). In addition, if adults are not included in the analysis, *Leucrocuta* also does not exhibit a significant response in the 1.0 treatment. Both *Leucrocuta* and *Stenacron* were present at site UP but absent or present in reduced numbers at sites DN1—DN4 (Table 2-5). Once again, the microcosm responses accurately reflect the distributions of these organisms in the field.

The mayfly genus *Baetis* was represented by two species *B. intercalaris* and *B. amplus*. These species colonized the artificial substrates in relatively low numbers (Fig.

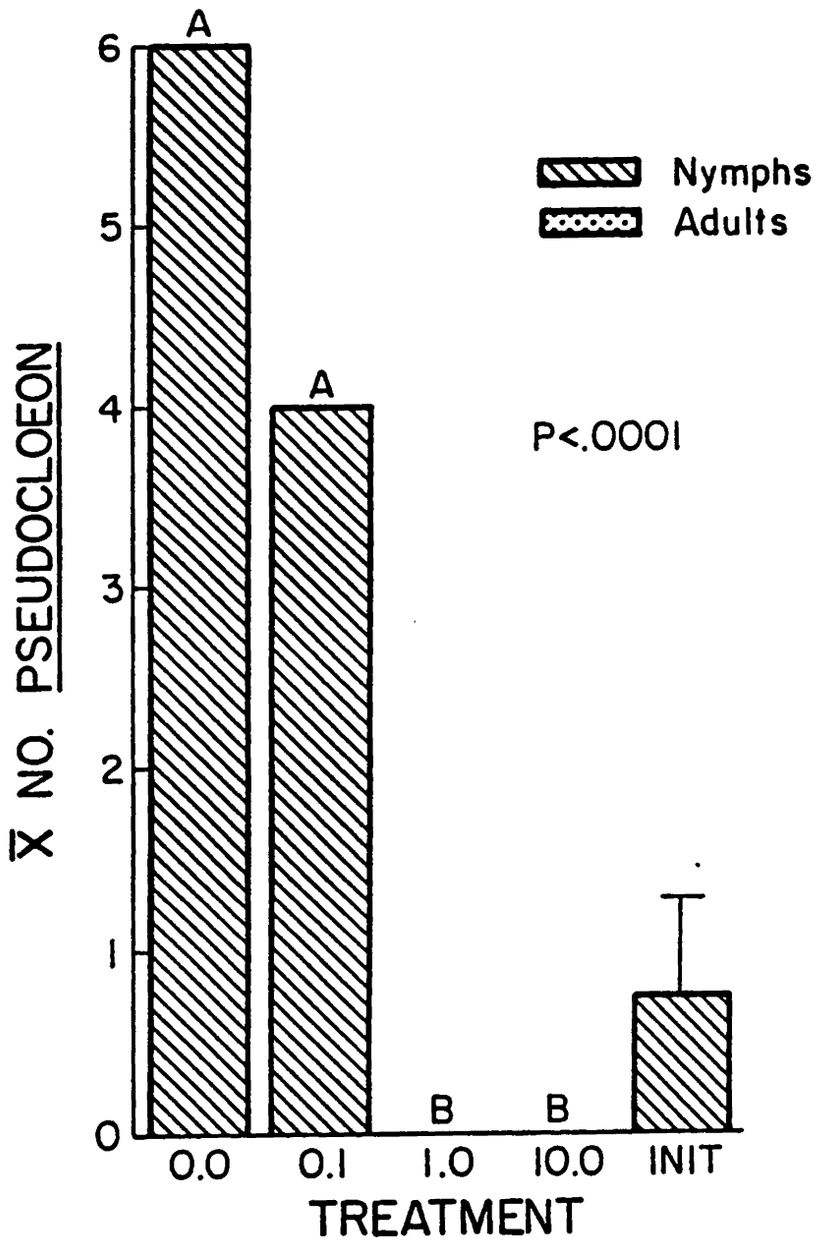

Figure 2-2. Mean number of *Pseudocloeon* (Ephemeroptera:Baetidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

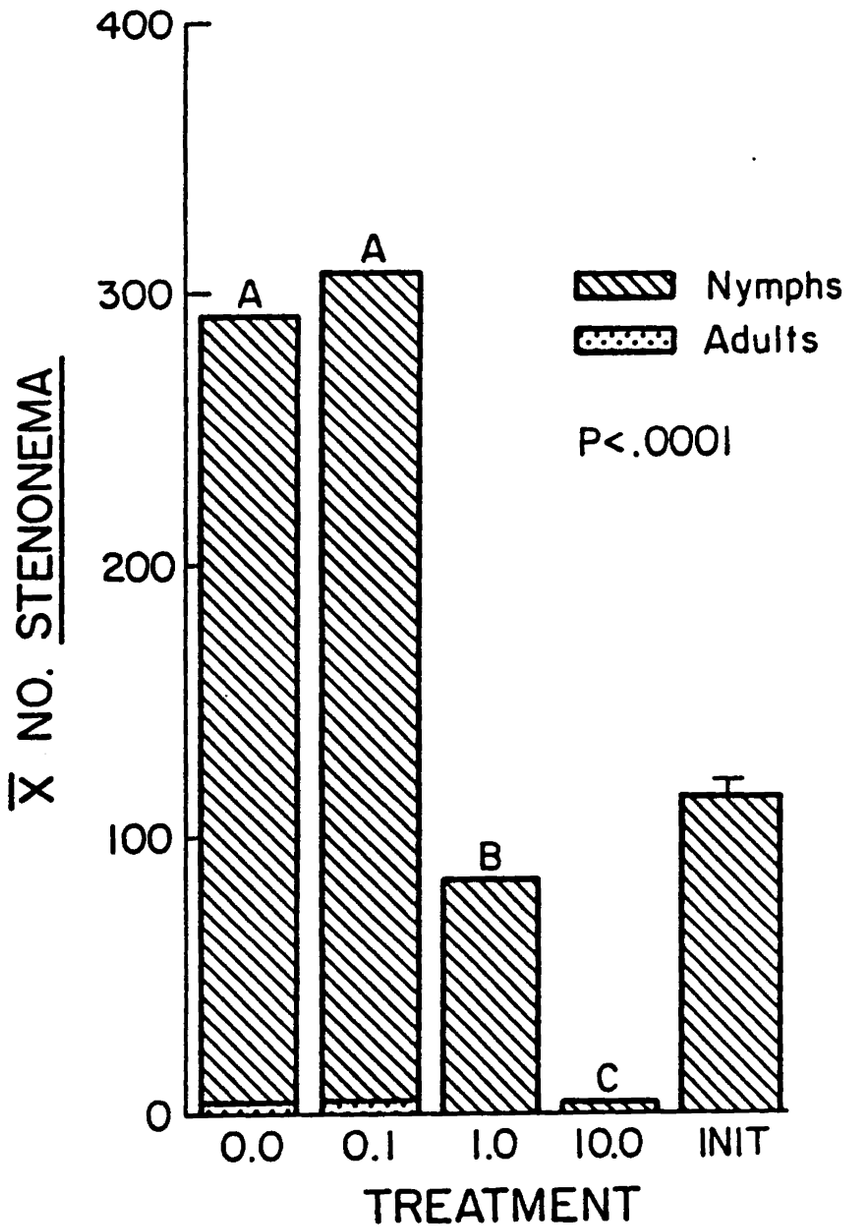

Figure 2-3. Mean number of *Stenonema* (Ephemeroptera:Heptageniidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

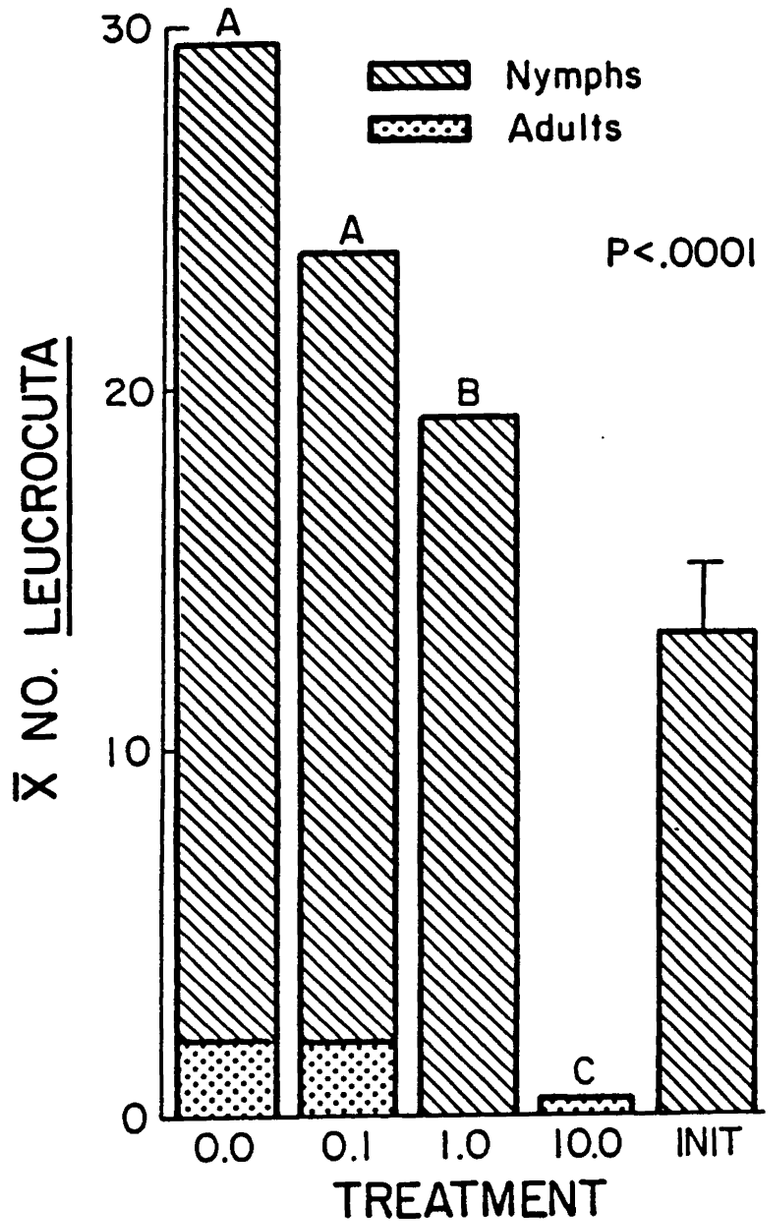

Figure 2-4. Mean number of *Leucrocuta* (Ephemeroptera:Heptageniidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

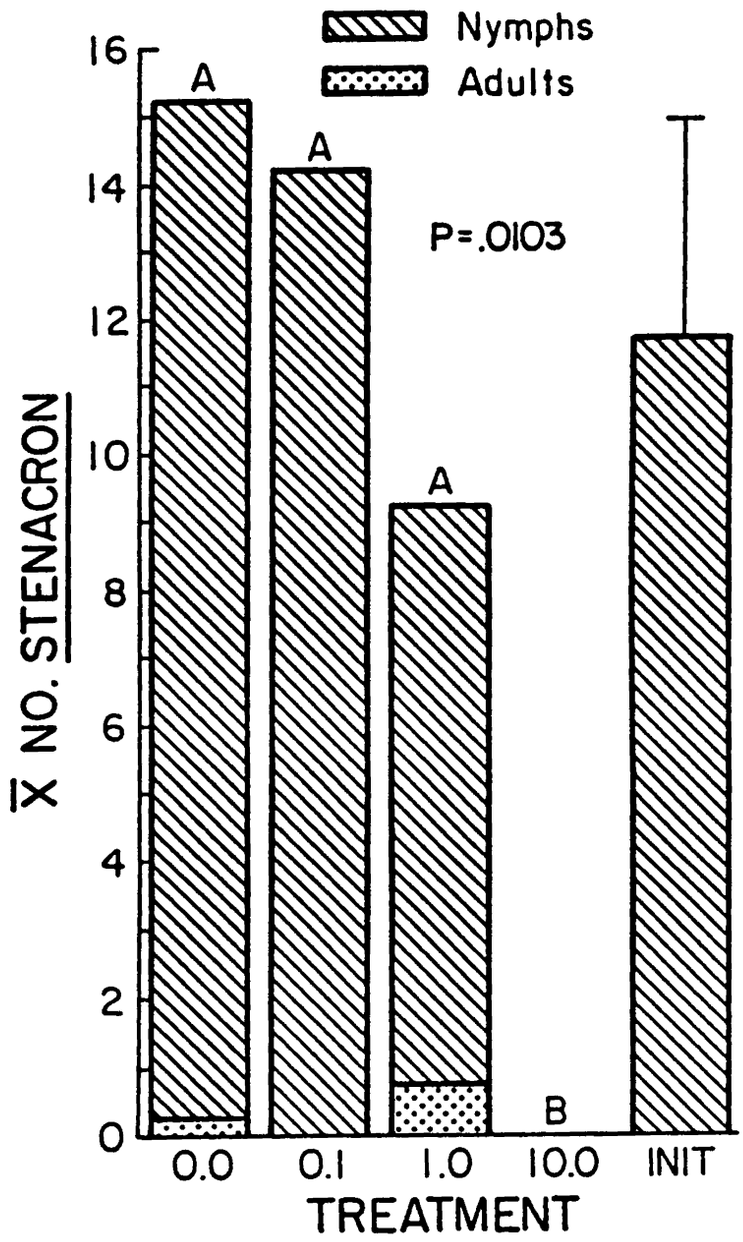

Figure 2-5. Mean number of *Stenacron* (Ephemeroptera:Heptageniidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

Table 2-5. Results from the kick-samples taken at each field site. UP = upstream reference site; DN1-DN4 = sites downstream of the discharge. Only taxa with densities ≥5 from at least one site are included as individual taxa.

Taxon	UP	DN1	DN2	DN3	DN4
Total Taxa	16	9	14	17	27
Taxa in Common with UP	16	5	7	11	14
Total Insects	115	280	2600	484	250
% Chironomids	33	98	95	85	30
% Mayflies	55	0	0	1	32
Isonychia	11	0	0	2	34
Stenonema	44	0	0	0	24
Stenacron	11	0	2	2	6
Leucrocuta	9	0	0	2	3
Baetis	1	0	0	0	12
Caenis	6	0	0	2	2
Optioservus	2	0	4	4	34
Hydropsyche	4	2	100	48	32
Cheumatopsyche	2	0	12	8	6
Brachycentrus	0	0	16	0	4
Chironominae	18	208	2304	288	42
Orthocladiinae	0	64	64	16	2
Tanypodinae	3	0	0	16	4
Atherix	6	1	76	96	22
Hemerodromia	0	0	12	2	0
Gomphus	1	1	6	2	4

2-6), probably because of low numbers on the natural substrate at site UP (Table 2-5). *Baetis* density was slightly reduced in the 1.0 treatment, and only a few adults collected during the first week of the experiment survived the 10.0 treatment. Maki *et al.* (1975) and Federle and Collins (1976) also observed an increase in insect mortalities with longer exposures to various pesticides. Although in this case the numbers of emergent adults in the 0.1 and 1.0 treatments were not significantly different from the control, the techniques used in this experiment could have detected a toxicant-induced inhibition (c.f., Clubb et al. 1975) or stimulation (c.f., Maki et al. 1975) of emergence. Twenty-five years ago, Surber and Thatcher (1963) suggested that the monitoring of adult emergence is necessary during bioassays with aquatic insects, but their advice is often not heeded. If adult emergence had not been monitored in this investigation, results would have in-

dicated significantly ($P \leq 0.05$) lower *Baetis* densities in the control microcosms than in the 0.1 microcosms, and no significant difference between the control and 10.0 microcosms. In a previous experiment, a much higher percentage of *Baetis* emerged during a 30-day experiment (Pontasch and Cairns, in press), indicating the necessity of collecting adults during the course of ecological or toxicological experimentation with this widespread genus.

Only the 10.0% effluent treatment had an effect on the mayflies *Caenis* sp. and *Isonychia bicolor* (Fig. 2-7 and 2-8). However, based on the distribution of these organisms in the natural stream (Table 2-5), they apparently cannot tolerate effluent concentrations much higher than 1.0%. Although the low density of *Caenis* at site DN4 is surprising, the response of these taxa correlates well with the observed response in the field.

The control stream microcosms were able to maintain all mayfly taxa at or above initial levels during this 20-day experiment. As expected, the mayflies as a group were relatively intolerant of the effluent. Four of the mayfly taxa had significantly lower densities at 1.0% effluent concentration, while four other taxa were affected by the 10.0% concentration. Some of the more sensitive mayflies had nonsignificant decreases in density in the 0.1 microcosms. Two other hemimetabolous orders of insects, Plecoptera and Odonata, did not colonize the artificial substrates in densities great enough to allow statistical analysis of their response. However, several *Acroneuria* sp. (Perlidae:Plecoptera) were collected from the 10.0% effluent streams at the end of the experiment, and one was found in the kick-sample from site DN1, so *Acroneuria* sp. may be relatively tolerant of the effluent. In addition, *Gomphus* sp. (Gomphidae:Odonata) was found at all field sites (Table 2-5), indicating that it also may be relatively tolerant.

Hydropsyche morosa and *Cheumatopsyche* spp. (Hydropsychidae:Trichoptera) were the only caddisflies that colonized the artificial substrates in numbers sufficient to permit an evaluation of their response to the effluent (Fig. 2-9 and 2-10). Both of these hydropsychids are apparently quite tolerant of the effluent. The much higher numbers of *H. morosa* and another species, *H. sparna*, at sites DN2—DN4 (Table 2-5) are probably

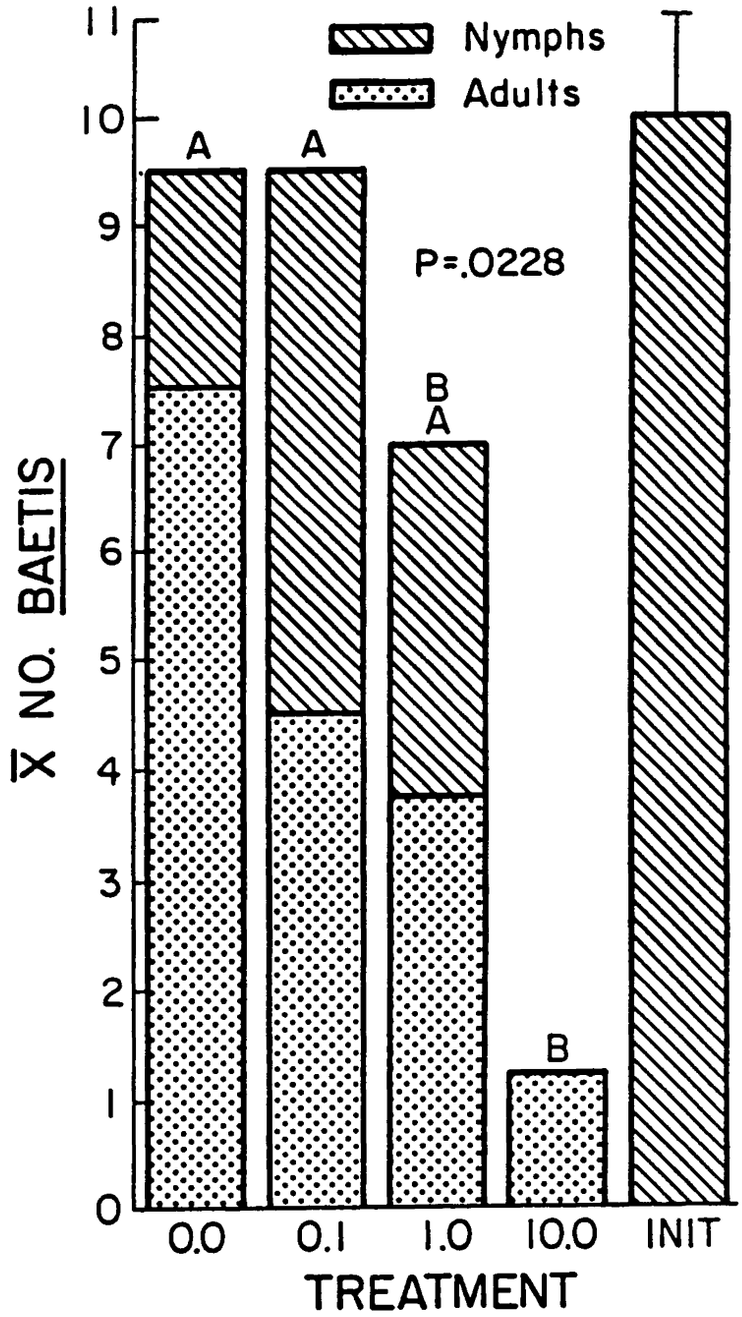

Figure 2-6. Mean number of *Baetis* (Ephemeroptera:Baetidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

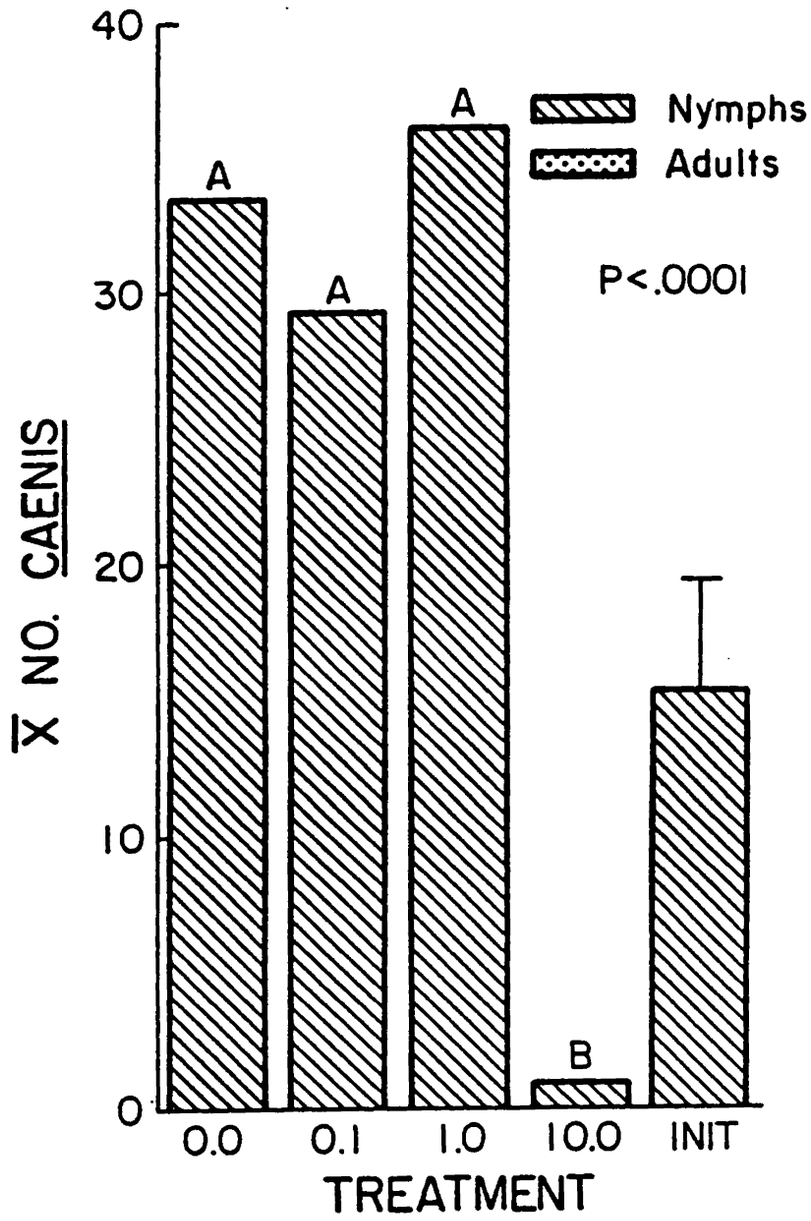

Figure 2-7. Mean number of *Caenis* (Ephemeroptera:Caenidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

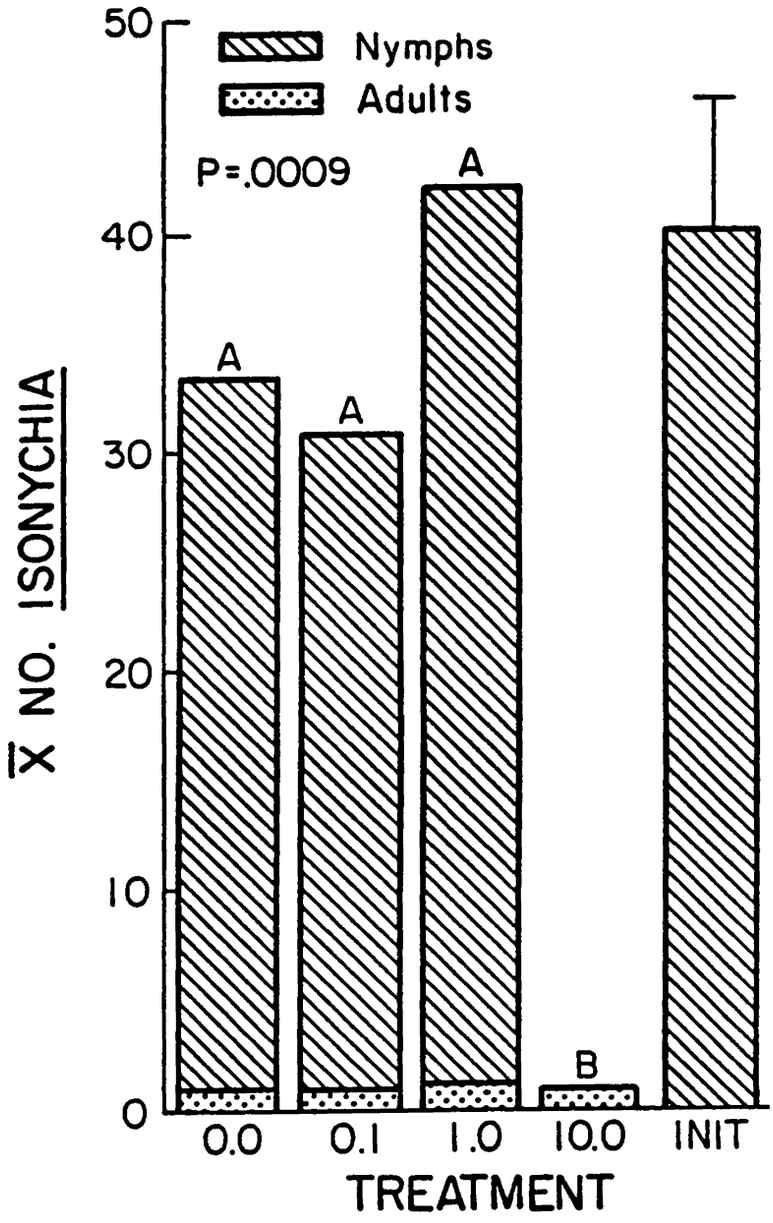

Figure 2-8. Mean number of *Isonychia* (Ephemeroptera:Oligoneuriidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

because of an increase in food supply at these downstream sites. Apparently, these filter-feeders were either feeding directly on the organic matter present in the effluent or on other taxa (e.g., see Chironomidae below) that were present in great numbers at these impacted sites. The stream microcosms were unable to reflect the high numbers present at sites DN2—DN4 because the length of the test did not allow recruitment of these relatively slow growing species. Similarly, the higher density at site DN2 of *Brachycentrus* sp. (Brachycentridae), another collector-filterer, was not reflected by the stream microcosms. In a previous experiment, the hydropsychids colonized the artificial substrates in very large numbers (INIT > 700) but only 30% of this density could be maintained over the 30-day experiment (Pontasch and Cairns, in press). In this experiment, even though initial densities were much lower and the artificial streams were seeded with periphyton 2 weeks earlier, *H. morosa* still could not be maintained at initial densities. The cause of this problem is probably related to the quantity or quality of food available in the seston of the stream microcosms. However, the algal, detrital, and animal food base (*c.f.*, Coffman *et al.* 1971) appeared to be sufficient, especially in the treated microcosms where periphyton and chironomids (see below) were stimulated.

The three main subfamilies of the Chironomidae (Diptera), Chironominae, Orthocladiinae, and Tanypodinae, were all present in relatively high numbers in the stream microcosms (Fig. 2-11, 2-12, and 2-13) The Chironominae were particularly abundant on the artificial substrates and were maintained at their initial densities in the control streams. Over 95% of the Chironominae were *Paratanytarsus* sp.; a parthenogenetic species with a life cycle of approximatey 13 days (USEPA, unpublished data). The significantly higher densities of *Paratanytarsus* sp. in the stream microcosms receiving effluent corresponded to the much higher densities of this insect at sites downstream of the effluent outfall (Table 2-5). Effluent stimulation of their periphytic food source (see above) probably caused the increases in *Paratanytarus* sp. densities. The ability to detect such ecological interactions is one advantage multispecies tests have over more conventional single species tests. The higher numbers of Orthocladiinae present at sites downstream of the outfall would also have been predicted by the higher

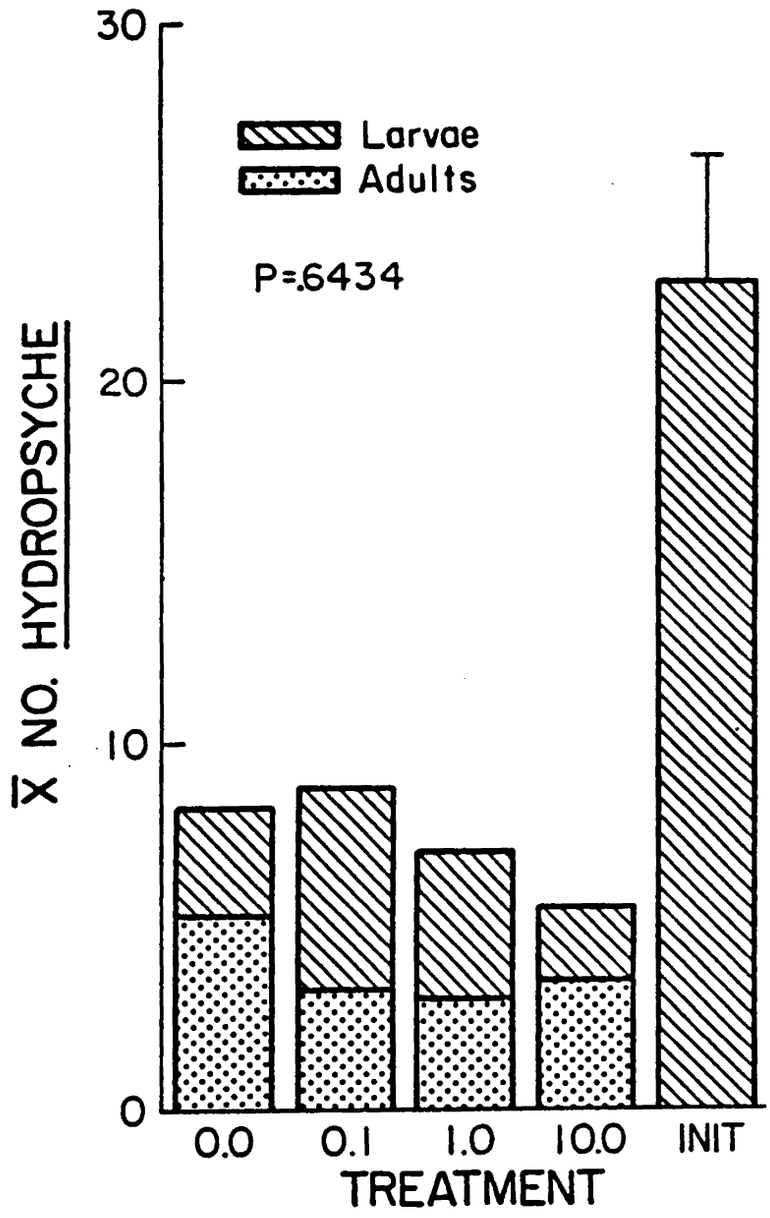

Figure 2-9. Mean number of *Hydropsyche* (Trichoptera:Hydropsychidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included).

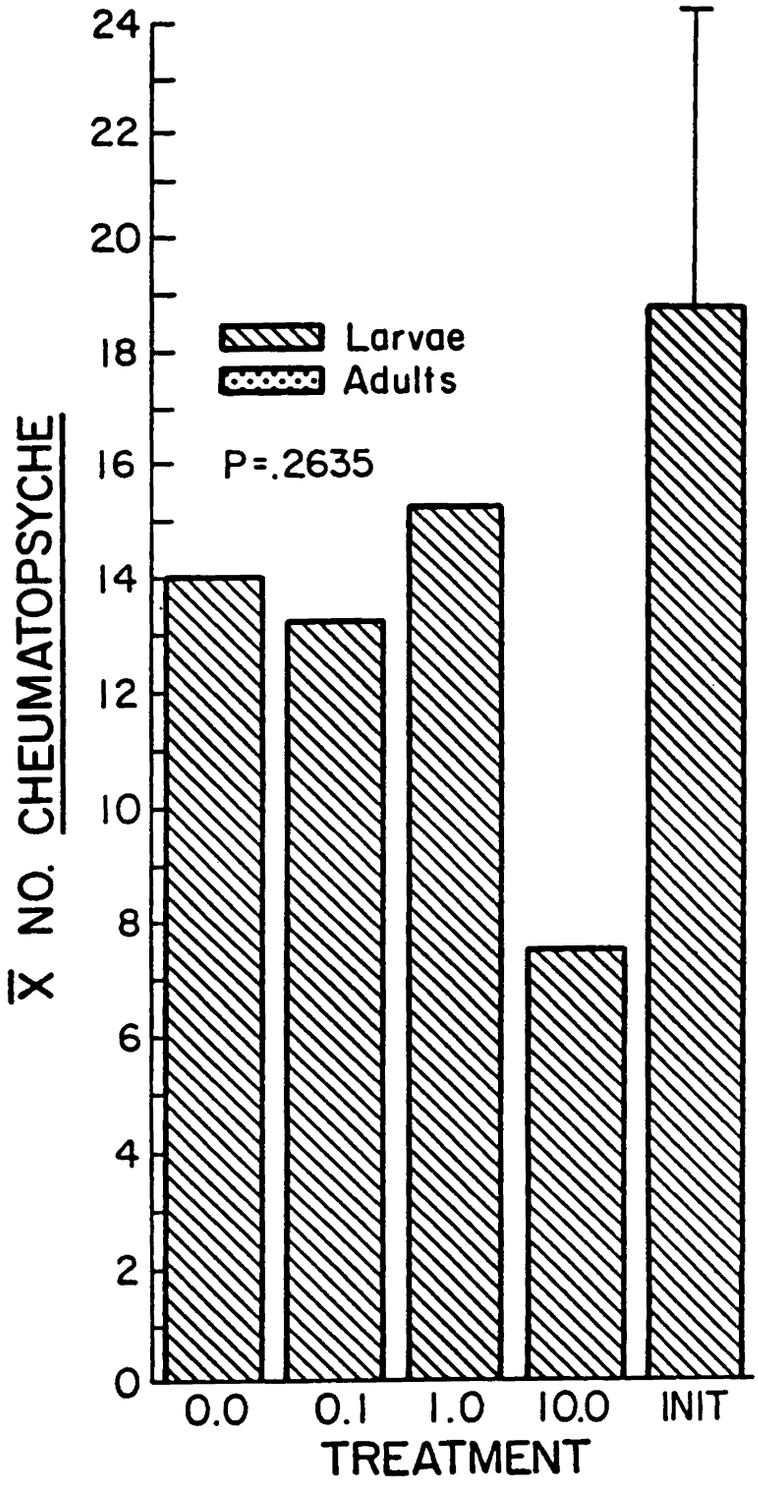

Figure 2-10. Mean number of *Cheumatopsyche* (Trichoptera:Hydropsychidae) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included).

numbers of Orthocladiinae in the stream microcosms receiving effluent. The data suggest that most, if not all, Orthocladiinae initially placed in the stream microcosms emerged during the course of the experiment. Therefore, the larvae collected at the end of the experiment probably either hatched from eggs carried in on the artificial substrates, entered as larvae <500 μ, or hatched from eggs oviposited by adults during the experiment. If emergence traps had not been in place, the trend towards higher numbers may have been obscured due to adults from treated streams ovipositing in the control streams. The data for Tanypodinae was rather variable both in the stream microcosms and the field. Although Tanypodinae also tended to increase in numbers in the treated microcosms, this subfamily was not maintained at initial levels in the control or 0.1 treatment.

Two other families of Diptera, Empididae and Athericidae, were represented respectively by *Hemerodromia* sp. and *Atherix lantha*. The larvae of both of these species are predacious, and their increase in numbers at DN2 (Table 2-5) is probably a response to the large number of chironomid prey available at that site. In addition, *Hemerodromia* adults were regularly observed feeding on adult *Paratanytarsus* in all stream microcosms. Both species were tolerant of the effluent concentrations used in the stream microcosms (Fig. 2-14 and 2-15).

Summary

The stream microcosms were able to maintain most taxa at or above initial densities. Depending on species, mayflies in the microcosms exhibited significant decreases in density at 1.0 or 10.0% effluent, although some species had nonsignificant decreases in the 0.1% concentration. Mayflies responded at much lower concentrations than acute tests with *D. magna* would have predicted. Hydropsychids were not affected by the effluent, and chironomids and periphyton were stimulated. Overall, the stream microcosms accurately predicted macroinvertebrate and periphyton response in the field. Although stimulation of periphyton and some chironomids may occur, the microcosm results suggest that a concentration of 0.1% effluent in the receiving stream

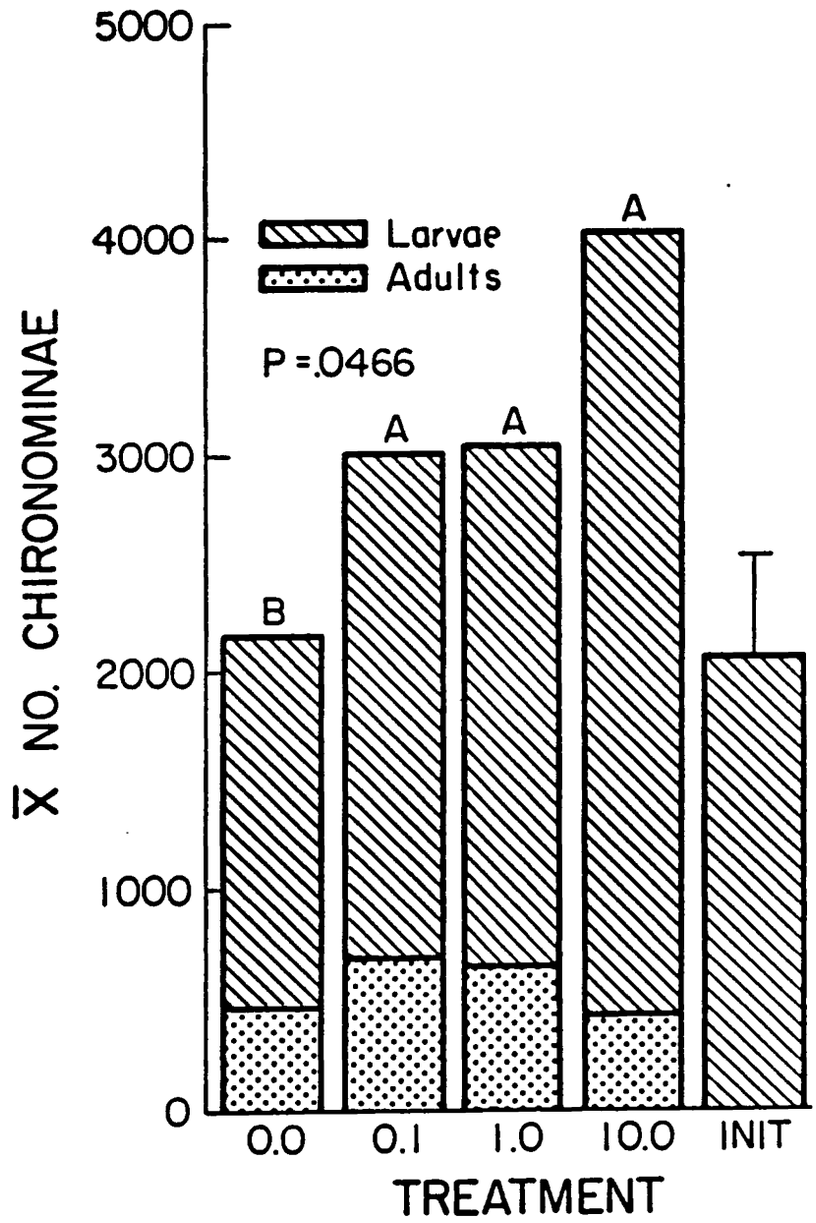

Figure 2-11. Mean number of Chironominae (Diptera) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

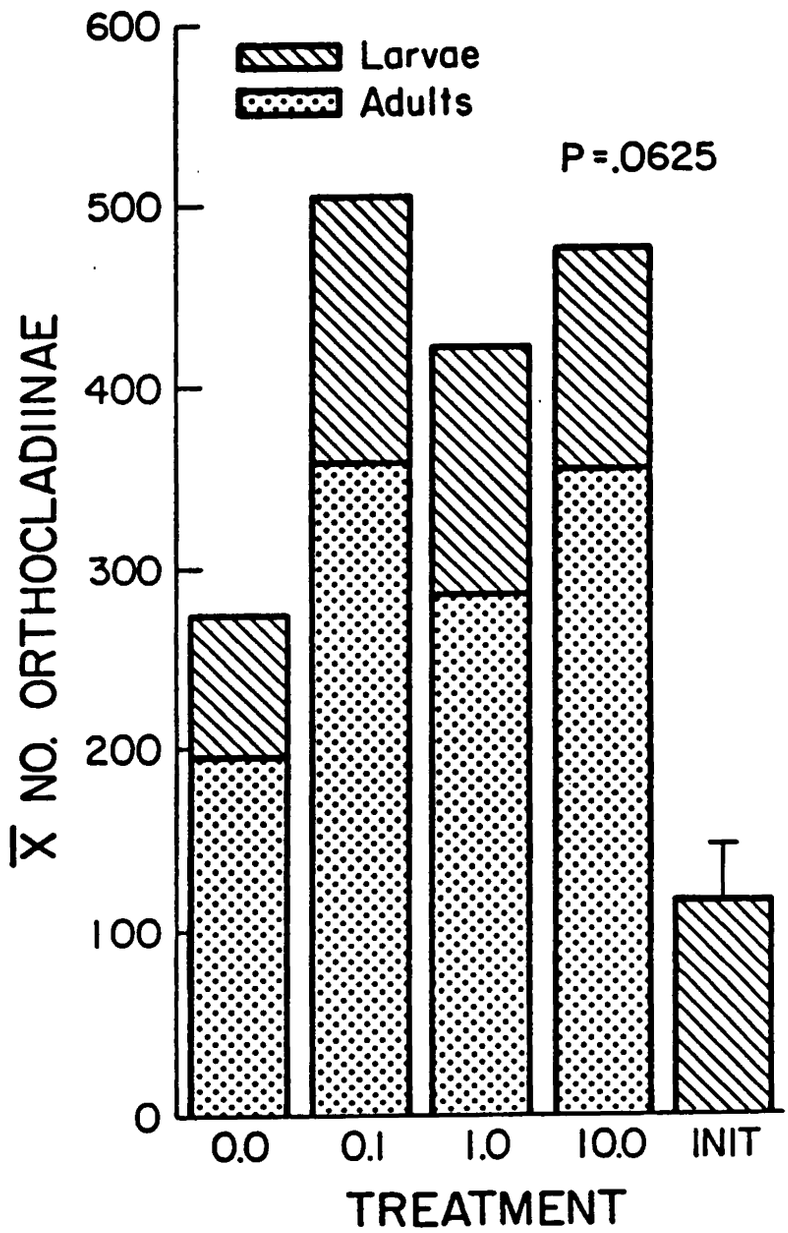

Figure 2-12. Mean number of Orthocladiinae (Diptera) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included).

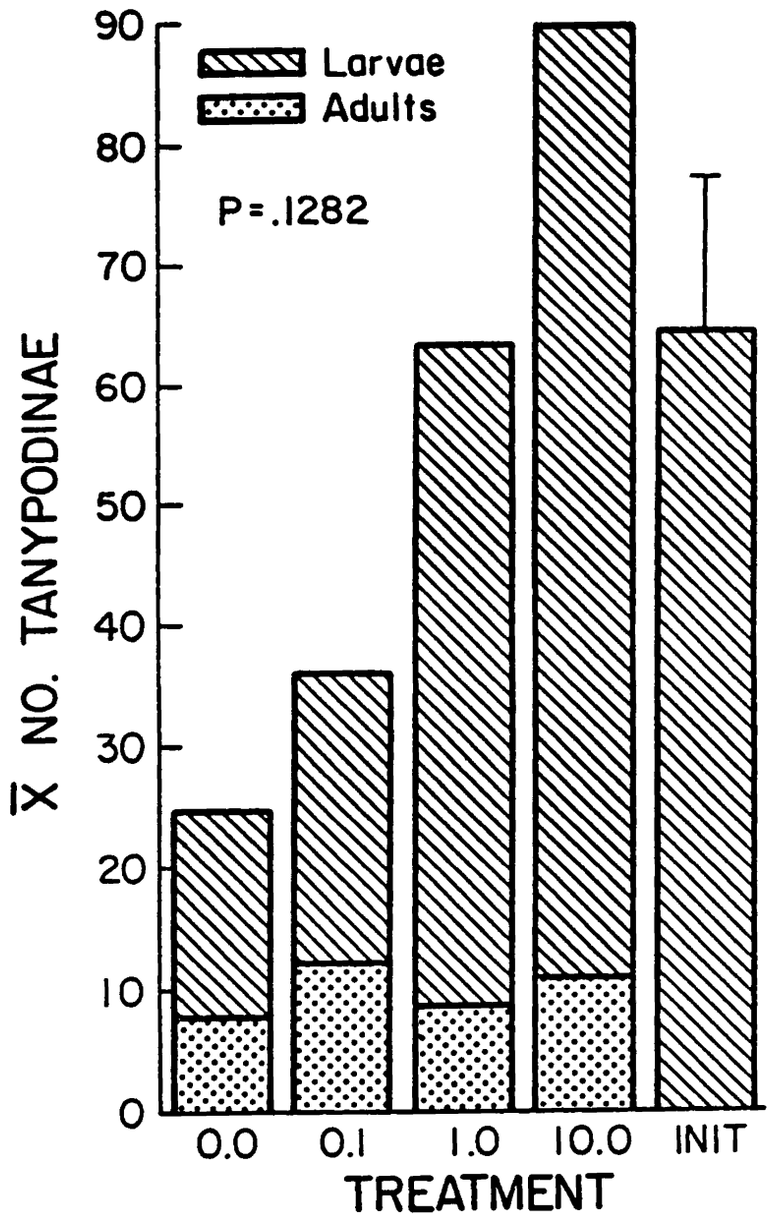

Figure 2-13. Mean number of Tanypodinae (Diptera) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included).

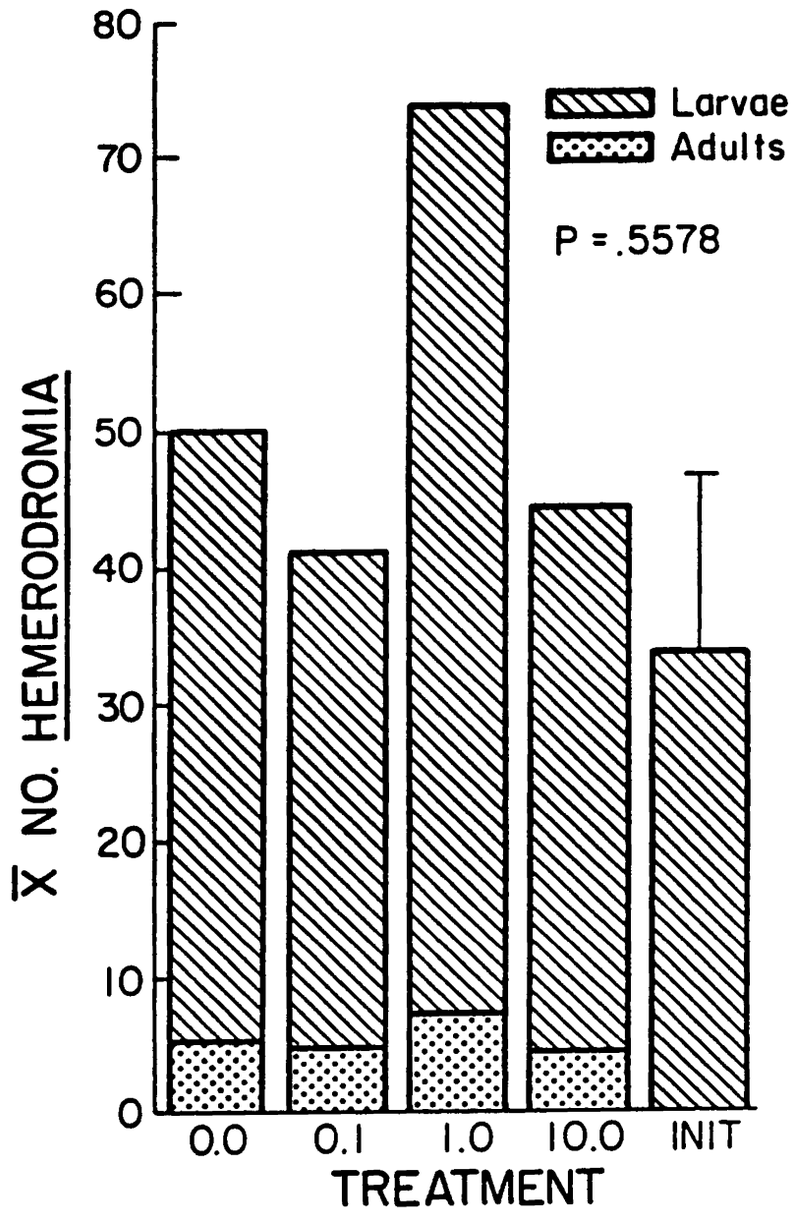

Figure 2-14. Mean number of *Hemerodromia* (Empididae:Diptera) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included).

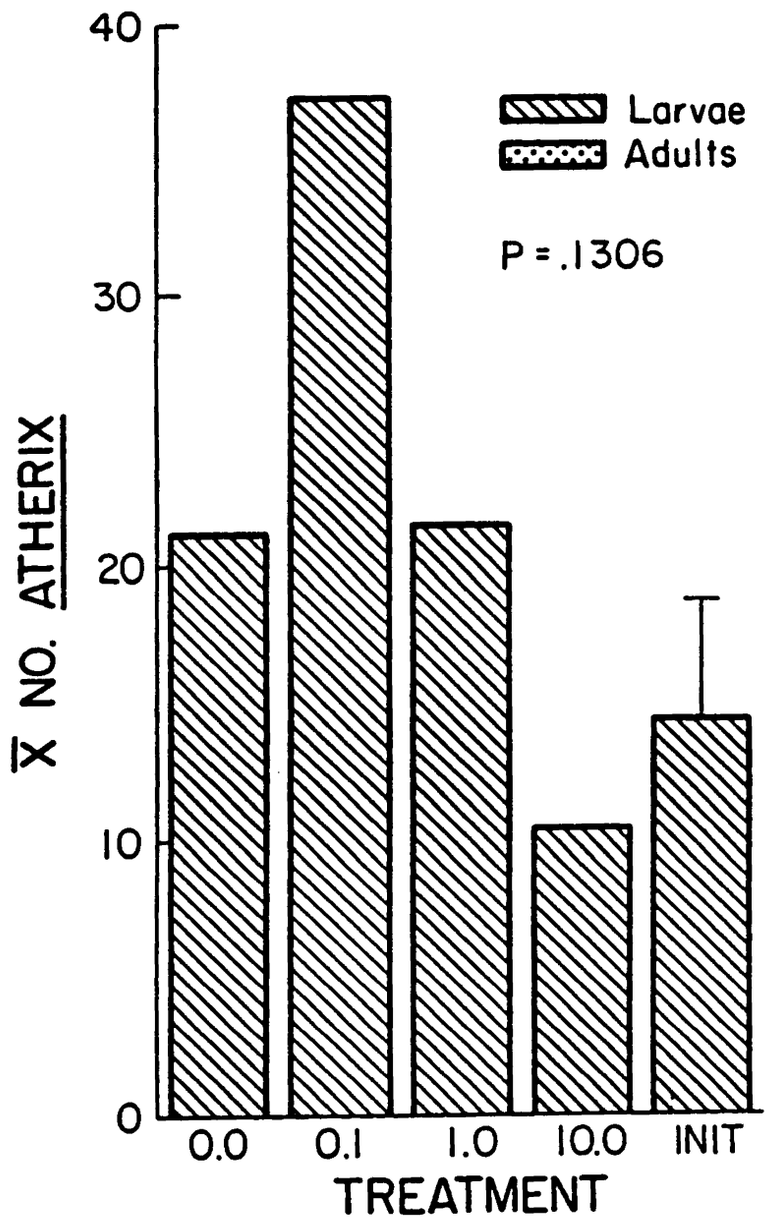

Figure 2-15. Mean number of *Atherix* (Athericidae:Diptera) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included).

should produce no adverse effects on the indigenous biota. Further research should be undertaken in an effort to develop standardized protocols for conducting multispecies toxicity tests with lotic macroinvertebrates and periphyton. These protocols would provide environmentally realistic and scientifically justifiable alternatives to conventional tests and would greatly reduce the amount of uncertainty in environmental hazard assessment.

References

Allard, M., and G. Moreau. 1987. Effects of experimental acidification on a lotic macroinvertebrate community. Hydrobiologia 144:37-49.

American Public Health Association, American Water Works Association, and Water Pollution Control Federation. 1985. Standard methods for the examination of water and wastewater. 16th ed. Washington, DC. 1268 p.

Burks, S. L., and J. L. Wilhm. 1977. Bioassays with a natural assemblage of benthic macroinvertebrates, p. 127-136. *In* F. L. Mayer, and J. L. Hamelink [ed.] Aquatic toxicology and hazard evaluation, STP 634. American Society for Testing and Materials, Philadelphia, PA.

Cairns, J., Jr. [ed.] 1985. Multispecies toxicity testing. Pergamon Press, New York, NY. 261 p.

Cairns, J., Jr. [ed.] 1986. Community toxicity testing, STP 920. American Society for Testing and Materials, Philadelphia, PA. 350 p.

Cairns J., Jr., and J. R. Pratt. 1987. Ecotoxicological effect indices: a rapidly evolving system. Water. Sci. Tech. 19:1-12.

Clubb, R. W., A. R. Gaufin, and J. L. Lord. 1975. Acute cadmium toxicity tests upon nine species of aquatic insects. Environ. Res. 9:332-341.

Coffman, W. P., K. W. Cummins, and J. C. Wuycheck. 1971. Energy flow in a woodland stream ecosystem: I. Tissue support trophic structure of the autumnal community. Arch. Hydrobiol. 68:232-276.

Federle, P. F., and W. J. Collins. 1976. Insecticide toxicity to three insects from Ohio ponds. Ohio J. Sci. 76:19-24.

Giesy, J.P. [ed.] 1980. Microcosms in ecological research. U.S. Department of Energy, Washington, DC. 1110 p.

Giesy, J. P., and P. M. Allred. 1985. Replicability of aquatic multispecies test systems, p. 187-247. *In* J. Cairns, Jr. [ed.] Multispecies toxicity testing. Pergamon Press. New York, NY. 261 p.

Maki, A. W., L. Geissel, and H. E. Johnson. 1975. Comparative toxicity of larval lampricide (TFM: 3-triflouromethyl-4-nitrophenol) to selected benthic macroinvertebrates. J. Fish. Res. Board Can. 32:1455-1459.

Niederlehner, B. R., J. R. Pratt, A. L. Buikema, Jr., and J. Cairns, Jr., 1986. Comparisons of estimates of hazard derived at three levels of complexity, p. 30-48. *In* J. Cairns, Jr. [ed.] Community toxicity testing, STP 920. American Society for Testing and Materials, Philadelphia, PA. 350 p.

Odum, E. P. 1984. The mesocosm. Bioscience 34:558-562.

Perez, K. T., and G. E. Morrison. 1985. Environmental assessments from simple test systems and a microcosm: Comparison of monetary costs, p. 89-95. *In* J. Cairns, Jr. [ed.] Multispecies toxicity testing. Pergamon Press. New York, NY. 261 p.

Perkins J. L. 1983. Bioassay evaluation of diversity and community comparison indexes. J. Water. Pollut. Control Fed. 55:522-530.

Pontasch, K. W., and J. Cairns, Jr. (in press). Establishing and maintaining laboratory-based microcosms of riffle insect communities: their potential for multispecies toxicity tests. Hydrobiologia.

Rosenberg, D. M., and V. H. Resh. 1982. The use of artificial substrates in the study of freshwater benthic macroinvertebrates, p. 175-235. *In* J. Cairns, Jr., [ed.] Artificial substrates. Ann Arbor Science Publishers, Ann Arbor, MI. 279 p.

Shaw, D. W., and G. W. Minshall. 1980. Colonization of an introduced substrate by stream macroinvertebrates. Oikos 34:259-271.

Stout, R. J., and W. E. Cooper. 1983. Effect of p-cresol on leaf decomposition and invertebrate colonization in experimental outdoor streams. Can. J. Fish. Aquat. Sci. 40:1647-1657.

Surber, E. W. and T. O. Thatcher. 1963. Laboratory studies of the effects of alkyl benzene sulfonate on aquatic invertebrates. Trans. Am. Fish. Soc. 92:152-160.

United States Environmental Protection Agency. 1985. Methods for estimating the acute toxicity of effluents to freshwater and marine organisms. Environmental Monitoring and Support Laboratory, Cincinnati, OH. EPA/600/4-85/013, 215 p.

Wall, T. M., and R. W. Hanmer. 1987. Biological testing to control toxic water pollutants. J. Water Pollut. Control Fed. 59:7-12.

Weber, C. 1973. Recent developments in the measurement of the response of plankton and periphyton to changes in their environment, p. 119-138. *In* G. Glass [ed.] Bioassay techniques and environmental chemistry. Ann Arbor Science Publishers, Ann Arbor, MI.

CHAPTER THREE

COMPARISONS OF SINGLE SPECIES, MICROCOSM, AND FIELD RESPONSES TO A COMPLEX EFFLUENT

Abstract

Short term, single species toxicity tests are used to identify, regulate, and monitor the discharge of problem effluents into receiving ecosystems. The ability of single species tests to protect entire ecosystems has been questioned, and microcosm tests have been proposed as better tools for reducing uncertainty in environmental risk assessment. We compared responses to a complex effluent from microcosms of indigenous macroinvertebrates and protozoans to responses observed in acute tests with *Daphnia magna*, *Ceriodaphnia dubia* and *Pimephales promelas* and chronic survival and reproductive tests with *C. dubia* The predictive utility of these various tests was then evaluated against observed effects in the receiving stream.

The acute LC_{50}s (% effluent) from the acute tests were 63.09 for *Pimephales promelas*, 18.8 to 31.3 for *Daphnia magna* and 54.7 for *Ceriodaphnia dubia*. Results from 7-day chronic tests indicated that *C. dubia* survival was significantly ($P \leq 0.05$) affected at 30% effluent and reproduction was affected at concentrations ≥3% effluent. In the protozoan microcosms, community composition was significantly ($P \leq 0.05$) changed at 1.0%; while protozoan species richness was significantly reduced at 3.0% effluent. In the macroinvertebrate microcosms, mayflies were the only macroinvertebrates adversely affected by the effluent. Four mayfly taxa had significant ($P \leq 0.05$) reductions in denstiy in the 1.0% treatment and four others were significantly reduced at 10.0%; while some chironomids had significant increases in density at 0.1% and above. Microcosm responses corresponded very well to observed effects in the field. The microcosms not only were the most sensitive indicators of effluent toxicity, they also correctly predicted which indigenous organisms would be lost and which would be stimulated at various ambient concentrations of the effluent.

Introduction

Although toxicity testing has been advocated as a tool for predicting environmental safety and harm since at least the 1940's (Hart *et al.* 1945), management strategies for many years ignored biological response as an assessment tool. In more recent years, toxicity test data finally has been incorporated into management decisions in discharge permitting, and this has resulted in an improved control of toxics.

Short-term, single species toxicity tests were initially chosen to identify, regulate, and monitor environmental problems, primarily because of familiarity with these tests and reports of their ease and low cost (USEPA 1985ab). However, in the years between the development of single species tests and their formal application to management problems, appreciation for the complexity of intact ecosystems has grown, and exclusive reliance on single species toxicity tests for predictions of environmental safety and harm has been criticized (e.g., NRC 1981; Cairns 1983; Odum 1984). Although single species tests can efficiently examine relative toxicity, they may not be the most accurate or efficient method for predicting responses in receiving ecosystems because single species tests: 1) do not take into account interactions among species; 2) utilize genetically homogeneous laboratory-stock test populations which may lack the adaptive capability of heterogenous natural populations; 3) utilize species of unknown relative sensitivity; 4) are often conducted under physical and chemical conditions which lack similarity to natural habitats; and 5) utilize species which are usually not indigenous to the receiving ecosystem, thus complicating field-validation of the test.

Studies of the effectiveness of short-term, single species biological testing of effluents under the EPA's Complex Effluent Toxicity Testing Program, as summarized by Wall and Hanmer (1987), have found correlations between response of single species in laboratory tests and certain indices of field response. However, single species tests are often overprotective when large application factors are employed, and accurate predictions of how much and what kind of damage to ecosystem structure and function to expect are not possible. Greater predictive accuracy is necessary in many management

applications such as premanufacture licensing where a confirmatory field survey is not possible and/or where costs of inaccurate predictions would be unacceptable.

Thus, whether current assessments of single species toxicity test constitute adequate validation of their use as predictors of ecosystem effects, as suggested by Wall and Hanmer (1987), depends on hotly disputed judgements of what is sufficient accuracy in predictions of environmental safety, and what ecosystem properties and behaviors are the essential objects of environmental protection. It is unlikely that a majority of ecologists would be satisfied that entire ecosystems were, in fact, being protected based on 7 day tests with three single species. An obvious way to reduce uncertainty in environmental risk assessment is by incorporating tests that directly examine those properties of ecosystems that are the object of environmental legislation (e.g., species richness). Such multispecies tests are being developed (see Giesy 1980; Cairns 1985, 1986; USEPA 1987), but have not been generally accepted in the regulatory framework.

Some multispecies tests utilize generic assemblages of organisms that may or may not occur together in nature or be indigenous to a given receiving system. These generic microcosms, in an effort to increase standardization and replicability, forfeit the environmental realism that makes multispecies tests so attractive. We believe, that the "communities" used in multispecies tests should either be derived from the proposed or actual receiving stream in the site-specific case or from regional reference ecosystems if standardized communities are needed for more general tests.

The objective of the research reported here was to evaluate the predictive utility of two types of multispecies tests and five types of conventional single species tests by comparing predictions from the various tests with a complex effluent to observed effects in the receiving system. Comparisons such as this are rarely made, but are essential for the development of a hierarchy of toxicity testing protocols.

Materials and Methods

Study area

Several receiving systems for possible use in this investigation were evaluated. Factors considered in evaluating receiving systems included: 1) presence of similar morphometric characteristics (e.g., substrate, flow, depth) at all (upstream and downstream) intensive study sites; 2) presence of a relatively unperturbed reference reach upstream from the point of impact; 3) presence of a consistent impact; and 4) cooperation from the discharging industry. Although several promising sites were identified, cooperation from only one discharging industry could be secured. That company cooperated with the understanding that they would not be identified.

The work described here was conducted concurrently in August, 1987 on a third order stream receiving a complex effluent and at the Ecosystem Simulation Laboratory at VPI&SU. Additional components of the original study plan including a quantitative benthic survey, *in situ* bioassays with macroinvertebrates and toxicity tests with protozoan microcosms and *Ceriodaphnia* using ambient water from each field site were interrupted when drought conditions forced the industry to stop discharging. Components of the effluent thought to be major contributors to its in-stream effects included chlorides, ammonia, phenols, and heavy metals including lead and copper. The surrounding watershed is largely forested and sparsely populated. There are no other point source discharges for >20 km upstream.

An upstream control site (UP) was located immediately above the effluent outfall. The first downstream site (DN1) was located ~5 m below the outfall in the effluent plume. The second downstream site (DN2) was located ~0.25 km below the discharge after thorough mixing. A third downstream site (DN3) was established ~1.4 km below the outfall. The fourth downstream site (DN4) was located ~6.4 km downstream, below a convergence with another third order stream that provides substantial dilution and a new source of colonists. All intensive study sites were located on riffle habitats and relatively unshaded by riparian vegetation. The substrate at all sites was primarily cobbles (6-25 cm) with a few boulders (>25 cm) ~25% embedded in pebbles (1-5 cm) and sand (0.5-3.0

mm). During the study period the average riffle width was ~7 m, except at site DN4 where the width was ~12 m. The average depth was 20 cm.

Effluent for laboratory tests was collected as a grab sample on 5, 8 and 15 August at a point immediately before discharge into the stream. Samples were pumped into 210-ℓ drums and transported to the laboratory for testing.

Single species tests

Acute (48 h) toxicity tests (USEPA 1985a) with *Daphnia magna* were conducted at the beginning and end of use of each effluent batch. These acute, static tests were conducted using both laboratory and stream water as the diluent for the final effluent batch to determine if the source of diluent had an effect on the toxicity of the effluent. Short-term chronic toxicity tests with *Ceriodaphnia dubia* were conducted using the test method described by the USEPA (1985b) except that a single effluent grab sample was used for all renewals. Effluent for this test was stored in a closed container at 4°C in the dark, and upstream water filtered through a 26 μ mesh net was used as diluent. Acute tests were also conducted with *Ceriodaphnia dubia* at the beginning and end of the chronic test. In addition we report the results of an acute test with fathead minnows (*Pimephales promelas*) that was conducted by the industry as part of the required biological testing under their NPDES permit.

Differences in *Ceriodaphnia dubia* and *Daphnia magna* survival were analyzed using Fisher's exact test. Differences in young produced by *C. dubia* were evaluated by a one-way ANOVA followed by Dunnet's procedure for the separation of means. The industry evaluated their test with *Pimephales promelas* using Probit analysis.

Macroinvertebrate microcosms

Macroinvertebrate communities were developed in rock-filled plastic containers (10.6 X 10.6 X 8.3 cm) with six circular holes (12 mm dia) in each side. The artificial substrates were secured on wooden frames which had previously been anchored to the stream bottom with iron rods and cement blocks. The substrates were colonized for 30 d in the riffle area at UP. On 2 August 1987, 16 colonized substrates were sampled to provide an estimate of initial macroinvertebrate species-abundances placed in the arti-

ficial streams. In addition, a 0.25 m^2 kick-sample was taken from the substrate at each of the field sites. On 3 August, additional substrates were removed from the stream by placing a dip net (mesh size ~ 350 μ) behind the container as it was transferred to one of 32 coolers (7 ℓ capacity), filled with river water, for transportation back to the laboratory.

Four colonized substrates were placed in each of 16 oval artificial streams (1.7 X 0.24 m channel) constructed of molded fiberglass. A standpipe in each stream maintained depth at 13 cm; a 500 μ mesh screen prevented the loss of macroinvertebrates down the standpipe. On August 5, the streams were dosed in quadruplicate at 0.0, 0.1, 1.0 and 10.0% effluent by pouring an appropriate amount of 100% effluent into each stream. Subsequent dosage (10 ml min^{-1}) was delivered by peristaltic pumps through pvc tubing. Dilution water (90 ml min^{-1} charcoal dechlorinated tap water) was supplied to each artificial stream from a headbox system. Current (25 cm $sec.^{-1}$) was provided by paddle wheels attached to a 1 cm iron rod which was turned at 30 rpm by a 0.25 hp electric motor. Two 120 cm Durotest Vita-Lites® over each stream provided daylight equivalent light. The photoperiod corresponded to sunrise and sunset on day 15 of the test. Each stream was covered by a 1.00 X 0.75 X 0.30 m emergence trap (mesh size ~1.0 mm).

Adult insects were collected from each artificial stream with an aspirator every 48-72 h throughout the course of the experiment. Due to the discontinuation of discharge, the test was terminated 10 days early. Therefore, after 20 days of exposure, the four macroinvertebrate artificial substrates were removed from the stream microcosms in a dip net and washed through a 500 μ sieve. Additional macroinvertebrates and fish were collected by draining the artificial streams through the dip net. The organisms and debris collected were preserved in labeled jars containing 70 percent ETOH. Sorting was done by hand, and insects, with the exception of some midges (Chironomidae:Diptera), were identified to the lowest possible taxonomic unit using appropriate references. A more detailed description of the techniques used in establishing and testing the macroinvertebrate microcosms has been reported elsewhere (Chapter 2).

Total insect density (adults and young combined) per taxon was determined for each stream microcosm. Macroinvertebrate taxa with mean densities ≥four per stream microcosm for any experimental group were considered a core taxon. The density of individuals in each core taxon compared over all experimental groups was used to determine treatment effects. The data were log-transformed and then analyzed by a one-way ANOVA in conjunction with Duncan's Multiple Range Test for the separation of means. Mean macroinvertebrate densities placed in each artificial stream (INIT) were determined by randomly combining the 16 substrates sampled on 2 August into four, four-substrate groups. These composited substrates were not included in the ANOVA, but the mean and standard error for each core taxon are reported to give an estimate of initial densities.

The quantitative macroinvertebrate data from *in situ* bioassays were lost when the industry discontinued discharge without notice. Therefore, field-validation of the laboratory microcosm results is based on results from the kick-samples taken at each field site on 2 August. Although quantitative comparisons between field and laboratory are therefore not possible, the kick-sample results do accurately reflect the gross differences in macroinvertebrate distributions among the sites during the study period.

Protozoan microcosms

Protozoan communities were developed on polyurethane foam artificial substrates (6 X 5 X 4 cm) anchored to cement blocks at site UP. The substrates were colonized for 13 days and then either transplanted to downstream field sites for *in situ* tests or returned to the laboratory in an insulated container partially filled with stream water. The ability of these communities to replicate themselves by colonizing barren substrates has been found to be a sensitive indicator of toxic stress (Cairns *et al.* 1985; Niederlehner *et al.* 1985; Pratt *et al.* 1987). Success of colonization of barren substrates (islands) from naturally derived communities (epicenters) was monitored.

At the laboratory, three substrates were examined immediately. The remaining substrates were randomly assigned to treatments; three replicate tanks at at each of five effluent concentrations (0.3, 1.0, 3.0, 10.0 and 30.0%) and a control. Charcoal

dechlorinated tap water was used as diluent. Test vessels were high-density polyethylene tanks (35 X 28 X 15 cm) containing 7.5 ℓ of solution. Light was provided by Durotest Vita-Lite® full-spectrum flourescent bulbs. Light intensity was 5000 lux and the photoperiod was 16 hour light/8 hour dark. Temperature was not controlled, and ranged from 19 to 22°C over the length of the test. Solutions were renewed using a grab sample of the effluent after 4 and 10 days.

A single colonized substrate (epicenter) was suspended from a hook in the bottom of the container at the center of each tank and surrounded by six initially barren (island) substrates. Island substrates from each tank were sampled after 7 and 21 days by squeezing the contents into a polystyrene sample cup. Two or three drops of the material that settled to the bottom of the cup were pipetted onto a microscope slide, covered with a No.1, 22 mm cover glass and examined for the relative abundance and kinds of protozoan species present. The entire cover glass was scanned at 200 and/or 450X total magnification. Protozoa were identified to the lowest practical taxon (usually genus or species) based on gross morphology and behavior using standard taxonomic keys. When taxonomic identification was uncertain, observations and drawings of characteristics were made to insure consistent identification over the course of the experiment. Subsamples were prepared and examined until an asymptotic number of species was reached. Generally, two subsamples were required.

The three colonized artificial substrates transplanted to each field site were collected after seven days at their new locations and placed in sterile plastic sampling bags. Stream water was added to cover the sponges. The bags were transported in a small insulated container to the laboratory for analysis. These substrates were analyzed in the same manner described above for substrates from the laboratory tests. In addition, these samples were examined for abundances of cladocerans and copepods by placing a 1 ml subsample in a Sedgewick-Rafter counting cell and scanning at 40X total total magnification.

Differences in taxonomic composition of protozoan communities based on species presence or absence were examined with Hendrickson's M statistic (Hendrickson 1978)

which compares the number of taxa in common for all pairs of substrates. Comparisons to controls were made stepwise by recalculating the statistic after sequential elimination of the highest treatment group until heterogeneity in the number of taxa in common was no longer significant. Crustacean responses were analyzed by a one-way ANOVA followed by Duncan's Multiple Range Test for the separation of means.

General

Temperature, dissolved oxygen, pH, hardness, conductivity, current velocity, and other appropriate parameters were monitored in the various test systems and the natural receiving system. Effluent concentrations in the microcosms and at the field sites were estimated by: 1) regressing conductivity values on known concentrations from serial dilutions of effluent for each effluent batch; and then 2) using the regression equation to predict concentration from conductivity measured in the laboratory microcosms and at the field sites. Both laboratory and field diluent water were used to determine separate regression equations for the laboratory and field estimates.

Results

General

Because of the large quantities needed and the 3.5 h travel time to the field site, dilution water for use in the laboratory microcosms could not be transported from the receiving stream, and effluent was usually collected at weekly intervals. Characteristics of the laboratory diluent, river water at UP and DN2 and the effluent are reported in Table 3-1. Average hardness of the laboratory diluent was twice that of the upstream river water. However, hardness of the laboratory diluent and the upstream river water was variable, overlapped in time and the differences apparently had no effect on organisms in the control microcosms.

The experiment was conducted during a period of extreme low flow in the receiving stream. Temperature at site UP on 1 August ranged from a low of 15°C at night to 25°C in the afternoon; while current velocities at the field sites on 5 August ranged from 11.4 to 19.6 cm sec^{-1}. Macroinvertebrate microcosm temperatures were set at 18°C and

Table 3-1. Water quality characteristics of river water (UP = upstream reference reach; DN2 = downstream site 2), laboratory diluent, and whole effluent. Means for the month of August, 1987. Data (except lab diluent) was provided by the discharging industry.

Characteristic	UP	DN2	Diluent	Effluent
pH (units)	7.09	7.11	8.00	6.80
Hardness (mg ℓ^{-1} $CaCO_3$)	37.70	44.3	80.00	860.00
Chlorides (mg ℓ^{-1})	3.43	76.2	8.67	3,131.70
NH_4—N (mg ℓ^{-1})	0.39	20.82	0.06	121.30
Total lead (μg ℓ^{-1})	3	3	<1	106

current velocity at 25 cm sec^{-1} to better approximate August temperature and flow conditions reported in previous years (unpublished industry data).

An strong correlation (r^2>0.99) existed between conductivity and effluent concentration for both types of diluent water. The estimated effluent concentrations in both types of laboratory microcosms were very close to the nominal concentration targeted for each microcosm, and were similar to concentrations in the receiving stream. However, industry analysis of water at DN2 found concentrations of some effluent components corresponding to 0.9 to 2.5% whole effluent rather than the 4.1% based on conductivity measures, suggesting that the degradation of some components may not be reflected in conductivity measures.

Single species tests

The LC_{50} from the industry tests with *Pimephales promelas* was 63.09±3.44% effluent. The LC_{50}s for *Daphnia magna*, as calculated by binomial method, were 31.3 for the effluent collected on 8 August and 18.8 for the effluent collected on 15 August (Table

3-2), indicating some between sample variability. However, the results from the tests with *D. magna* indicated no measureable degradation of the effluent's toxicity over the one week storage period, and only a slight difference between the two types of diluent (Table 3-2). Indigenous fishes were present at site DN1 which had an effluent concentration of 14.1 percent, but it is not known whether these indigenous fishes could have survived the higher concentrations suggested by the acute test with *Pimephales promelas*. The complete absence of crustaceans at DN1 (see Table 3-5), suggests that indigenous crustaceans were more intolerant of chronic exposure to the effluent than acute tests with *D. magna* would have predicted.

All tests with *Ceriodaphnia dubia* were conducted with the effluent collected on 8 August. In both acute toxicity tests with *Ceriodaphnia dubia*, one conducted at the beginning and one at the end of the chronic test, there were no survivors in 100% effluent and no mortalities in the next lower concentration (30%) (Table 3-3). The LC_{50}, as calculated by binomial method, was 54.7% effluent. Once again, this consistency over time suggests no change in the toxicity of the stored effluent over a seven day period. Results from the chronic test indicate that survival was affected at 30% effluent, and reproduction was affected at concentrations ≥3% effluent (Table 3-3). Slight stimulation was observed at lower effluent concentrations. Only the results from the reproduction test would have predicted the complete absence of crustaceans at DN1.

Macroinvertebrates

Other than insects, the only macroscopic organisms regularly observed in either field or laboratory samples were Decapoda and a darter (*Etheostoma* sp. :Percidae). These organisms colonized the macroinvertebrate substrates in numbers insufficient to allow a statistical analysis of treatment effects. However, these crayfish and darters were present in three of four samples from the high effluent microcosms and at all field sites, suggesting that they are relatively tolerant of the effluent.

Thirty-five insect taxa were collected from the stream microcosms at the end of the experiment. However, because of the lack of either larval (e.g., *Caenis* spp.) or adult (e.g., *Baetis* spp. and female Heptageniidae) keys to species, some taxa were analyzed

Table 3-2. Percent survival in acute tests with *Daphnia magna*. The code in parentheses refers to the August date when the effluent grab sample was taken, the second number is the age (days) of the effluent and the letter refers to source of the diluent used (S = upstream; L = laboratory)

Concentration (Percent)	Percent Survival					
	(8,1,L)	(8,6,L)	(15,1,L)	(15,1,S)	(15,7,L)	(15,7,S)
Control	100	100	100	100	100	100
6.25	100	100	100	100	100	100
12.50	100	100	100	100	100	100
25.00	100	100	5*	20*	0*	5*
30.0	NC	NC	0*	5*	0*	0*
37.50	NC	0*	0*	0*	0*	0*
50.00	0*	0*	0*	0*	0*	0*
100.00	0*	0*	NC	NC	NC	NC

NC = not conducted at that concentration
* = significantly different from the control (alpha = 0.05)

at the generic level. Fifteen of these taxa were considered to be core taxa (mean density≥4 per microcosm), and the following analysis is based on their response to the effluent.

Eight of the core taxa were mayflies (Ephemeroptera). The control streams maintained all mayfly taxa at or above initial levels during this 20-d experiment. The mayfly genera *Pseudocloeon* (Baetidae), *Caenis* (Caenidae), *Stenonema* and *Leucrocuta* (Heptageniidae) all had densities in the control streams that were considerably higher than initial densities, suggesting that recruitment from eggs (or nymphs <500μ) occurred during the course of the experiment. The mayflies were the only insect group that was adversely affected by the effluent in the stream microcosms (Fig. 3-1). Although the marked decrease in density in the 1.0 treatment is not significant for mayflies as a group, four of the mayfly core taxa (*Paraleptophlebia swannanoa*, *Pseudocloeon* spp., *Stenonema* spp. and *Leucrocuta* sp.) had significantly lower densities at 1.0% effluent concentration, while the other four core taxa (*Stenacron* sp., *Isonychia bicolor*, *Baetis* spp. and *Caenis* sp.) had significantly lower densities in the 10.0% concentration. Some of the more sensitive mayflies had nonsignificant decreases in density in the 0.1

Table 3-3. Percent survival and young per surviving female in acute and chronic tests with *Ceriodaphnia dubia*. Diluent water for all tests was upstream river water. The tests were run with effluent collected on 8 August, 1987.

Concentration (Percent)	Chronic survival	Young per surviving female[1]	Acute survival[2]
Control	100	18.9 (5.0)	100
0.1	100	20.4 (4.1)	100
0.3	100	21.8 (5.1)	100
1.0	100	20.0 (5.0)	100
3.0	100	12.9 (2.9)*	100
10.0	90	2.1 (3.2)*	100
30.0	60*	0*	100
100.0	0*	NO	0*

[1]treatment mean and standard deviation
[2]results were the same for acute tests conducted at the beginning and end of the chronic tests.
* = significantly different from the control (alpha = 0.05)
NO = not observed

microcosms. Nearly 80% of *Baetis* spp. emerged from the control streams during this 20 day experiment, and an even higher percentage emerged during a previous 30 day experiment (Pontasch and Cairns, in press). If adults had not been collected, the control stream densities of *Baetis* spp. would not have been significantly different from the 10.0 treatment, indicating that monitoring of adult emergence is necessary to accurately interpret the results of toxicity tests with aquatic insects. The response of mayflies in the stream microcosms corresponds very well both qualitatively and quantitatively to their response in the receiving stream (Table 3-4). Mayflies were either absent or had greatly reduced densities at sites DN1-DN3 where the effluent concentration (based on conductivity measures) ranged from 3.5 to 14.1%. At site DN4, where the effluent concentration was estimated to be 1.1%, mayflies were present, although the more sensitive species (e.g., *Stenonema* spp.) were apparently reduced in numbers. Two other hemimetabolous orders of insects, Plecoptera and Odonata, did not colonize the artificial substrates in densities great enough to allow statistical analysis of their response. However, several *Acroneuria* sp. (Perlidae:Plecoptera) were collected from the 10.0%

effluent streams at the end of the experiment, and one was found in the kick-sample from site DN1, so *Acroneuria* sp. may be relatively tolerant of the effluent. In addition, *Gomphus* sp. (Gomphidae:Odonata) was found at all field sites (Table 3-4), indicating that it also may be relatively tolerant.

The caddisflies (Trichoptera) were apparently quite tolerant of the effluent (Fig. 3-2). *Hydropsyche morosa* and *Cheumatopsyche* spp. (Hydropsychidae) were the only caddisflies considered to be core taxa, and neither taxon had a significant response in the stream microcosms. The stream microcosms were unable to reflect the higher densities of *H. morosa* and another species, *Hydropsyche sparna* at sites DN2—DN4 (Table 3-4), apparently because the length of the test did not allow recruitment of these relatively slower growing species. Similarly, the higher densities at site DN2 of *Brachycentrus* sp. (Brachycentridae), another collector-filterer, was not reflected by the stream microcosms. Apparently, these filter-feeders were either feeding directly on the organic matter present in the effluent, or on other taxa (e.g., see Chironomidae below) that were present in great numbers at these impacted sites. In a previous experiment the hydropsychids colonized the artificial substrates in very large numbers (INIT > 700), but only 30% of this density could be maintained over the 30-d experiment (Pontasch and Cairns, in press). In this experiment, even though initial densities were much lower, *H. morosa* still could not be maintained at initial densities. The cause of this problem is probably related to the quantity or quality of food available in the seston of the stream microcosms.

In addition to *Atherix lantha* (Athericidae) and *Hemerodromia* sp. (Empididae), the dipteran core taxa included the three main subfamilies of the Chironomidae; Chironominae, Orthocladiinae and Tanypodinae. Although all dipteran core taxa were present in relatively high numbers in the stream microcosms (Fig. 3-3), The Chironominae were particularly abundant on the artificial substrates, and were maintained at their initial densities in the control streams. Over 95% of the Chironominae were *Paratanytarsus* sp.; a parthenogenetic group with a life cycle of approximatey 13 days (USEPA, unpublished data). The significantly higher densities of *Paratanytarsus* sp.

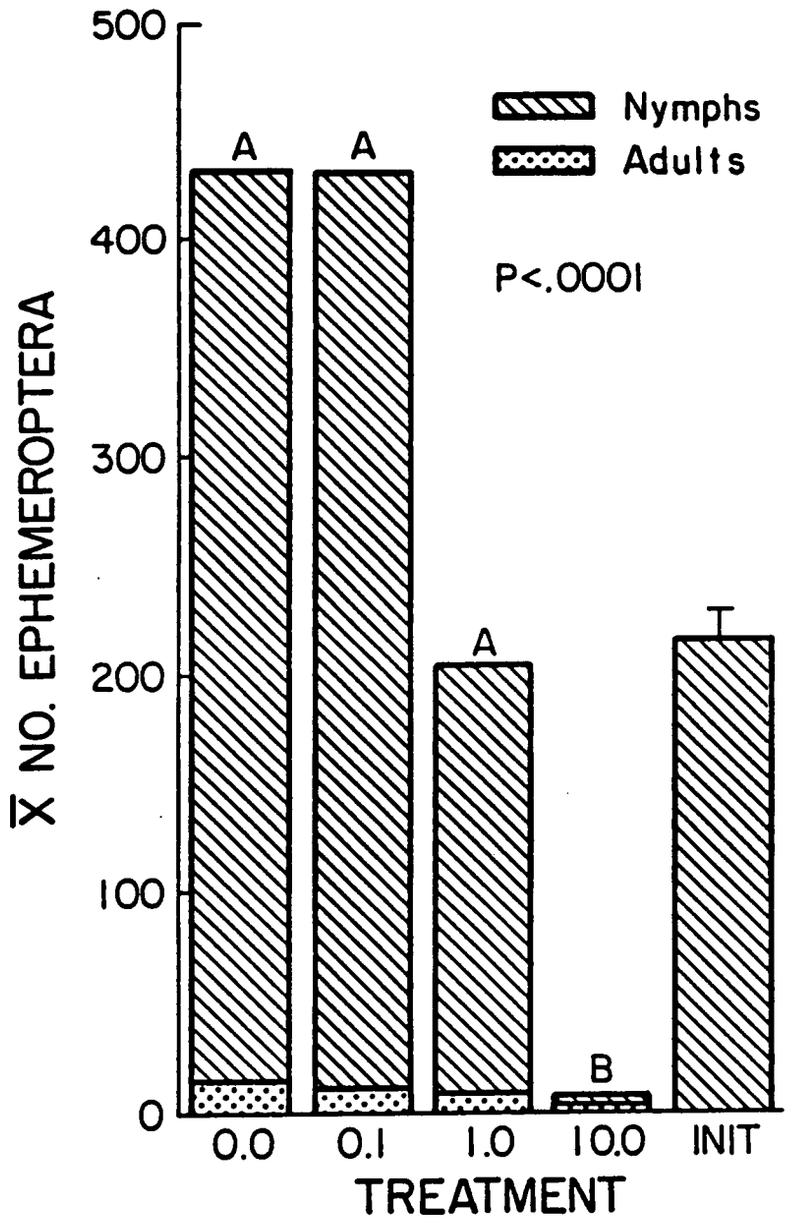

Figure 3-1. Mean number of mayflies (Ephemeroptera) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

Table 3-4. Results from the kick-samples taken at each field site: UP = upstream reference site; DN1-DN4 = sites downstream from the discharge. Only taxa with densities ≥5 from at least one site are included as individual taxa.

Taxon	UP	DN1	DN2	DN3	DN4
Effluent Concentration	0.0	14.1	4.1	3.5	1.1
Total Taxa	16	9	14	17	27
Taxa in Common with UP	16	5	7	11	14
Total Insects	115	280	2600	484	250
Percent Chironomids	33	98	95	85	30
Percent Mayflies	55	0	0	1	32
Isonychia	11	0	0	2	34
Stenonema	44	0	0	0	24
Stenacron	11	0	2	2	6
Leucrocuta	9	0	0	2	3
Baetis	1	0	0	0	12
Caenis	6	0	0	2	2
Optioservus	2	0	4	4	34
Hydropsyche	4	2	100	48	32
Cheumatopsyche	2	0	12	8	6
Brachycentrus	0	0	16	0	4
Chironominae	18	208	2304	288	42
Orthocladiinae	0	64	64	16	2
Tanypodinae	3	0	0	16	4
Atherix	6	1	76	96	22
Hemerodromia	0	0	12	2	0
Gomphus	1	1	6	2	4

in all stream microcosms receiving effluent corresponded to the much higher densities of this insect at sites downstream from the effluent outfall (Table 3-4). Effluent stimulation of their periphytic food source (Chapter 2), probably caused the increases in *Paratanytarus* sp. densities. Similarly, the higher numbers of Orthocladiinae present at sites downstream from the outfall, would have been predicted by the higher numbers of Orthocladiinae in the stream microcosms receiving effluent. Most, if not all, Orthocladiinae initially placed in the stream microcosms emerged during the course of the experiment. If emergence traps had not been in place, the trend towards higher numbers may have been obscurred due to adults from treated streams ovipositing in the control streams. The data for Tanypodinae was rather variable both in the stream

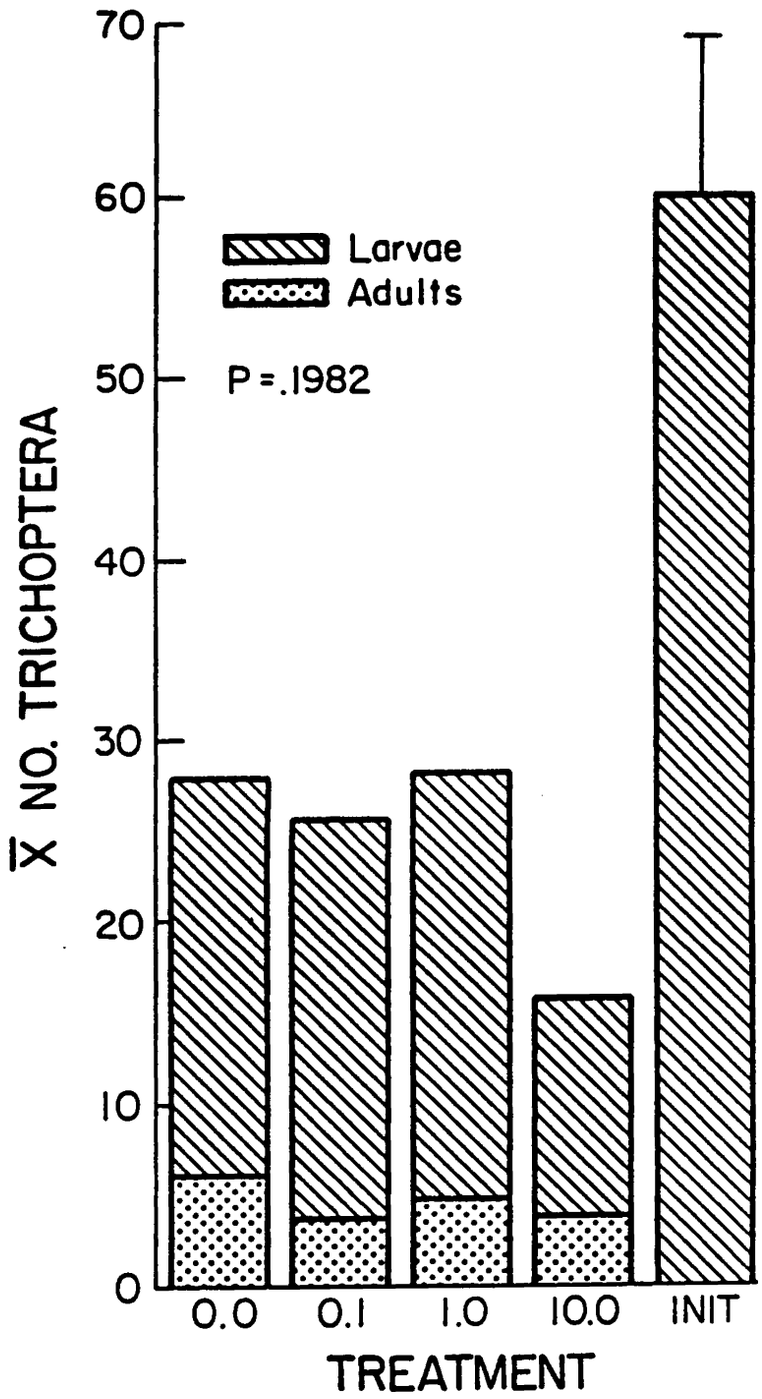

Figure 3-2. Mean number of caddisflies (Trichoptera) per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included).

microcosms and the field. However, the Tanypodinae also tended to increase in numbers in the treated microcosms.

Two other families of Diptera, Empididae and Athericidae, were represented respectively by *Hemerodromia* sp. and *Atherix lantha*. The larvae of both of these species are predacious, and their increase in numbers at DN2 (Table 3-4) is probably a response to the large number of chironomid prey available at that site. Both species were tolerant of the effluent concentrations used in the stream microcosms.

Protozoans

Results of the microcosm toxicity test are summarized in Table 3-5. Similar trends were observed in islands sampled after 21 days, but, because no effluent was available for renewals after 10 days, only the 7-day results are presented here. Species richness of protozoan communites was reduced at effluent concentrations ≥3%. The taxonomic composition of the protozoan communities was significantly different from controls at concentrations ≥1% effluent. No crustaceans were observed in concentrations ≥10% effluent.

The benthic microorganisms present downstream from the outfall were typical of a stream receiving large amounts of organic wastes. Thick mats of "sewage fungus" (i.e. *Sphaerotilus natans*) covered the substrate at site DN1. Within one week, these growths also covered the substrates transferred to DN1 from UP at the beginning of the test. This station also supported large numbers of a few heterotrophic protozoan species including *Paramecium caudatum, Epistylis* sp. and *Loxocephalus pagius*. Site DN2 also contained these protozoans and traces of "sewage fungus" were present, but this site was primarily characterized by thick mats of filamentous green algae (possibly *Cladophora* sp.). Periphyton samples taken from the natural substrate at UP and DN2 indicated a significant increase in chlorophyll *a* and total biomass (ash-free dry weight) at DN2 (Chapter 2). By 23 August, six days after the industry halted discharge, there was no visible evidence of the "sewage fungus" community present at DN1.

Results of analyses on artificial substrates transplanted from UP to the downstream sites for *in situ* tests are summarized in Table 3-5. Species richness of protozoan com-

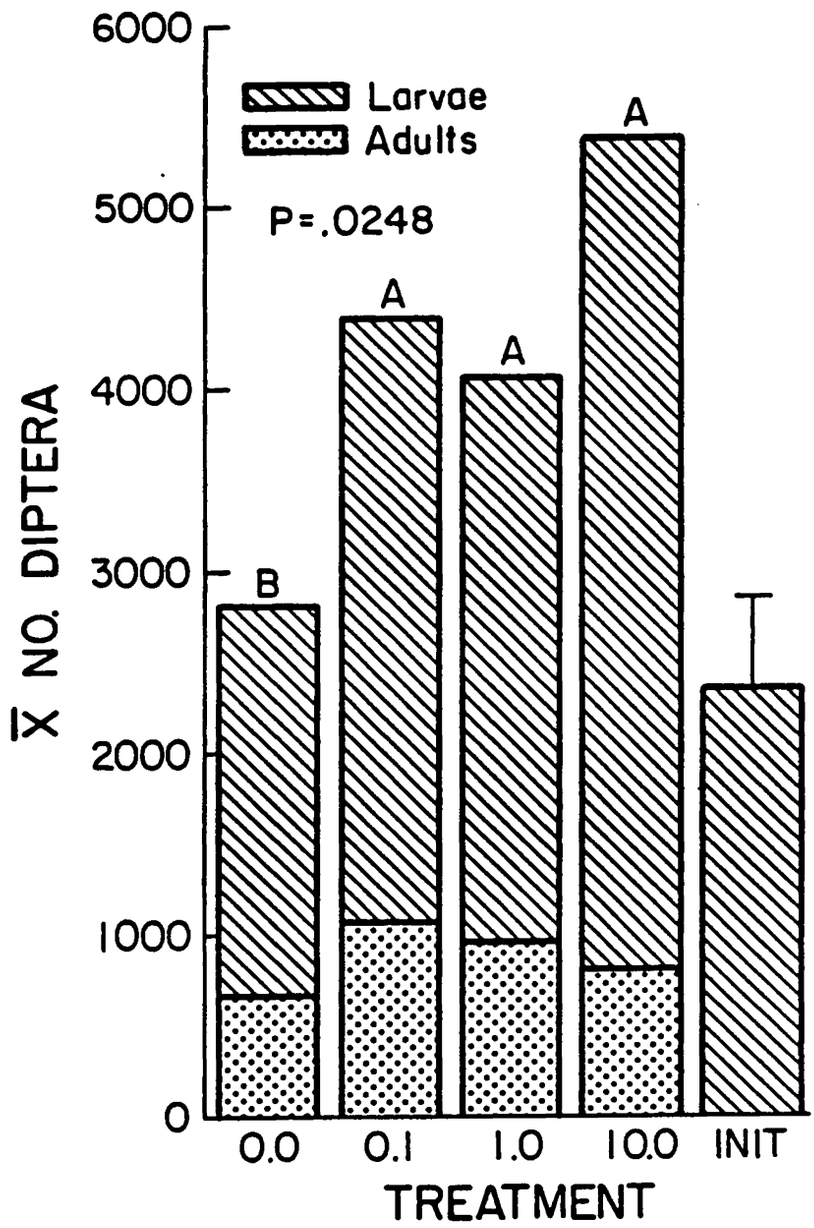

Figure 3-3. Mean number of Diptera per stream after 20 d in stream microcosms receiving four effluent levels. 0.0 = control; 0.1—10.0 = percent effluent concentration; INIT = is the estimated number of organisms placed in each stream at the beginning of the experiment (error bar is one S.E.). Statistical analysis is based on combined densities of adults and young. P value is from the overall ANOVA (INIT not included). Treatments with the same letter are not significantly different ($P \leq 0.05$)—analyzed by Duncan's multiple range test.

Table 3-5. Means and standard deviations of protozoan and crustacean responses in field and microcosm studies.

Microcosms			Field			
Treatment Conc. (Percent)	Protozoan species richness	Taxa in common w/control	Site-Conc. (Percent)	Protozoan species richness	Taxa in common w/UP	No. of Crustaceans per ml
Control	39.3 1.2	24.7 1.2	UP-0.0	39.0 3.5	24.0 2.6	5.33 0.58
0.3	37.3 3.8	20.0 3.2	DN4-1.1	34.7 2.3	19.2 2.8	10.75 3.59
1.0	34.7 4.6	18.8* 3.0	DN3-3.5	33.7 4.0	17.0 1.7	1.67 1.15
3.0	22.3* 2.9	12.6* 2.9	DN2-4.1	36.9 4.9	18.7 2.2	5.00 1.00
10.0	20.7* 4.0	10.0* 2.4	DN1-14.1	22.3* 5.1	7.1* 2.1	0.00* 0.00
30.0	17.7* 1.5	6.1* 1.2				

* = significantly different from the control (alpha=0.05)

munities appeared to be affected only at station DN1. There was substantial recovery in species numbers at DN2 and continuing downstream. Comparisons of protozoan community compostion indicate significant differences only at DN1. Crustacean abundance in substrate communities was severely affected at DN1 and recovered at DN2. Station DN4 showed greatly increased crustacean abundance.

The areas of agreement between response in the protozoan microcosms and the field include the appearance of heterotrophic species typical of organic enrichment, and elimination of protozoan and crustacean taxa at high concentrations of effluent. However, the significant reduction in protozoan species richness at 3.0% effluent in the laboratory microcosms was not observed at stations DN2 or DN3 which had effluent concentrations (measured by conductivity) greater than 3.0%. Similarly, the microcosms

predicted a significant difference in protozoan community composition at effluent concentrations ≥1.0%, but significant changes in protozoan community composition were only observed at DN1 where the effluent concentration was estimated to be 14.1%. However, the use of changes in protozoan community composition as a criterion, would have predicted the loss of some mayfly species that occurred at 1.0% effluent.

Discussion

The significant decrease in density of some mayflies, and the significant change in protozoan community composition at an effluent concentration of 1.0% were the most sensitive indicators of effluent toxicity. One mayfly, *Pseudocloeon* sp., was extirpated in the 1.0% treatment. Conversely, significant increases in dipteran density at 0.1% concentration, and chlorophyll *a* and periphytic biomass (ash-free dry weight) at 1.0% (Chapter 2), were observed in the macroinvertebrate microcosms. Without the use of an "application factor" the microcosm tests suggest that effluent concentrations of 0.1 to 0.3% should have no adverse effects on the indigenous macroinvertebrate and protozoan communities. Benthos samples taken in March, 1987, when higher flows reduced the effluent concentration in the plume at DN1 to 0.9%, did, in fact, contain the mayflies *Isonychia bicolor* and *Ephemera guttulata* (unpublished data).

The acute tests with *Daphnia magna*, *Ceriodaphnia dubia* and *Pimephales promelas*, if an application factor of 0.13 is used (USEPA 1985c), would suggest that an effluent concentration around 5% would produce no acute adverse effects on the indigenous biota. Therefore, these tests, in conjunction with an application factor, would have prevented the loss of crustaceans and some protozoans at DN1, but would not have prevented the elimination of most mayflies at DN1-DN3. The chronic survival of *Ceriodaphnia dubia* was similar to that in acute exposures. Conversely, tests of *C. dubia* reproduction successfully predicted elimination of crustaceans at DN1, and the elimination of mayflies at DN1-DN3. In another study, Carlson *et al.* (1986) found reproduction tests with *C. dubia* to be predictive of copper toxicity in a Connecticut river. The reproductive test with *C. dubia* is much more sensitive and biologically meaningful than most

conventional tests. However, the time and costs of conducting these reproductive tests with *C. dubia* are very similar to the protozoan microcosm tests (Niederlehner *et al.* 1986; Cairns and Pratt 1987) which were more sensitive and provided more information on how much and what kind of damage to expect at a given effluent concentration; for example, reductions in species richness or stimulation of periphyton.

In terms of time and costs, the macroinvertebrate microcosms suffer from the need to sort the organisms after sampling. However, these added costs may, in certain situations, be compensated by the wealth of information that can be gleaned from the test. For example, effluent-induced inhibition (*c.f.*, Clubb *et al.* 1975) or stimulation (*c.f.*, Maki *et al.* 1975) of adult insect emergence, and the periphyton-*Paratanytarsus* interaction observed in this study could not be detected by most conventional tests. In addition, The large number of indigenous taxa that are tested allows a more direct prediction of effects in the receiving stream. Although they are more applicable than tests with species that are not indigenous to the receiving stream, short-term, single species tests with indigenous insects would not be as sensitive as the protocol used in this investigation. Short-term bioassays with aquatic insects often indicate that they are more tolerant than fish or other invertebrates to environmental stressors such as heavy metals (Rehwoldt *et al.*, 1973; Clubb *et al.*, 1975; Sloof, 1983), ammonia (Williams *et al.*, 1986a), and various hydrocarbons (Sloof, 1983; Millemann *et al.*, 1984; Green *et al.*, 1986). However, aquatic insects, possibly because of lower resistance during molting, tend to exhibit increased sensitivity with longer exposures (Clubb *et al.*, 1975; Maki *et al.*, 1975; Spehar *et al.*, 1978). In addition, short-term bioassays often use late-instar larvae because they are more easily collected in the field and manipulated in the laboratory. Several studies have suggested that earlier instars are more sensitive than later ones to pesticides (Jensen & Gaufin, 1964; Sanders & Cope, 1968; Maki *et al.*, 1975;) and heavy metals (Clubb *et al.*, 1975; Gauss *et al.*, 1985; Williams *et al.* 1986b). Therefore, in order to conduct ecologically meaningful bioassays with aquatic insects, a relatively long exposure period is necessary to ensure that most species undergo one or more molts during testing and that some species have complete life cycles during the period of exposure. Although

they are slightly more time consuming and expensive, the biological significance of multispecies tests with indigeous insects and their applicability to field effects is much more obvious than that of single-species tests.

It seems apparent that the microcosm tests not only provided a more accurate prediction of the NOEC, but also indicated which indigenous organisms would be lost and which would be stimulated if the NOEC was exceeded to varying degrees. This type of information, which cannot be obtained from even the most sensitive single species tests or from tests with gnotobiotic microcosms, would be valuable in many management applications such as premanufacture licensing where a confirmatory field survey is not possible and/or where the costs of inaccurate predictions would be unacceptable.

The efficiency of single species tests in determining the relative toxicity of individual chemicals and complex mixtures is not in question. What is in doubt is the ability of these tests to accurately predict effects on receiving ecosystems. Our results indicate that the relatively pristine mountain stream investigated in this study would be better served by incorporating multispecies tests with indigenous organisms into the regulatory framework. Further research should be undertaken in an effort to develop standardized protocols for conducting multispecies toxicity tests with lotic communities. These protocols would provide environmentally realistic and scientifically justifiable alternatives to conventional tests, and would greatly reduce the amount of uncertainty in environmental hazard assessment.

References

Cairns, J. Jr. 1983. Are single species toxicity tests alone adequate for estimating environmental hazard? Hydrobiologia 100:47-57.

Cairns, J., Jr. ed. 1985. Multispecies toxicity testing. Pergamon Press, New York, 261 p.

Cairns, J., Jr. ed. 1986. Community toxicity testing. ASTM STP 920, American Society for Testing and Materials. Philadelphia. 350 p.

Cairns J., Jr., J.R. Pratt and B.R. Niederlehner 1985. A provisional multispecies toxicity test using indigenous organisms. J. Test. Eval. 13:316-319

Cairns, J., Jr. and J. R. Pratt. 1987. Ecotoxicologial effect indices: a rapidly evolving system. Wat. Sci. Tech. 19:1-12.

Carlson, A. R., H. Nelson, and D. Hammermeister. 1986. Development and validation of site-specific water quality criteria for copper. Environ. Toxic. and Chem. 5:997-1012.

Clubb, R. W., A. R. Gaufin & J. L. Lord, 1975. Acute cadmium toxicity tests upon nine species of aquatic insects. Envir. Res. 9:332-341.

Gauss, J. D., P. E. Woods, R. W. Winner & J. H. Skillings, 1985. Acute toxicity of copper to three life stages of *Chironomus tentans* as affected by water hardness-alkalinity. Envir. Pollut. 37:149-157.

Giesy, J.P. ed. 1980. Microcosms in ecological research. U.S. Department of Energy. 1110 p.

Giesy, J.P. and P.M. Allred. 1985. Replicability of aquatic multispecies test systems, p.187-247. In J. Cairns, Jr., ed. Multispecies toxicity testing. Pergamon Press. New York 261 p.

Green, D. W. J., K. A. Williams & D. Pascoe, 1986. Studies on the acute toxicity of pollutants to freshwater macroinvertebrates 4. Lindane (γ—Hexachlorocyclohexane). Arch. Hydrobiol. 106:263-273.

Hendrickson, J.A., Jr. 1978. Statistical analysis of the presence-absence component of species composition data. In K.L. Dickson, J Cairns, Jr. and R.J. Livingston, eds., Biological Data in Water Pollution Assessment: Quantitative and Statistical Analyses, STP 652. American Society for Testing and Materials, Philadelphia, PA, pp.113-124.

Jensen, L. D. & A. R. Gaufin, 1964. Effects of ten organic insecticides on two species of stonefly naiads. Trans. am. Fish. Soc. 93:27-34.

Maki, A. W., L. Geissel & H. E. Johnson, 1975. Comparative toxicity of the larval lampricide TFM (3 triflouromethyl-4-nitrophenol) to selected benthic macroinvertebrates. J. Fish. Res. Bd Can. 32:1455-1459.

Millemann, R. E., W. J. Birge, J. A. Black, R. M. Cushman, K. L. Daniels, P. J. Franco, J. M. Giddings, J. F. McCarthy & A. J. Stewart, 1984. Comparative acute toxicity to aquatic organisms of components of coal-derived synthetic fuels. Trans. am. Fish. Soc. 113:74-85.

National Research Council 1981. Testing for the effects of chemicals on ecosystems. National Academy Press, Washington, DC.

Niederlehner, B.R., J.R. Pratt, A.L. Buikema, Jr. and J. Cairns, Jr. 1985. Laboratory tests evaluating the effects of cadmium on freshwater protozoan communities. Environ. Toxicol. Chem. 4:155-165.

Niederlehner, B. R., J. R. Pratt, A. L. Buikema, and J. Cairns, Jr. 1986. Comparison of estimates of hazard derived at three levels of complexity. In Cairns, J., Jr. ed. 1986. Community toxicity testing. ASTM STP 920, American Society for Testing and Materials. Philadelphia. 350 p.

Odum, E.P. 1984. The mesocosm. Bioscience 34:558-562.

Pontasch, K.W. and J. Cairns, Jr. Establishing and maintaining laboratory-based microcosms of riffle insect communities: their potential for multispecies toxicity tests. Hydrobiologia (in press).

Pratt, J.R. B.R. Niederlehner and N.J. Bowers J Cairns, Jr. 1987. Prediction of permissible concentrations of copper from microcosm toxicity tests. Tox. Assess. 2:417-436.

Rehwoldt, R., L. Lasko, C. Shaw & E. Wirhowski, 1973. The acute toxicity of some heavy metal ions toward benthic organisms. Bull. Envir. Contam. Toxicol. 10:291-294.

Sanders, H. O. & O. B. Cope, 1968. The relative toxicities of several pesticides to naiads of three species of stonefly. Limnol. Oceanogr. 33:112-117.

Slooff, W., 1983. Benthic macroinvertebrates and water quality assessment: some toxicological considerations. Aquat. Toxicol. 4:73-82.

Spehar, R. L., R. L. Anderson & J. T. Fiandt, 1978. Toxicity and bioaccumulation of cadmium and lead in aquatic invertebrates. Envir. Pollut. 15:195-208.

U.S. Environmental Protection Agency. 1985a. Methods for estimating the acute toxicity of effluents to freshwater and marine organisms. Environmental Monitoring and

Support Laboratory, EPA/600/4-85/013 National Technical Information Service, Springfield, VA.

U.S. Environmental Protection Agency. 1985b. Methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. EPA/600/4-85/014 National Technical Information Service, Springfield, VA. Cincinnati, OH. EPA/600/4-85/014

U.S. Environmental Protection Agency. 1985c. Technical support document for water quality-based toxics control. EPA/440/4-85/032, National Technical Information Service, Springfield, VA. Environmental Monitoring and Support Laboratory, Cincinnati, OH. EPA/600/4-85/014

U.S. Environmental Protection Agency. 1987. Toxic substances control act test guidelines. Fed. Reg. 52:36334-36371

Wall, T.M. and R.W. Hanmer. 1987. Biological testing to control toxic water pollutants. J. Wat. Pollut. Cont. Fed. 59:7-12.

Williams, K. A., D. W. J. Green & D. Pascoe, 1986a. Studies on the acute toxicity of pollutants to freshwater macroinvertebrates, 3. Ammonia. Arch. Hydrobiol. 106:61-70.

Williams, K. A., D. W. J. Green, D. Pascoe & D. E. Gower, 1986b. The acute toxicity of cadmium to different larval stages of *Chironomus riparius* (Diptera: Chironomidae) and its ecological significance for pollution regulation. Oecologia 70:362-366.

CHAPTER FOUR

DIVERSITY INDICES, COMMUNITY COMPARISON INDICES AND CANONICAL DISCRIMINANT ANALYSIS: INTERPRETING THE RESULTS OF MULTISPECIES TOXICITY TESTS

Abstract

Canonical discriminant analysis, diversity indices, and community comparison indices were evaluated to determine their utility in quantifying macroinvertebrate response to a complex effluent in laboratory microcosms. The Shannon and Simpson diversity indices and the following community comparison indices were analyzed: 1) Renkonen's percentage similarity; 2) Bray-Curtis index; 3) average χ^2 index; 4) Stander's SIMI index; 5) Pinkham-Pearson index; and 6) Canberra metric index. Statistical inferences were made with the diversity measures using an ANOVA in conjunction with Duncan's test. A permutation or randomization procedure was used to test the hypothesis of no treatment effect based on the community comparison indices. The canonical discriminant analysis was primarily influenced by the drop in mayfly densities in the higher dose microcosms and indicated that the high-dose microcosms were different from the other microcosms. The Shannon and Simpson diversity measures indicated that both the medium and high dose microcosms were significantly different ($P \leq 0.05$) from the control microcosm, but this response was due solely to the increase in density of chironomids. When the diversity indices were calculated without chironomids, they failed to reflect adverse effects on mayflies in the medium dose microcosms. The statistical inferences for all community comparison indices were similar, and indicated that the high dose microcosms were significantly ($P \leq 0.06$) different from the control. However, the Bray-Curtis index provided the most meaningful condensation of the data.

Introduction

As appreciation for the complexity of intact ecosystems has grown, the exclusive reliance on single species toxicity tests for predictions of environmental safety and harm has been criticized (*e.g.*, NRC 1981; Cairns 1983; Odum 1984; Kimball and Levin 1985). Many researchers believe that uncertainty in environmental risk assessment could be substantially

reduced by incorporating tests that directly examine those properties of ecosystems that are the object of protective legislation (*e.g.*, species richness). Such multispecies toxicity tests are being developed (see Taub 1976; Giesy 1980; Cairns 1985, 1986). Early work suggests that they are not only more realistic models of exposure than single species toxicity tests, but they also may be no more expensive (Perez and Morrison 1985; Niederlehner *et al.* 1986; Cairns and Pratt 1987) and just as replicable (Giesy and Allred 1985) as conventional tests. Although multispecies tests have not been generally accepted, some tests have recently begun to be incorporated into the regulatory framework (USEPA 1987).

Multispecies tests have been criticized because they are not "interpretable" (Tebo 1985) or "decisive" (Mount 1985), and therefore, are not useful in a regulatory framework. These criticisms are due, in part, to the lack of sound statistical procedures for analyzing the community structure and composition data that emerges from these studies.

Given mortality as an endpoint, several approaches can be employed in analyzing data from multispecies tests: 1) conventional measures (*e.g.*, LC_{50}s) can be applied to individual taxa; 2) changes in density, biomass, etc. for individual taxa can be evaluated using univariate techniques such as ANOVA in conjunction with means separation procedures; 3) community level effects can be evaluated using multivariate techniques, such as canonical discriminant analysis, with species as variables; 4) changes in community structure can be measured using diversity indices; or 5) changes in community composition can be measured using community comparison indices. Decisions on which approach to use depends on what questions are being asked. If the interest is to determine responses of individual species, a univariate approach should be employed. Similarly, if the interest is evaluating community level response, either multivariate techniques should be applied directly to the data, or the data should be summarized using some summary measure of diversity or community composition. In the past the utility of these summary measures has usually been subjectively evaluated by the investigators (*e.g.*, Wilhm and Dorris 1968; Perkins 1983; Reinke 1986; Pontasch and Brusven 1988). Although this subjective approach may be appropriate for some situations, it is necessary to be able to make objective statistical inferences when "decisive" answers are required for regulatory purposes. Statistical inferences from diversity measures can be made

using a one-way ANOVA in conjunction with a means separation procedure. Inferential analysis of the community composition indices requires a less conventional approach (see Methods).

This study evaluates the utility of canonical discriminant analysis, the Shannon (Shannon 1948) and Simpson (Simpson 1949) diversity indices, and the percentage similarity (Renknonen 1938), cosine (Stander 1970), Bray-Curtis (Bray and Curtis 1957), Canberra metric (Lance and Williams 1967), average χ^2 (Parrish and Wagner 1983), and Pinkham-Pearson (Pinkham and Pearson 1976) community comparison indices in quantifying macroinvertebrate response to a complex effluent in laboratory stream microcosms. The diversity and community comparison indices were selected on the basis of common usage or reported utility. This selection process was the only major subjective step in this investigation. Although some are more promising than others, no single diversity or community comparison index has been proven effective in quantifying all of responses of aquatic communities to stressors. Therefore, it is necessary that investigators interpret their results in light of the limitations of each index used. For example, Johannsson and Minns (1987) suggested that the percentage similarity index, a popular measure which utilizes relative abundance, was the best index for "reflecting compositional differences" in phytoplankton communities. However, Pinkham and Pearson (1976) and Pontasch and Brusven (1988) point out that this index could result in a conclusion of complete similarity between communities that differed by orders of magnitude in absolute abundance. Such differences in absolute abundance often occur in comparisons between stressed and reference communities. It is clear that investigators using any index must use sound ecological judgement in their interpretation of the results.

Methods

The data used in the comparisons are from a study designed to determine if contaminant-induced changes in macroinvertebrate communities in laboratory stream microcosms could be used to predict macroinvertebrate responses in a natural stream receiving the same contaminant. Test organisms were transferred to the laboratory streams in rock-filled containers previously colonized for 30 days upstream of the outfall of a complex effluent. Four containers were placed in each artificial stream. The microcosms were covered by emergence traps and daylight equivalent lights provided a natural photoperiod. The 16 microcosms were continuously dosed in quadruplicate at four (0.0, 0.1, 1.0, and 10.0%) effluent concentrations. Adult insects were collected every 48-72 h from each microcosm. At the end of 20 days the remaining macroinvertebrates were censused. Macroinvertebrates with the exception of some midges (Diptera:Chironomidae) were identified to genus or species using appropriate keys. The data represent the combined densities of adult and young insects. In order to avoid interference from rare species, only those taxa that had a mean density ≥ 4 per microcosm were considered core taxa and included in the data set. Fifteen of the 35 taxa found in the microcosms at the end of the test were thus included, resulting in a 15 by 16 data matrix (Table 4-1). Each core taxon was subjected to a one-way ANOVA in conjunction with Duncan's multiple range procedure for the separation of means. The ability of the various indices tested to detect community-level responses was evaluated against population-level responses detected by the ANOVA.

The number of replicates was not sufficient to permit using all core taxa in the canonical discriminant analysis, so the three chironomid subfamilies were lumped as chironomids, the three genera of heptageniid mayflies (Ephemeroptera) as Heptageniidae, the two genera of baetid mayflies as Baetidae, and the two genera of hydropsychid caddisflies (Trichoptera) as Hydropsychidae resulting in a 9 by 16 data matrix. Due to the number of zeros, the data were normalized using $\log_{10}(X+1)$ transformation prior to canonical discriminant analysis. The analysis was run on the Statistical Analysis Sytem (SAS®) version 5.

The Shannon and Simpson diversity indices (Table 4-2) were calculated using the original 15 by 16 data matrix without a transformation. Pielou's evenness (Pielou 1966) was also

Table 4-1. Original data from a multispecies test with microcosms of indigenous stream insects.

	Dose (percent effluent)															
Taxon	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	1.0	1.0	1.0	1.0	10.0	10.0	10.0	10.0
Species richness	15	15	15	15	15	15	15	15	13	14	13	14	10	10	9	8
Ephemeroptera																
Isonychia bicolor	42	47	14	31	36	12	52	24	19	86	52	12	3	1	0	0
Caenis sp.	22	26	44	42	29	16	66	6	46	46	30	22	0	4	0	0
Pseudocloeon spp.	4	12	2	6	5	4	4	3	0	0	0	0	0	0	0	0
Baetis spp.	3	7	1	27	10	7	12	9	8	11	5	4	3	1	0	1
Paraleptophlebia sp.	20	6	14	6	4	4	20	6	0	6	2	4	0	0	0	0
Stenonema spp.	290	312	292	276	248	277	368	340	91	97	102	50	6	3	5	0
Stenacron sp.	9	6	32	14	2	20	30	5	8	13	0	16	0	0	0	0
Leucrocuta sp.	30	13	31	44	18	36	24	17	32	4	21	20	0	0	2	0
Trichoptera																
Hydropsyche morosa	6	9	9	9	8	2	10	15	5	4	8	11	7	0	6	9
Cheumatopsyche spp.	8	16	16	16	17	8	22	6	24	6	15	16	10	6	8	6
Chironomidae (Diptera)																
Paratanytarsus sp.	2037	2655	1886	2071	2739	2731	2976	3605	2658	3432	4046	2037	3655	3361	4759	4342
Orthocladiinae	257	304	188	349	271	501	726	523	425	314	664	284	372	526	453	557
Tanypodinae	32	33	10	24	7	56	33	49	37	62	121	34	91	137	84	47
Other Diptera																
Hemerodromia sp.	34	131	6	30	38	35	42	50	30	84	145	36	58	21	49	50
Atherix lantha	20	33	16	16	24	16	77	32	4	26	25	31	14	4	4	20

calculated to augment the interpretation of Shannon function values, because use of this index to assess community response to perturbation is based on the assumption that individuals in a polluted environment are less evenly distributed among the species present. Index values for each stream microcosm were considered "data points" and were analyzed by a one-way ANOVA and Duncan's multiple range procedure for the separation of means.

The mathematical complements of the Bray-Curtis, Canberra metric and average χ^2 indices were calculated, as indicated in Table 4-3, so that values from the various indices would be comparably scaled. The index values were computed for each possible pairwise combination of data vectors, resulting in a 16 by 16 matrix of similarity values. Some of these values are from comparisons within a given treatment (say the control), while others are from comparisons between treatments. Boyle *et al.* (1984) suggested testing for differences between two groups (*e.g.*, control and 1.0% effluent) by testing for differences between the within and between similarities using a two sample t-test (treating the within and the between similarities as two groups of observations). Similarly, Brock (1977) and Hruby (1987) suggested using an ANOVA to test for overall differences among within and between similarities followed by t-tests to detect group differences. The problem with these approaches is that the measures of similarity are not independent. In the two group case (with four replicates per group) 28 "data points" would be used to estimate the difference between the mean within (12) and mean between (16) similarities, while there are only 8 true replicates.

The procedure used in this study was based on a randomization or permutation analysis. The overall effect of dose was tested using a randomization of the data to see the affect of switching observations on the community measures (Mielke 1981; Zimmerman *et al.* 1985; Genter *et al.* 1987). For simplicity suppose there are two doses. Under the null hypothesis of no dose effect, the data collected from one dose should, within the limits of statistical error, be indistinguishable from the second dose. We can therefore compute the community similarity matrix using the actual data, switch some of the data among the doses, recompute the matrix and notice how the matrix changes. To evaluate change we can summarize the matrix by computing the mean similarity between replicates from different doses (say $\bar{B}_{data}$). We switch some observations, chosen in a random fashion, and recompute the mean similarity

Table 4-2. Diversity Indices

Shannon's (1948)	$H = -\sum_{k=1}^{S} p_k \log_2 p_k$
Simpson's (1949)	$1 - \lambda = 1 - \sum_{k=1}^{S} p_k^2$
Pielou's evenness (1966)	$J = \frac{H}{\log_2 S}$

Where:
S = number of species
p_k = proportion of individuals in species k

between replicates of different doses (call this $\bar{B}_{permute}$). The process is computed 1000 times. Then we have 1000 $\bar{B}_{permute}$ values and the actual value ($\bar{B}_{data}$). If their is no dose effect, $\bar{B}_{data}$ should be numerically similar to the $\bar{B}_{permute}$ values. If their is an effect, $\bar{B}_{data}$ should be smaller than most or all of the $\bar{B}_{permute}$ values. A test statistic would be the number of $\bar{B}_{permute}$ values smaller than $\bar{B}_{data}$. If the measure is a distance measure, the statistic will be the number greater than $\bar{B}_{data}$.

The second step in the analysis was to study differences between doses. This was done by comparing dose levels using the multiple comparison procedure of Foutz *et al.* (1985). With four treatments, there are six pairwise comparisons; for each of these a submatrix of between measures and a summary statistic $\bar{B}$. When the data are permuted, a new set of six summary statistics is computed. By appropriately ordering this set of six statistics for the 1000 permutations, experimentwise and comparisonwise error rates may be controlled and critical values for multiple comparisons determined (see Foutz *et al.* 1985 for details).

Table 4-3. Community comparison or similarity indices.

Percentage similarity (Renkonen 1938)	$PS_{ij} = \sum_{k=1}^{S} \min\ (p_{ik}, p_{jk})$
Cosine or SIMI index (Stander 1970)	$SIMI_{ij} = \frac{\sum n_{ik}\, n_{jk}}{\sqrt{\sum n_{ik}^2 \sum n_{jk}^2}}$ All the sums are from k = 1 to S
Bray and Curtis (1957)	$1 - BC_{ij} = 1 - \frac{\sum_{k=1}^{S} \lvert n_{ik} - n_{jk} \rvert}{\sum_{k=1}^{S} (n_{ik} + n_{jk})}$
Canberra metric (Lance and Williams 1967)	$1 - CM_{ij} = 1 - 1/S \sum_{k=1}^{S} \lvert n_{ik} - n_{jk} \rvert \,/\, (n_{ik} + n_{jk})$
Average χ^2 (Parrish and Wagner 1983)	$1 - \chi^2_{ij} = 2 \sum_{k=1}^{S} \frac{(O_k - E_k)^2 \,/\, E_k}{N}$
Pinkham and Pearson (1976)	$B_{ij} = 1/S \sum_{k=1}^{S} \frac{\min\ (n_{ik}, n_{jk})}{\max\ (n_{ik}, n_{jk})}$

Where:
S = number of core taxa present in microcosms i and j
p_{ik} = proportion of individuals in taxon k in microcosm i
p_{jk} = proportion of individuals in taxon k in microcosm j
n_{ik} = number of individuals in taxon k in microcosm i
n_{jk} = number of individuals in taxon k in microcosm j
N = total number of individuals in the core taxa sampled from microcosms i and j
O_k = number of individuals in taxon k present in one of the microcosms
E_k = sum of the individuals in taxon k from microcosms i and j divided by 2

The third step involved evaluation of the effect of individual species on the test statistic for overall differences. By removing a species and computing the new mean between similarity one can assess the influence a species has on the test. As a measure of influence one may use:

$$I_i = 100\,\frac{\bar{B}_{-i} - \bar{B}}{\bar{B}}$$

where

I_i measures the influence of species i

$\bar{B}_{-i}$ is the mean between similarity when species i is removed

$\bar{B}$ is the mean between similarity using all the species

If a species is adversely affected by the toxicant, I_i will generally be large and positive. That is, if the species was not sampled, the communities would appear to be closer together. A value near zero indicates that the species is redundant; its removal does not alter the information provided by the test. A large negative value of I_i suggests that species i is unaffected or positively affected by the toxicant. With measures based on abundance, negative values may also occur if the species is dominant and the toxicant does not affect it, or the effect is positive. Therefore, in this investigation, when the "influence" of *Paratanytarsus* (see below) was high relative to the remaining species, the data were $\log_{10}(X + 1)$ transformed and the entire analysis was rerun.

Results and Discussion

Other than insects, the only macroscopic organisms regularly observed in either field or laboratory samples were crayfish (Decapoda) and a darter (*Etheostoma* sp.:Percidae). These organisms colonized the substrates in insufficient numbers to allow an analysis of treatment effects. Eight of the 15 core taxa were mayflies (Ephemeroptera), and mayflies were the only group that was adversely affected by the effluent in the stream microcosms. Four of the mayfly core taxa *Paraleptophlebia swannanoa (Leptophlebiidae), Pseudocloeon* spp.

(Baetidae), *Stenonema* spp., and *Leucrocuta* sp. (Heptageniidae) had significantly ($P \leq 0.05$) lower mean densities at 1.0% effluent concentration, while the other four mayfly taxa *Stenacron* sp. (Heptageniidae), *Isonychia bicolor* (Oligoneuriidae), *Baetis* spp. (Baetidae), and *Caenis* sp. (Caenidae) had significantly ($P \leq 0.05$) lower mean densities in the 10.0% concentration. Some of the more sensitive mayflies had nonsignificant decreases in density in the 0.1 microcosms. The caddisflies (Trichoptera) *Cheumatopsyche* spp. and *Hydropsyche morosa* (Hydropsychidae) were apparently not adversely affected by the effluent; their mean densities were not significantly different in the various treatments. The dipteran *Paratanytarsus* sp. (Chironomidae) had very high densities in all stream microcosms (Table 4-1), and relative to the control microcosms, *Paratanytarsus* had significantly ($P \leq 0.05$) higher mean densities in all microcosms receiving effluent. The other chironomids present (Orthocladiinae and Tanypodinae) also had higher densities in the dosed microcosms; while *Atherix lantha* and *Hemerodromia* densities were variable (Table 4-1). Overall, the stream microcosms accurately predicted macroinvertebrate response in the field (Chapter 2). Although chironomids may be stimulated, the results suggest that an effluent concentration of 0.1% in the receiving stream should produce no adverse effects on the indigenous biota.

Canonical discriminant analysis

After lumping, the overall ANOVA's for the new groups were highly significant ($P \leq 0.0022$) for the five mayfly groups (*Caenis, Isonychia bicolor, Paraleptophlebia swannanoa*, Baetidae, and Heptageniidae), but none were significant for the four other groups (Hydropsychidae, Chironomidae, *Atherix lantha,* and *Hemerodromia* sp.). There were four treatments so there are (t-1) or three possible canonical variates. The first canonical variate (CAN1) explained over 85% of the variability, CAN2 12%, and CAN3 3%; only CAN1 was significant ($P = 0.0455$). The mayfly within group loadings all contributed to the separation along CAN1, but the heptageniid loading was dominant (Table 4-4). Since mayflies dominated CAN1 and were adversely affected by the effluent, the chironomid group loading for CAN1 was negative because chironomids had a positive response to the effluent. It is interesting to note that chironomids in general, and *Paratanytarsus* in particular, also had negative values when

"influence" on the various community comparison indices was calculated (see Table 4-6). These statistics are in some ways similar to each other. CAN1 separated the 10.0% effluent treatment from the other treatments (Fig. 4-1). This separation would be expected since there were very few mayflies found in the 10.0 microcosms. The need to "lump" the taxa (variables) in order to run this analysis may have reduced its sensitivity. This is a structural problem with using canonical discriminant analysis, not a result of poor experimental design. The use of the four artificial substrates in each stream as subsamples would have increased the number of "replicates" from four to 16, but the numerous adults and insects that colonized the sides of the streams could not be assigned to a particular artificial substrate.

Although canonical discriminant analysis may be useful in some multispecies testing applications, it is not recommended for use in analyzing data sets similar to that descibed here. The need to "lump" the taxa in order to run the analysis adds a degree of subjectivity and reduces the amount of community-level information being analyzed. In addition, if a dominant taxon exhibits a significant response to the stressor, that taxon will also dominate the canonical loadings at the expense of other less dominant but biologically significant taxa.

Diversity indices

The use of diversity indices to assess environmental perturbations has been in vogue for the past 20 years. Despite criticisms based on theoretical (Hurlbert 1971) and real world (Cook 1976; Godfrey 1978; Hughes 78; Murphy 1978; Perkins 1983; Pontasch and Brusven 1988) considerations, diversity indices, especially the Shannon function, have been and continue to be uncritically applied to nearly every investigation of benthic macroinvertebrate response to environmental stressors. Both Simpson's and Shannon's function are "heterogeneity" indices (*c.f.*, Peet 1974); both the number of species and the evenness with which individuals are distributed among the species contribute to the index values. Use of these indices in pollution assessment hinges upon the assumption that when aquatic communities are stressed, the number of taxa and the evenness with which individuals are distributed among the taxa are both reduced resulting in a lower index value. However, these indices do not consider the "kinds" of species present or their absolute abundance; a fact that can cause misleading re-

Table 4-4. Within canonical structure and standardized canonical coefficients for canonical variate one (CAN1).

Group	Within	Standardized
Isonychia	0.3173	0.1168
Caenis	0.4265	0.6847
Baetidae	0.2810	0.0273
Paraleptophlebia	0.3320	-0.2537
Heptageniidae	0.6948	3.5506
Hydropsychidae	0.1016	-0.0610
Chironomidae	-0.1254	-0.5201
Hemerodromia	0.0109	0.5608
Atherix	0.1382	0.6407

sults. In theory, the Brillouin index (Brillouin 1956) should be applied situations where the entire community can be censused; while the Shannon index should be used when diversity is to be estimated from samples from larger communities (Pielou 1975). However, these two indices usually respond similarly, so the more commonly used Shannon function was employed in this study.

Relative to the other methods used in this investigation, Simpson's and Shannon's diversity were the most sensitive in that they were significantly lower than the control at 1.0 percent effluent (Fig. 4-2). However, these results are misleading because they were due solely to the increase in chironomids in the treated streams. These indices were sensitive to the increase in chironomid densities because the higher densities resulted in a reduction in evenness. When the indices were calculated without chironomids, they were significantly higher in the 1.0 microcosms relative to the other treatments (Table 4-5) because the increase in evenness that resulted from the reduction in *Stenonema* densities (Table 4-1) was not offset. Therefore, in a microcosm test where enrichment did not occur, or where the test was too short to allow an enrichment response, these diversity indices are likely to be higher in stressed communities because stressor-induced reductions in species-abundances often result in a more even distribution of organisms (Pontasch and Brusven 1988). In other multi-

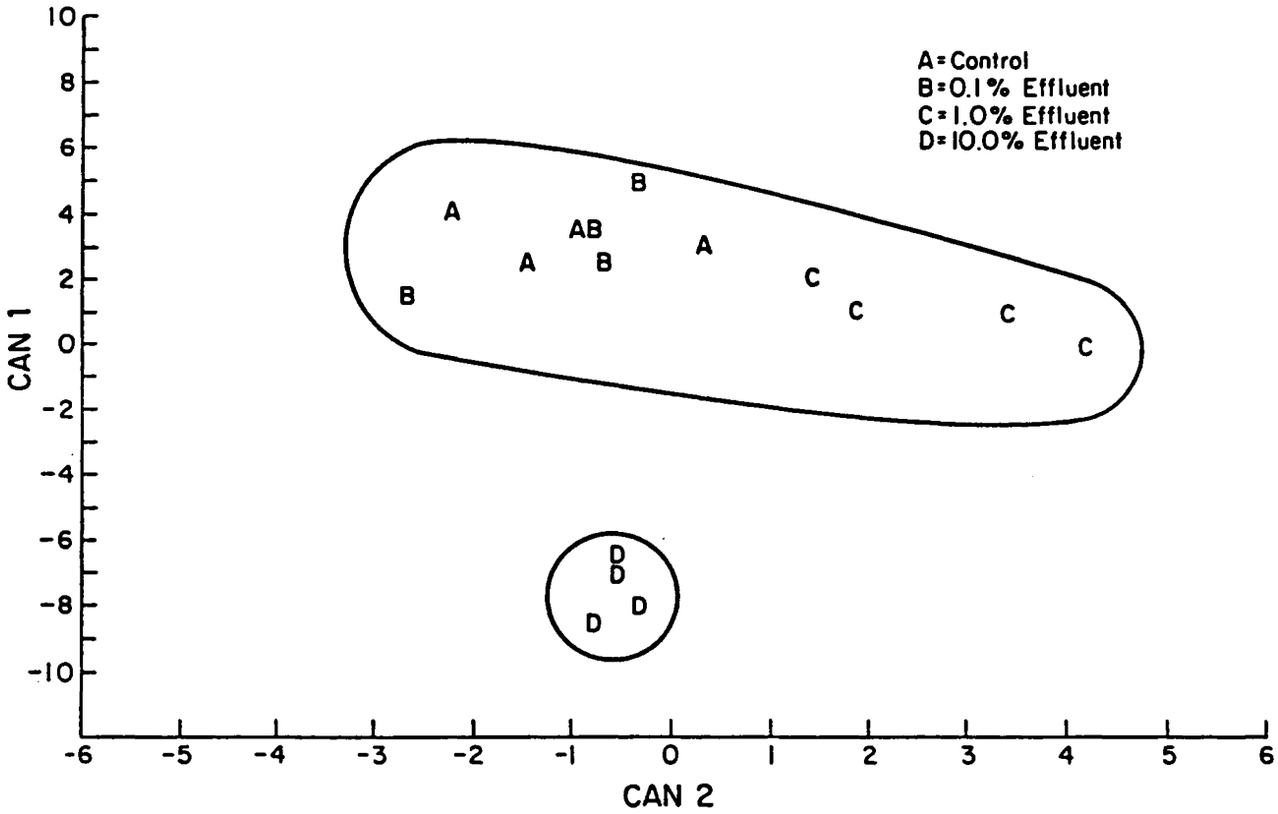

Figure 4-1. Results from a canonical discriminant analysis of multispecies data from a bioassay with indigenous insects. CAN1 = canonical variate 1; CAN2 = canonical variate 2. Only CAN1 was significant (P=0.0455). Mayflies dominated the CAN1 loadings.

species tests where diversity indices have been used to evaluate community repsonse to copper (Perkins 1983) and petroleum refinery wastewaters (Burks and Wilhm 1976) diversity either failed to decrease or increased because evenness increased in the stressed communities.

Community comparison indices

Community comparison (similarity) indices have been used to quantify benthic community response to environmental stressors for about the same period of time as diversity indices (*e.g.*, Cairns and Kaesler 1969), but only recently has their use begun to replace diversity measures. The similarity indices used in this study are all quantitative in that they utilize either relative or absolute abundance data. These indices quantify differences in the kinds of species present, and their relative or absolute densities. The method by which these differences are estimated and summed over the two communities being compared causes differences among the various index values. Some indices are biased for or against rare species, some are sensitive to zeros, and some are sensitive to dominant species. None of the indices has emerged as "the best", but some are considered better than others (Perkins 1983; Hruby 1987; Pontasch and Brusven 1988). In most water pollution studies, it appears that the index used is the one for which computer software is available. Although there are exceptions, most investigators make the assumption (never stated) that the index used is accurately quantifying actual differences between the two communities being compared. However, the various indices often respond differently to the same set of data (*e.g.*, Perkins 1983; Pontasch and Brusven 1988), so a great deal of caution should be employed in selection and interpretation. Clustering techniques are often employed to compare a series of communities, but in an impact study where deviation from a control or reference community is of interest, it is best to compare each impacted site or treated microcosm to the unperturbed community (Pontasch and Brusven 1988). In addition, statistical inference is necessary where subjective evaluations are not acceptable. All of these factors were taken into consideration during selection and evaluation of the community comparison indices used in this investigation.

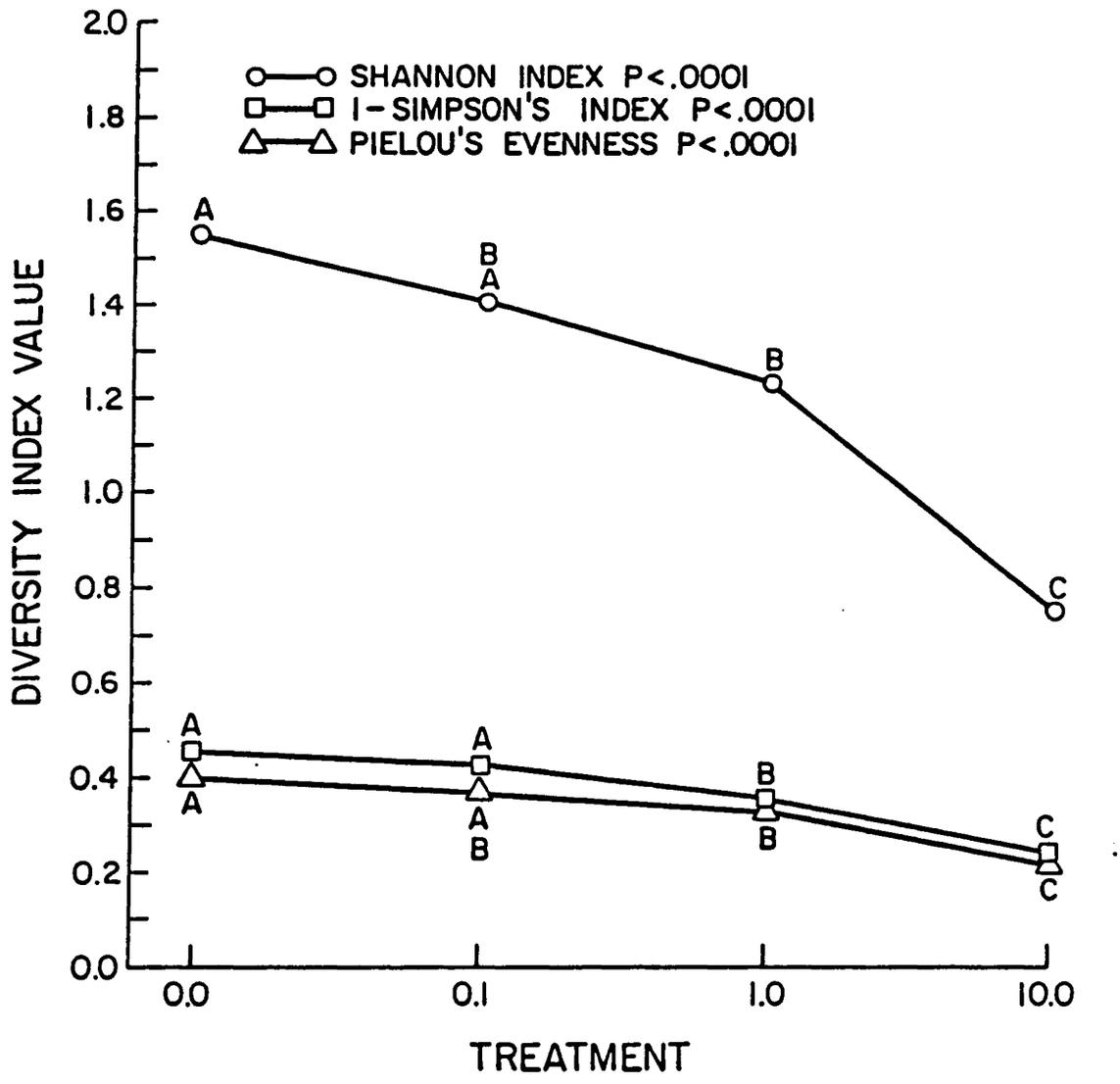

Figure 4-2. Results from diversity index analysis of multispecies data from a bioassay with indigenous insects. Treatments represent percent effluent concentrations in the microcosms. For a given index, points with the same letter are not significantly different ($P \leq 0.05$). Analysed by ANOVA/Duncans.

Table 4-5. Mean diversity index values calculated without chironomids. Means with the same letter are not significantly different ($P \leq 0.05$). P values are from the overall ANOVA.

Treatment (effluent conc.)	Simpson's	Shannon's	Evenness
0.0	0.651 b	2.295 b	0.640 b
0.1	0.621 b	2.212 b	0.617 b
1.0	0.819 a	2.805 a	0.827 a
10.0	0.608 b	1.850 c	0.702 b
P value	0.0006	0.0008	0.0011

Renkonen's (1938) percent similarity index is identical to the proportional similarity index of Whitaker (1975). This index measures differences in the proportional contribution each species makes to total numbers. The percent similarity index (run with original data) was strongly influenced by the large numbers of *Paratanytarsus* sp. present in the microcosms (Table 4-6). This influence resulted in relatively high mean similarities between the control and 0.1 (PS = .92), 1.0 (PS = .89), 10.0 (PS = .83) microcosms because the proportional contribution of *Paratanytarsus* was high in all treatments and dominated all comparisons. Index values from the run on the original data indicated that the 1.0 treatment was not significantly different ($P \leq 0.06$) from the 10.0 treatment, but the the 10.0 treatment was significantly different from the 0.0 and 0.1 treatments. After a $\log_{10}(X+1)$ transformation, the mean similarity between the 0.0 and 10.0 microcosms decreased from 0.83 to 0.61, suggesting that the reduction in the influence of *Paratanytarsus* allowed the reduction or extirpation of mayflies in the 10.0 microcosms to be reflected in the index value. However, the similarities between the control and 0.1 and 1.0 treatments remained high (Fig. 4-3) when in fact, there was a significant increase in *Paratanytarsus* densities in the 0.1 microcosms and a significant decrease in densities of four mayfly taxa in the 1.0 microcosms. In terms of the statistical inference, the percentage similarity index performed as well as most of the other indices tested. However, it does not quantify differences in absolute abundance between communities; a flaw that

cannot be overlooked when comparing impacted communities to a control. However, the percentage similarity index is usefull when absolute abundances from two different types of samples cannot be directly compared. For example, when comparing insect communities from artificial substrate samples to those sampled from a known surface area of the natural substrate, it is impossible to define a suitable correction factor for the differences in surface area and volume of substrate sampled. In this type of comparison the percentage similarity index would be the index of choice.

The Pinkham-Pearson (Pinkham and Pearson 1976) and Canberra metric (Lance and Williams 1967) indices were not unduly influenced by the large numbers of chironomids (Table 4-6). The reason for this is that all species are weighted equally by these two indices (see Table 4-3). If the object of a regulatory legislation is to treat each species as an entity to be protected, then, on the surface, these two indices would appear to be appropriate. However, if the protection of individual species is of concern, then a species-level analysis is more appropriate than any of the methods investigated in this paper. Neither of these indices had high within control similarities (Fig. 4-3). Inferences based on these indices indicated that the communities present in the 10.0 microcosms were weakly significantly different ($P \leq 0.06$) from controls. Neither index responded to the increases in *Paratanytarsus* densities in the treated microcosms. Wolda (1981) and Bloom (1981) demonstrated that the Canberra metric index has a nonlinear response to changes in species-abundances so use of this index in evaluating multispecies toxicity test results is questionable. The Pinkham-Pearson index was developed for pollution studies and is often used for impact assessment. These results support its use in most circumstances, but if enrichment is possible the Pinkham-Pearson index may not reflect the presence of dominant species.

The SIMI or cosine index generated values above 0.99 for all comparisons when the original data was used (Fig. 4-4). The influence of *Paratanytarsus* was high for this index also, but as was the case for most other community comparison indices, the SIMI index failed to reflect the significant increase in *Paratanytarsus* densities in the treated microcosms. Even after a $\log_{10}(X+1)$, transformation the similarity between the control and 0.1 and 1.0 microcosms was still very high (Fig. 4-5). The average χ^2 index and the Bray-Curtis indices

Table 4-6. Influence (see Methods) of each core taxon on similarity values. PS = percentage similarity; SIMI = cosine index; BC = Bray-Curtis index CM = Canberra metric index; χ^2 = average χ^2 index; PP = Pinkham-Pearson index. The O and L refer to original versus $Log_{10}(X + 1)$ data.

Taxon	PS-O	PS-L	SIMI-O	SIMI-L	BC-O	BC-L	χ^2-O	χ^2-L	CM-O	PP-O
Ephemeroptera										
Isonychia bicolor	0.523	1.434	0.007	0.832	0.292	1.369	0.242	1.044	8.183	10.294
Caenis sp.	0.545	1.590	0.007	0.948	0.233	1.395	0.206	1.263	8.054	9.950
Pseudocloeon spp.	0.091	1.589	0.000	0.451	0.048	1.254	0.046	1.226	9.276	8.307
Baetis spp.	0.150	0.413	0.001	0.237	0.074	0.677	0.056	0.302	7.122	9.180
Paraleptophlebia sp.	0.197	1.769	0.001	0.588	0.082	1.360	0.076	1.256	9.001	9.428
Stenonema spp.	4.938	1.395	0.396	1.361	2.009	1.536	1.791	0.815	8.822	11.027
Stenacron sp.	0.297	2.017	0.002	0.847	0.129	1.615	0.117	1.504	9.202	10.175
Leucrocuta sp.	0.476	1.712	0.005	0.881	0.167	1.395	0.155	1.256	8.066	9.812
Trichoptera										
Hydropsyche morosa	0.081	-0.317	0.000	0.007	0.013	-0.202	0.012	-0.201	4.311	6.043
Cheumatopsyche spp.	0.149	-1.167	0.001	-0.235	0.025	-0.696	0.011	-0.744	3.556	5.822
Chironomidae (Diptera)										
Paratanytarsus sp.	-26.200	-3.874	-14.27	-4.191	-17.978	-5.016	-14.570	-3.329	2.317	4.234
Orthocladiinae	0.0534	-2.784	0.147	-1.969	0.145	-3.272	-0.105	-2.325	2.856	4.993
Tanypodinae	0.053	-0.985	0.009	-0.236	0.344	-1.046	0.223	-1.187	5.177	7.698
Other Diptera										
Hemerodromia sp.	0.382	-1.490	0.012	-0.510	0.267	-1.216	0.165	-1.190	4.387	6.739
Atherix lantha	0.329	-1.151	0.004	-0.211	0.128	-0.439	0.090	-0.680	5.005	7.332

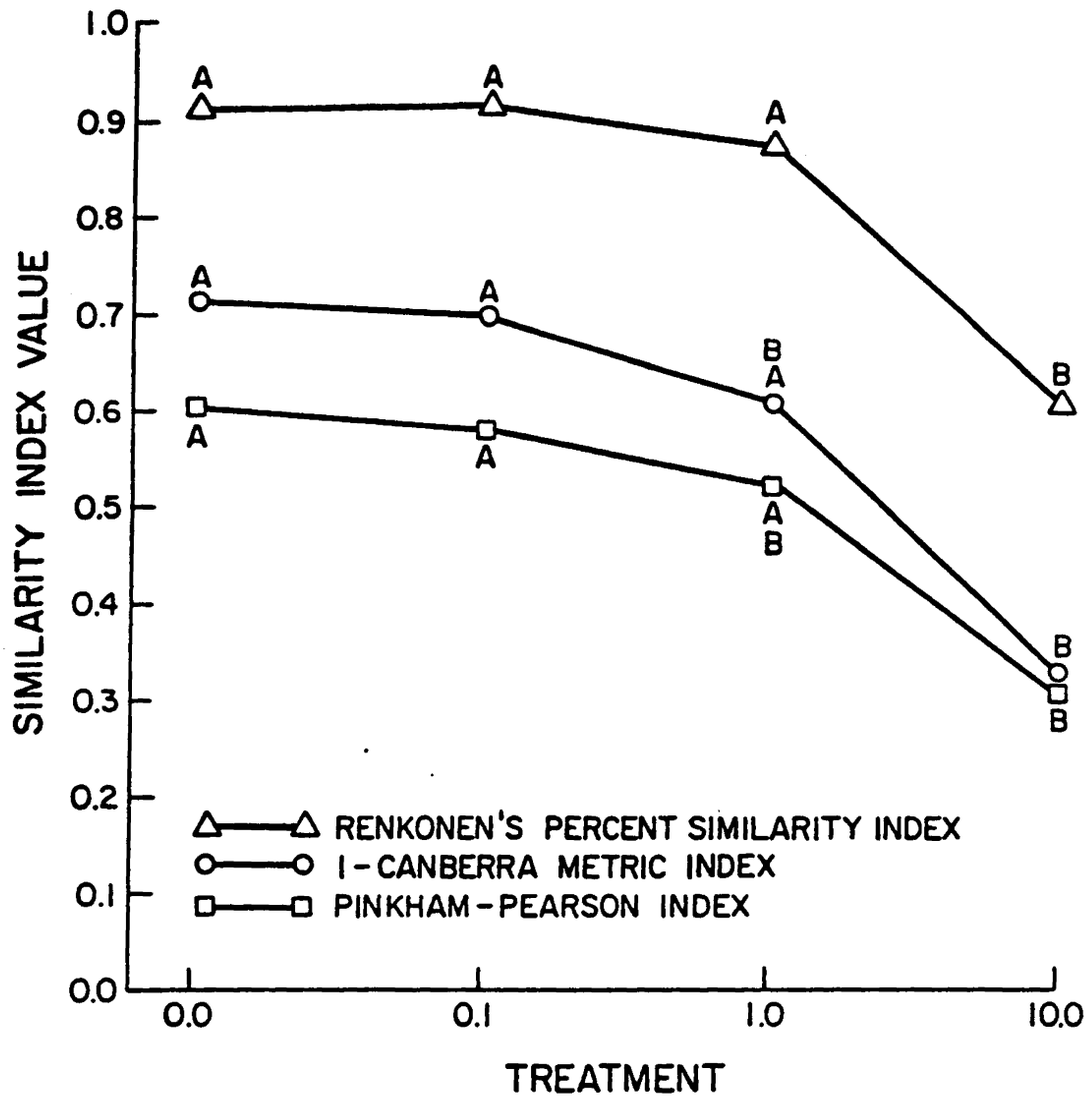

Figure 4-3. Results from community comparison index analysis of multispecies data from a bioassay with indigenous insects. Treatments represent percent effluent concentrations in the microcosms. Renkonen's percent similarity was calculated using $\log_{10}(X+1)$ transformed data. The Canberra metric and Pinkham-Pearson indices were calculated using the original data. For a given index, points with the same letter are not significantly different ($P \leq 0.06$). Analysed by a randomization/permutation procedure (see text).

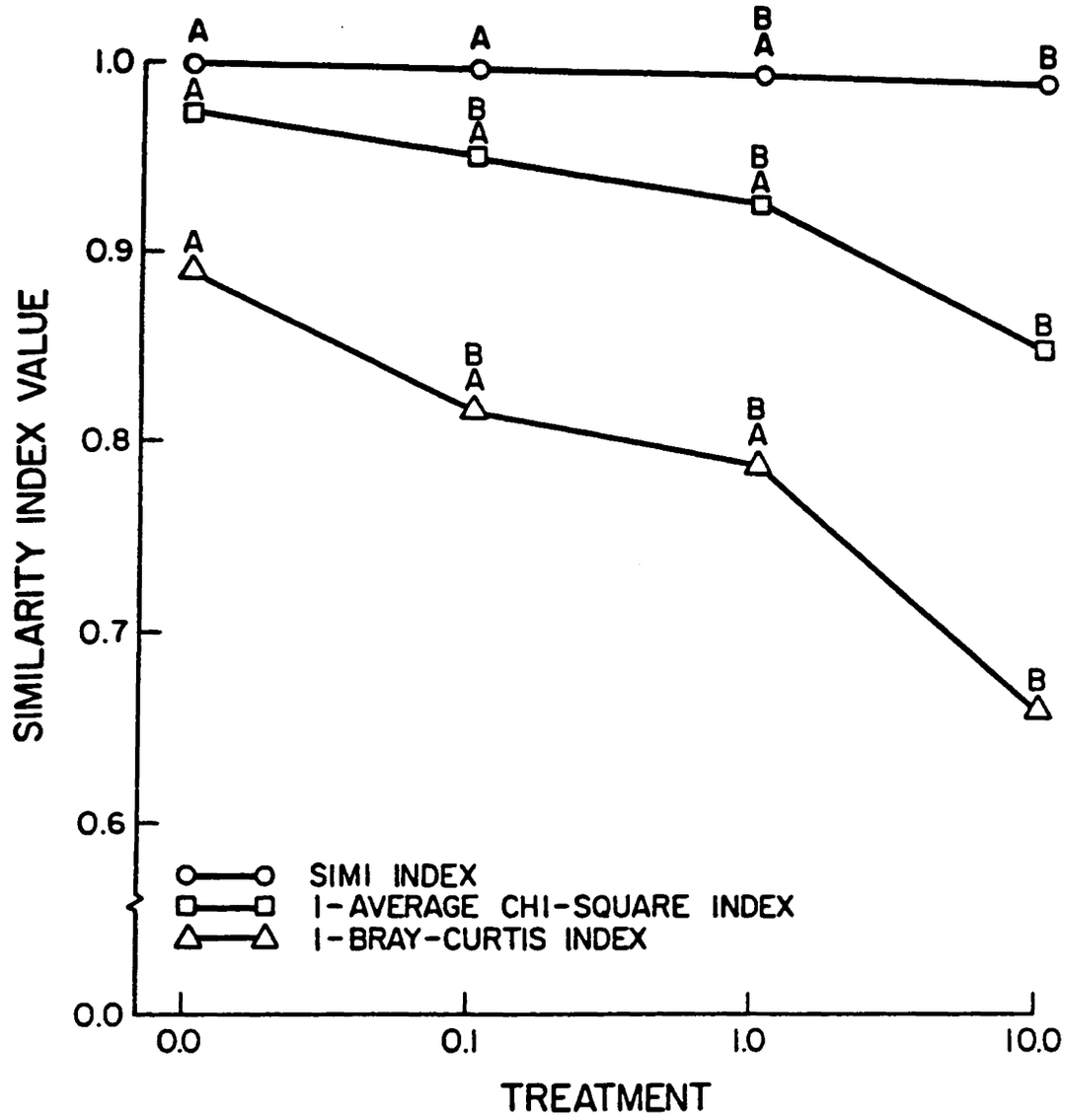

Figure 4-4. Results from community comparison index analysis of multispecies data from a bioassay with indigenous insects. Treatments represent percent effluent concentrations in the microcosms. All indices were calculated using the original data. For a given index, points with the same letter are not significantly different ($P \leq 0.06$). Analysed by a randomization/permutation procedure (see text).

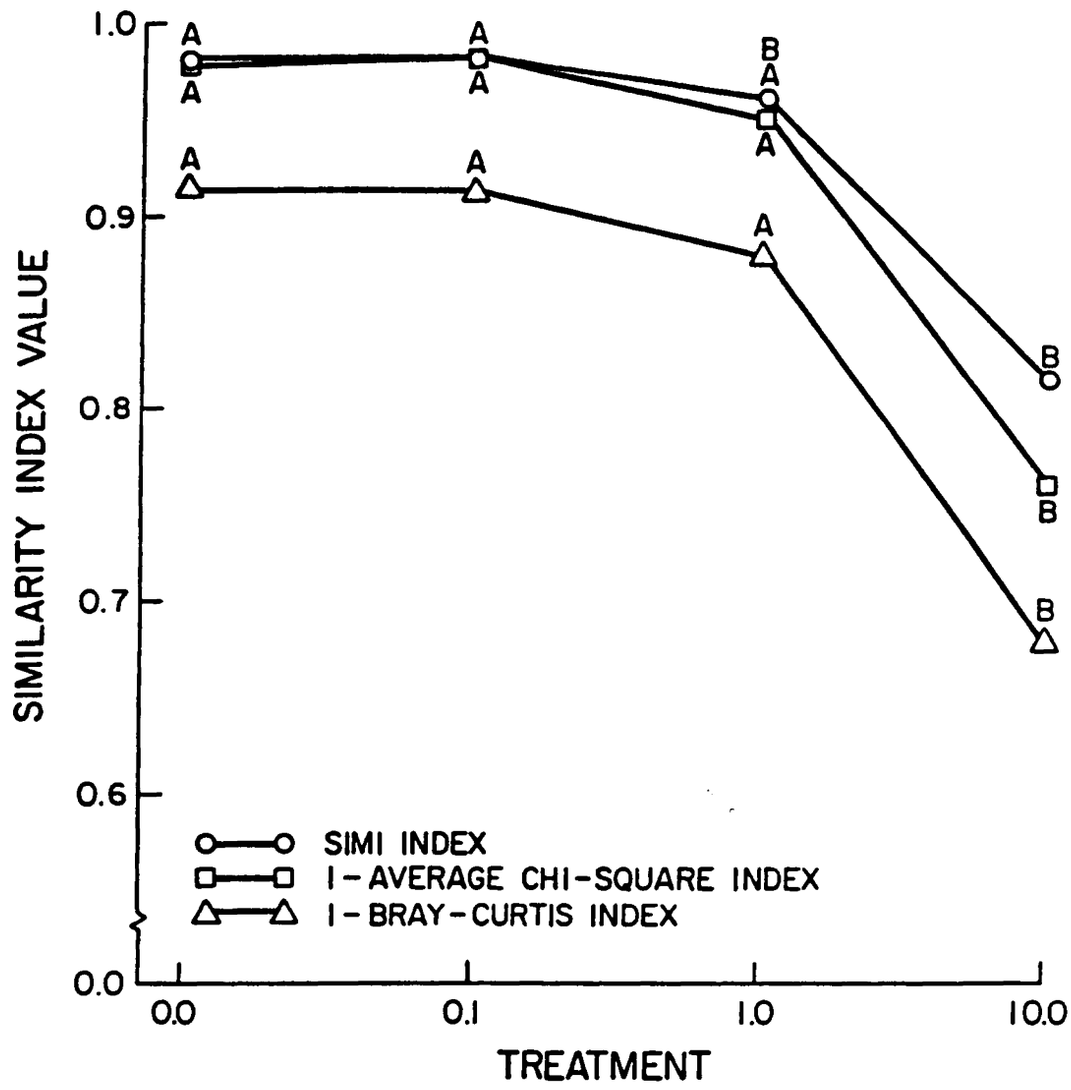

Figure 4-5. Results from community comparison index analysis of multispecies data from a bioassay with indigenous insects. Treatments represent percent effluent concentrations in the microcosms. All indices were calculated using $\log_{10}(X+1)$ transformed data. For a given index, points with the same letter are not significantly different ($P \leq 0.06$). Analysed by a randomization/permutation procedure (see text).

were also strongly influenced by *Paratanytarsus* (Table 4-6), but they did decrease in response to the significant increase in *Paratanytarsus* densities in the 0.1 microcosms (Fig. 4-4). Of all indices tested, the Bray-Curtis index had the largest drop in similarity between the control and 0.1 microcosms because it is sensitive to dominant taxa (Clifford and Stevenson 1975). However, the Bray-Curtis index also decreased at the 1.0 and 10.0 treatment levels, suggesting that its sensitivity to high numbers does not make it insensitive to less dominant species. After a $\log_{10}(X+1)$ transformation the Bray-Curtis and average χ^2 indices failed to decline at the 0.1 treatment level (Fig. 4-5).

Based on statistical inferences, none of the community comparison indices can be judged superior. Every index indicated that the communities exposed to the 10.0% effluent treatment were significantly different from the control communities. Thus, in terms of statistical inference, none of the indices detected the significant shifts in densities of some taxa that occurred at the 0.1 and 1.0 treatment levels. It appears that the community-level data was too variable within treatments to allow for significant differences at lower effluent concentrations. However, the response of the Bray-Curtis index to the untransformed data most closely reflected the results from the ANOVA's at the population level. In a number of other investigations, the Bray-Curtis index performed better than the other indices tested (Field *et al.* 1982; Perkins 1983; Hruby 1987; Pontasch and Brusven 1988). The Bray-Curtis index, therefore, often can provide a meaningful condensation of the data, but evaluations of community-level responses should always include considerations of population-level responses. None of the measures used in this investigation will ever replace sound ecological judgement in the interpretation of multispecies test results.

References

Bloom S. A. (1981) Similarity indices in community studies: potential pitfalls. *Mar. Ecol. Prog. Ser.* **5**, 125-138.

Boyle T. P., Sebaugh J. and Robinson-Wilson E. (1984) A hierarchical approach to the measurement of changes in community structure induced by environmental stress. *J. of Testing and Evaluation* **12**, 241-245.

Bray J.R. and Curtis J.T. (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27**, 325-349.

Brock D. A. (1977) Comparison of community similarity indexes. *J. Wat. Pollut. Control Fed.* **49**, 2488-2494.

Brillouin L. (1956) *Science and Information Theory.* Academic Press, New York.

Burks S. L. and Wilhm J. L. (1977) Bioassays with a natural assemblage of benthic macroinvertebrates. In *Aquatic Toxicology and Hazard Evaluation* (Edited by Mayer F.L. and Hamelink J. L.) pp. 127-136, ASTM STP 634. American Society for Testing and Materials.

Cairns J. Jr. (1983) Are single species toxicity tests alone adequate for estimating environmental hazard? *Hydrobiologia* **100**, 47-57.

Cairns J. Jr. [Editor] (1985) *Multispecies toxicity testing.* Pergamon Press, New York.

Cairns J. Jr. [Editor] (1986) *Community toxicity testing* ASTM STP 920. American Society for Testing and Materials.

Cairns J. Jr. and Kaesler R. L. (1969) Cluster analysis of Potomac river survey stations based on protozoan presence-absence data. *Hydrobiologia* **34**, 414-432.

Cairns J. Jr. and Pratt J. R. (1987) Ecotoxicological effect indices: a rapidly evolving system. *Water. Sci. Tech.* **19**, 1-12.

Clifford H. T. and Stephenson W. (1975) *An introduction to numerical classification.* Academic Press, New York.

Cook S. E. (1976) Quest for an index of community structure sensitive to water pollution. *Envir. Pollut.* **11**, 269-288.

Field J. G., Clark K. R. and Warwick R. M. (1982) A practical strategy for analysing multispecies distribution patterns. *Mar. Ecol. Prog. Ser.* **8**, 37-52.

Foutz R. V., Jensen D. R. and Anderson G. W. (1985) Multiple comparisons in the randomization analysis of designed experiments with growth curve responses. *Biometrics* **41**, 29-37.

Genter R. B., Cherry D. S., Smith E. P. and Cairns J. Jr. (1987) Algal-periphyton population and community changes from zinc stress in stream mesocosms. *Hydrobiologia* **153**, 261-175.

Giesy J. P. [Editor] (1980) *Microcosms in ecological research* U.S. Department of Energy, Washington, DC.

Giesy J. P. and Allred P. M. (1985) Replicability of aquatic multispecies test systems. In *Multispecies Toxicity Testing.* (Edited by Cairns J. Jr.) pp. 187-247, Pergamon Press. New York.

Godfrey P. J. (1978) Diversity as a measure of benthic macroinvertebrate community response to water pollution. *Hydrobiologia* **57**, 111-122.

Hruby T. (1987) Using similarity measures in benthic impact assessments. *Environmental Monitoring and Assessment* **8**, 163-180.

Hurlbert S. H. (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology* **52**, 577-586.

Johannsson O. E. and Minns C. K. (1987) Examination of association indices and formulation of a composite seasonal dissimilarity index. *Hydrobiologia* **150**, 109-121.

Kimball K. D. and Levin S. A. (1985) Limitations of laboratory bioassays: the need for ecosystem level testing. *Bioscience* **35**, 165-171.

Lance G.N. and Williams W.T. (1967) Mixed data classificatory programs.1. Agglomerative systems. *Aust. Comput. J.* **1**, 82-85.

Mielke P. W., Berry K. J., Brockwell P. J. and Williams J. S. (1981) A class of nonparametric procedures based on multiresponse permutation procedures. *Biometrika* **68**, 720-724.

Mount D. I. (1985) Scientific problems in using multispecies toxicity tests for regulatory purposes. In *Multispecies Toxicity Testing.* (Edited by Cairns J. Jr.) pp. 13-18, Pergamon Press. New York.

Murphy P.M. (1978) The temporal variability in biotic indices. *Envir. Pollut.* **17**, 227-236.

National Research Council (1981) *Testing for the Effects of Chemicals on Ecosystems.* National Academy Press, Washington, DC.

Niederlehner B. R., Pratt J. R., Buikema A. L. Jr. and Cairns J. Jr. (1986) Comparisons of estimates of hazard derived at three levels of complexity. In *Community Toxicity Testing.* (Edited by Cairns J. Jr.) pp. 30-48, ASTM STP 920. American Society for Testing and Materials.

Odum E. P. (1984) The mesocosm. *Bioscience* **34**, 558-562.

Parrish F. K. and Wagner J.A. (1983) An index of Community structure sensitive to water pollution. *J. Freshwat. Ecol.* **2**, 103-107.

Peet R. K. (1974) The measurement of species diversity. *Ann. Rev. Ecol. Syst.* **5**, 285-307.

Perez K. T. and Morrison G. E. (1985) Environmental assessments from simple test systems and a microcosm: Comparison of monetary costs, In *Multispecies Toxicity Testing.* (Edited by Cairns J. Jr.) pp. 89-95, Pergamon Press. New York.

Perkins J. L. (1983) Bioassay evaluation of diversity and community comparison indexes. *J. Wat. Pollut. Control Fed.* **55**, 522-530.

Pielou E. C. (1966) Species-diversity and pattern-diversity in the study of ecological succession. *J. Theor. Biol.* **10**, 370-383.

Pielou E. C. (1975) *Ecological Diversity*. John Wiley and Sons, New York.

Pinkham C. F. A. and Pearson J. G. (1976) Applications of a new coefficient of similarity to pollution surveys. *J. Wat. Pollut. Control Fed.* **48**, 717-723.

Pontasch K. W. and Brusven M. A. (1988) Diversity and community comparison indices: assessing macroinvertebrate recovery following a gasoline spill. *Water Research.*

Reinke D. C. (1986) Evaluation of the use of community similarity techniques as applied to phytoplankton communities. In *Community Toxicity Testing.* (Edited by Cairns J. Jr.) pp. 6-17, ASTM STP 920. American Society for Testing and Materials.

Renkonen O. (1938) Statistisch-okologische Untersuchungen uber die terrestische Kaferwelt der finnischen Bruchmoore. *An. Zool. Soc. Zool-Bot. Fenn. Vanamo* **6**, 1-231.

Shannon C. E. (1948) A mathematical theory of comunication. *Bell System Tech. J.* **27**, 379-423;623-656.

Simpson E. H. (1949) Measurement of diversity. *Nature* **163**, 688.

Stander J. M. (1970) *Diversity and similarity of benthic fauna off the coast of Oregon.* M.S. Thesis, Oregon State Univ., Corvallis, Oregon.

Taub F. B. (1976) Demonstration of pollution effects in aquatic microcosms. *International Journal of Environmental Studies* **10**, 23-34.

Tebo L. B. Jr. (1985) Technical considerations related to the regulatory use of multispecies toxicity tests. In *Multispecies Toxicity Testing.* (Edited by Cairns J. Jr.) pp. 19-26, Pergamon Press. New York.

U.S. Environmental Protection Agency (1987) Toxic substances control act test guidelines. *Fed. Reg.* **52,** 36334-36371.

Whittaker, R.H. (1975) *Communities and Ecosystems.* Macmillan, New York.

Wilhm J. L. and Dorris T. C. (1966) Species diversity of benthic macroinvertebrates in a stream receiving domestic and oil refinery effluents. *Am. Midl. Nat.* **76,** 427-449.

Wolda H. (1981) Similarity indices, sample size and diversity. *Oecologia* **50,** 296-302.

Zimmerman G. M., Goetz H. and Mielke P. W. Jr. (1985) Use of an improved statistical method for group comparisons to study effects of prairie fire. *Ecology* **66,** 606-611.

EPILOGUE

Summary

Chapter 1 addressed some of the requirements for establishing and maintaining evironmentally realistic laboratory-based stream microcosms, and the degree to which the microcosms reflected the natural communities from which they were derived. The following conclusions were drawn from the experiment described in chapter 1:

- The FT-C and S-C stream designs maintained most taxa at or above initial densities, and even in the FT-NC and S-NC streams densities of some taxa were not significantly different ($P \leq 0.05$) from initial densities.

- Emergent adults comprised a large proportion of mayfly and chironomid densities and should be monitored during bioassays with aquatic insects.

- Microcosms of riffle insect communities can be maintained for at least 30 days with moderate current and minimal flow-through.

- Relative to benthic samples taken directly from the source riffle, the artificial substrates selected for collector-filterers and against collector-gatherers.

The goal of the research presented in Chapter 2 was to determine if effluent-induced changes in macroinvertebrate and periphyton communities in laboratory stream microcosms could be used to predict macroinvertebrate and peripyhton responses in a natural system receiving the same waste. The conclusions drawn were the following:

- The control stream microcosms were able to maintain most taxa at or above initial densities during a 20 day bioassay.

- Mayfly densities in the microcosms were significantly ($P \leq 0.05$) reduced at 1.0 or 10.0% effluent depending on species, and a similar mayfly response was observed in the receiving stream.

- Mayflies responded at much lower concentrations than acute tests with *Daphnia magna* would have predicted.

- Hydropsychids in the microcosms were apparently not affected by the effluent, but exhibited increased densities at DN2.

- Chironomids and periphyton were stimulated both in the microcosms and the receiving stream.

- Overall, the stream microcosms accurately predicted the macroinvertebrate and periphyton response observed in the field.

Standard single-species tests were conducted simultaneously with the multispecies tests, thus allowing comparisons of multispecies microcosm tests to other more conventional tests (Chapter 3). The comparisons reported in Chapter 3 are rarely made, but are important to the incorporation of multispecies toxicity tests in the regulatory framework and the development of a hierarchy of toxicty testing protocols. Results of the comparisons revealed that:

- The significant decrease in density of some mayflies, and the significant change in protozoan community composition at an effluent concentration of 1.0% were the most sensitive indicators of effluent toxicity.

- The most sensitive single species test was the reproductive test with *Ceriodaphnia dubia*, which indicated that reproduction was inhibited at an effluent concentration of 3.0%.

- The acute single species tests, even with a large application factor, would not have predicted the elimination of mayflies in the receiving stream.

One of the principle arguments against the use of multispecies toxicity tests is that they are harder to interpret than single-species tests (Loewengart and Maki 1985; Mount 1985; Tebo 1985). In chapter 4 I evaluated some mathematical and statistical techniques for handling the data sets generated during the course of multispecies toxicity testing. Relatively simple techniques that provide unambiguous interpretations of results must be developed before multispecies tests will be accepted in the regulatory framework. Comparisons among canonical discriminant analysis, diversity indices, and community comparison indices indicated that the Bray-Curtis index (Bray and Curtis 1957) may be the most meaningful summary measure of community response in multispecies toxicity tests. However, none of the indices or techniques tested will ever replace sound ecological judgement in interpreting multispecies toxicity test results. For this reason, incorporation of multispecies tests into the regulatory framework will probably be opposed by traditional toxicologists who have chosen to ignore the ecology of test organisms.

Putting some ecology into ecotoxicology

The idea of conducting multispecies toxicity tests with lotic macroinvertebrates is not new. Channels (mesocosms) established along streams for ecological and

toxicological investigations have been colonized by benthic macroinvertebrates from the main stream and then dosed with a toxicant (*c.f.*, Giesy *et al.* 1979; Stout and Cooper 1983; Allard and Moreau 1987) or used in an ecological experiment (*c.f.*, Mundie *et al.* 1983; Irvine 1985). Burks and Wilhm (1977) were the first to colonize insects on artificial substrates and transfer them to streamside channels for toxicity tests, and Perkins (1983) demonstrated that colonized artificial substrates could be transferred long distances. It was this latter paper, that motivated me to try to develop a laboratory-based protocol for multispecies toxicity tests with lotic macroinvertebrates.

In designing and testing the laboratory stream microcosms, I attempted to infuse some "ecology" into ecotoxicology. Too often in the past, workers have not considered the ecology of test organisms in designing, conducting, and interpreting ecotoxicological research. For example, insects collected from a cool mountain stream have been tested in static aquaria in ambient laboratory lighting at 20°C. Temperature, photoperiod, and nutrition all influence the fitness and life history of aquatic insects (see Sweeney 1984), and should be accounted for when conducting bioassays with aquatic insects. I controlled microcosm temperature and photoperiod so that these variables were very similar to ambient conditions in the source ecosystem. I also attempted to provide a natural food source by seeding the microcosms with periphyton prior to the introduction of macroinvertebrates. Minshall (1984) states that "the substratum is the stage upon which the drama of aquatic insect ecology is acted out." Although it is often neglected or poorly modeled in designing toxicity tests with lotic insects, substrate is important to aquatic insects as an attachment site, a source of case-building materials, and as a collector of detrital food (see Minshall 1984 and Culp *et al.* 1983; Culp and Davies 1985). In addition, the substrate provides shelter from hydrologic forces (see Newbury 1984), and vertebrate (Brusven and Rose 1981) and invertebrate (Pontasch 1988) predators. The substrate used in my microcosms was as similar to the natural substrate as size considerations would allow. In my opinion, the methods I used either reduced or eliminated many of the confounding effects of testing under unrealistic conditions. The following paragraphs

outline some other ecological considerations that must be dealt with when conducting meaningful bioassays with communities of lotic insects.

In a natural system, lotic insects require current for gathering food, respiration, predation avoidance, and colonization. During a short-term single species bioassay with lotic insects, the primary reason for providing a current would be to avoid respiratory stress. However, in some cases, stream-dwelling organisms have been tested in aquaria lacking a current (*e.g.*, Warnick and Bell 1969; Sanders and Chandler 1972 ; Sloof 1983; Lemke and Anderson 1984; Green *et al.* 1986; Williams *et al.* 1986). Although some lotic insects can survive without a current, my results indicate that current is necessary for maintaining most taxa, especially mayflies. Although flow-through is often provided during bioassays, my results indicate that flow-through is not as important as current in maintaining lotic insects.

Most investigators trained in aquatic entomology know that many species of aquatic insects emerge as adults throughout the summer months. Twenty-five years ago Surber and Thatcher (1963) pointed out the necessity of monitoring adult emergence during bioassays with aquatic insects. In a single species test if a known number of early instar insects is being tested, it is not necessary to monitor emergence, but during multispecies tests it is necessary to monitor adult emergence to insure accurate interpretations of test results. However, most investigators do not monitor insect emergence during multispecies bioassays (*e.g.*, Burks and Wilhm 1977; Giesy *et al.* 1979; Perkins 1983; Hedtke 1984; Stout and Cooper 1983; Allard and Moreau 1987). It is likely that some of the "mortality" observed in these investigations was due to the emergence of adults. One could argue that the control would prevent a misinterpretation of the data, but stressor-induced stimulation (*c.f.*, Maki *et al.* 1975) or inhibibition (*c.f.*, Clubb *et al.* 1975) of emergence could still confound correct interpretations.

Another ecological factor that has received considerable attention from aquatic entomologists is the drift of stream insects. Three major types of drift, catastrophic, behavioral and constant, were proposed by Waters (1972). It was surprising that a catastrophic drift response to the effluent was not found in this investigation. Although not

monitored, behavioral and constant drift surely occurred in the stream microcosms. Screens placed over the standpipe in each stream microcosm prevented the loss of drifting insects. Although it may seem like a minor detail, had this loss not been prevented, organisms that drifted out the the streams would have been counted as mortalities at the end of the experiment. In on-site studies, where stream water is used as dilution water, it is important to prevent the drift of insects into the stream mesocosms. Drift into and out of the test system has not been controlled in most previous multispecies bioassays with aquatic insects (*e.g.*, Burks and Wilhm 1977; Giesy *et al.* 1979; Perkins 1983; Stout and Cooper 1983; Allard and Moreau 1987). However, a study into the effects of volcanic ash from Mount St. Helens, Washington, that did investigate insect drift, found higher drift rates in stream mesocosms receiving the ash (personal observation).

An awareness of potential counfounding factors and the willingness to control for them is the essence of a good experimental design and therefore, "good science." Although not all confounding factors can be controlled, for instance recruitment took place in my microcosms, an investigator should only be limited by expertise, money, and manpower in accounting for all confounding factors. It is my hope that after certain portions of this manuscript are published in the peer-reviewed literature, ignorance will no longer be an excuse for failing to take ecological factors into account during toxicity testing.

References

Allard, M. & G. Moreau, 1987. Effects of experimental acidification on a lotic macroinvertebrate community. Hydrobiologia 144:37-49.

Bray, J.R. and Curtis J.T. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27:325-349.

Brusven, M. A. and S. T. Rose. 1981. Influence of substrate composition and suspended sediment on insect predation by the torrent sculpin *Cottus rotheus* Can. J. Fish. Aquat. Sci. 38:1444-1448.

Burks, S.L. and J.L. Wilhm. 1977. Bioassays with a natural assemblage of benthic macroinvertebrates. In. [F.L. Mayer and J.L. Hamelink eds.] Aquatic Toxicology and Hazard Evaluation, ASTM STP 634, American Society for Testing and Materials.

Clubb, R.W., A.R. Gaufin, and J.L. Lord. 1975. Acute cadmium toxicity studies upon nine species of aquatic insects. Environmental Research 9:332-341.

Culp, J. M., S. J. Walde & R. W. Davies, 1983. Relative importance of substrate particle size and detritus to stream benthic macroinvertebrate distribution. Can. J. Fish. Aquat. Sci. 40:1568-1574.

Culp, J. M., and R. W. Davies. 1985. Responses of benthic macroinvertebrate species to manipulation of interstitial detritus in Carnation Creek, British Columbia. Can. J. Fish. Aquat. Sci. 42:139-146.

Giesy, J.P., Jr. H.J. Kania, J.W. Bowling, R.I. Knight, S. Mashburn, and S. Clarkin. 1979. Fate and biological effects of cadmium introduced into channel microcosms. EPA-6002/3-79-039. USEPA, Athens, GA.

Green, D. W. J., K. A. Williams & D. Pascoe, 1986. Studies on the acute toxicity of pollutants to freshwater macroinvertebrates 4. Lindane (γ—Hexachlorocyclohexane). Arch. Hydrobiol. 106:263-273.

Hedtke, S. F., 1984. Structure and function of copper stressed aquatic microcosms. Aquatic Toxicology 5:227-244.

Irvine, J. R. 1985. Effects of successive flow perturbations on stream macroinvertebrates. Can. J. Fish. Aquat. Sci. 42:1922-1927.

Lemke, A. E., and R. L. Anderson. 1984. Insect interlaboratory toxicity test comparison study for the chironomid (*Paratanytarsus* sp.) Procedure. EPA-600/S3-84-054. USEPA, Duluth, MN.

Loewengart G., and A.W. Maki. 1985. Multispecies toxicity tests in the safety assessment of chemicals: necessity or curiosity?. In. [J. Cairns, Jr. Ed.] Multispecies toxicity testing. Pergamon Press. 261 p.

Maki, A.W., L. Geissel and H.E. Johnson. 1975. Comparative toxicity of larval lampricide (TFM: 3-triflouromethyl-4-nitrophenol) to selected benthic macroinvertebrates. Journal of the Fisheries Research Board of Canada 32:1455-1459.

Minshall, G. W. 1984. Aquatic insect-substratum relationships, p. 358-400. In: V.H. Resh and D.M. Rosenberg (ed.) The Ecology of Aquatic Insects. Praeger Publishers New York, NY.

Mount, D. I. 1985. Scientific problems in using multispecies toxicity tests for regulatory purposes. In Cairns, J., Jr. [ed.] 1985. Multispecies toxicity testing. Pergamon Press, New York, NY. 261 p.

Mundie, J. H., S. M. McKinnel, and R. E. Traber. 1983. Responses of stream zoobenthos to enrichment of gravel substrates with cereal grain and soybean. Can. J. Fish. Aquat. Sci. 40:1702-1712.

Newbury, R. W. 1984. Hydrologic determinants of aquatic insect habitats, p. 323-357. In: V.H. Resh and D.M. Rosenberg (ed.) The Ecology of Aquatic Insects. Praeger Publishers New York, NY.

Perkins, J. L. 1983. Bioassay evaluation of diversity and community comparison indexes. J. Wat. Pollut. Control Fed. 55:522-530.

Pontasch, K. W. 1988. Predation by *Paragnetina fumosa* (Banks) (Plecoptera:Perlidae) on mayflies. American Midland Naturalist.

Sanders, H. O., and J. H. Chandler. 1972. Biological magnification of a polychlorinated biphenyl (Aroclor® 1245) from water by aquatic invertebrates. Bull. Env. Cont. and Tox. 7:257-263.

Slooff, W., 1983. Benthic macroinvertebrates and water quality assessment: some toxicological considerations. Aquat. Toxicol. 4:73-82.

Stout, R.J., and W.E. Cooper. 1983. Effect of p-cresol on leaf decomposition and invertebrate colonization in experimental outdoor streams. Can J. Fish. Aquat. Sci. 40:1647-1657.

Surber, E. W. and T. O. Thatcher. 1963. Laboratory studies of the effects of alkyl benzene sulfonate on aquatic invertebrates. Trans. Am. Fish. Soc. 92:152-160.

Sweeney, B. W. 1984. Factors influencing the life-history of aquatic insects. p. 56-100. In: V.H. Resh and D.M. Rosenberg (ed.) The Ecology of Aquatic Insects. Praeger Publishers New York, NY.

Tebo, L. B., Jr. 1985. Technical considerations related to the regulatory use of multispecies toxicity tests. In Cairns, J., Jr. [ed.] 1985. Multispecies toxicity testing. Pergamon Press, New York, NY. 261 p.

Warnick, S. L., and H. L. Bell. 1969. The acute toxicity of some heavy metals to different species of aquatic insects. J. Wat. Pollut. Control Fed. 41:280-284.

Waters, T. F. 1972. The drift of stream insects. In: Annual Review of Entomology, Vol. 17 Ent. Soc. of Am. 555pp.

Williams, K. A., D. W. J. Green & D. Pascoe, 1986. Studies on the acute toxicity of pollutants to freshwater macroinvertebrates, 3. Ammonia. Arch. Hydrobiol. 106:61-70.

Appendix 1. SAS ® program to calculate the Bray-Curtis Canbera-Metric, Pinkham-Pearson, Euclidean Distance, Average chi-square, and SIMI community comparison indices, the Shannon and Simpson diversity indices and Pielou's evenness.

```
TITLE'SAS PROGRAM TO COMPUTE VARIOUS INDICES';
TITLE'PREPARED BY KURT W. PONTASCH 3/88';
*TA=CANBERRA METRIC TB=BRAY-CURTIS TC=EUCLIDIAN DISTANCE;
*CHI=AVERAGE CHI-SQUARE PP=PINKHAM-PEARSON SIMI=SIMI;
*SHANN= SHANNON'S DIVERSITY SIMP=SIMPSON'S DIVERSITY EQUIT=EQUITABILITY;
*INPUT THE DATA AS IN A REGULAR SAS DATA SET. THE TYPE VARIABLE BELOW
*REFERRED TO ADULTS VS LARVAE. THE VARIABLES ISO--DSUM ARE VARIOUS TAXA;
DATA GBR;
INPUT TRT $ STREAM $ TYPE $ ISO CAE PSE BAE PAR STA STN LEU ESUM
HYD CHE TSUM BOE ORT TAN CSUM HEM ATH DSUM @@;
CARDS;


*DATA GOES HERE


PROC SORT;
BY TRT;
PROC MEANS NOPRINT SUM;
BY TRT;
VAR ISO CAE PSE BAE PAR STA STN LEU
HYD CHE BOE ORT TAN HEM ATH;
OUTPUT OUT=TWO
SUM=ISOS CAES PSES BAES PARS STAS STNS LEUS
HYDS CHES BOES ORTS TANS HEMS ATHS;
*THE DATA SET TWO IS THE SUM OF ADULTS AND LARVAE OVER TREATMENTS;
*AND THE ABOVE STUFF COULD BE DELETED FOR MOST CASES;
PROC PRINT;
PROC TRANSPOSE DATA=TWO OUT=KEEP1 PREFIX=TRT NAME=SPECIES;
PROC PRINT;
DATA NEW;SET KEEP1 END=E;
* THIS PART IS FOR PAIRWISE TESTS;
ARRAY X1(I) TRT1-TRT5;ARRAY X2(J) TRT1-TRT5;
*THIS PROGRAM IS SET TO COMPARE 5 TREATMENTS PAIRWISE;
*SO THEIR ARE 10 COMPARISONS. HENCE THE ARRAYS BELOW GO FROM 1-10;
ARRAY TA(K) TA1-TA10;ARRAY TB(K) TB1-TB10 ;ARRAY TC(K) TC1-TC10;
ARRAY CHI(K) CHI1-CHI10; ARRAY PP(K) PP1-PP10;ARRAY SIMA(K) SIMA1-SIMA10;
ARRAY SIMB(K) SIMB1-SIMB10; ARRAY SIMC(K) SIMC1-SIMC10;
ARRAY SIMI(K) SIMI1-SIMI10;
ARRAY TOTA(K) TOTA1-TOTA10;ARRAY TOTB (K) TOTB1-TOTB10;
RETAIN TA1-TA10 TB1-TB10 TC1-TC10 TOTA1-TOTA10 SIMI1-SIMI10
TOTB1-TOTB10 CHI1-CHI10 PP1-PP10 SIMA1-SIMA10 SIMB1-SIMB10 SIMC1-SIMC10;
K=0;DO I=1 TO 4; DO J=I+1 TO 5; K+1;
*I ON THE LINE ABOVE SHOULD BE EQUAL TO THE NUMBER OF TRTS-1;
*J SHOULD = THE NUMBER OF TREATMENTS;
A=X1-X2;B=X1+X2;
ASQ=A*A;
C=B/2;D=X2-C;DSQ=D*D;DSQC=DSQ/C;
AA=MIN(X1,X2); BB=MAX(X1,X2);CC=AA/BB;
DD=X1*X2;EE=X1*X1;FF=X2*X2;
SIMA+DD;SIMB+EE;SIMC+FF; IF E=1 THEN SIMI=SIMA/SQRT(SIMB*SIMC);
TA+ABS(A/B); IF E=1 THEN TA=1-(TA/_N_);
```

```
TOTB+ABS(B);TOTA+ABS(A);IF E=1 THEN TB=1-(TOTA/TOTB);
TC+A*A;IF E=1 THEN TC=SQRT(TC);
CHI+DSQC;IF E=1 THEN CHI=1-(2*(CHI/TOTB));
PP+CC;IF E=1 THEN PP=PP/15;
*THE 15 ON THE LINE ABOVE IS THE NUMBER OF SPECIES PRESENT BETWEEN THE;
*TWO COMMUNITIES BEING COMPARED;
END;END;
DATA NEW; SET NEW END=E; IF E=1;
PROC PRINT; VAR TA1-TA10 TB1-TB10 TC1-TC10 CHI1-CHI10
PP1-PP10 SIMI1-SIMI10;
*NOW WE CALCULATE DIVERSITY MEASURES. (ISOS--ATHS) BELOW
*SHOULD BE CHANGED TO;
*MATCH YOUR FIRST AND LAST VARIABLES OF COMPARISON;
DATA TWO;SET TWO END=E; SUM=SUM(OF ISOS--ATHS); S=0;
ARRAY SP(I) ISOS--ATHS;ARRAY P(I) P1-P15; SUMP=0; SUMPSQ=0;
RETAIN P1-P15;
*THE ARRAYS ABOVE SHOULD EQUAL THE NUMBER OF TAXA IN THE COMMUNITY;
*IN THIS CASE 15;
DO OVER SP;P=SP/SUM;SUMPSQ+P*P;P=P*LOG2(P);SUMP+P;IF SP NE 0 THEN S+1;
END; SIMP=1-SUMPSQ; SHANN=-SUMP; EQUIT=-SUMP/LOG2(S);
PROC PRINT; VAR SIMP SHANN EQUIT;
*GOOD LUCK!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!;
```

Appendix 2: Fortran program used to compute similarity measures and carry out permutation tests to determine if significant differences exist among the communities being tested.

```
C
C THIS PROGRAM CARRIES OUT PERMUTATION TESTS FOR SIMILARITY MEASURES
C AND CARRIES OUT THE MULTIPLE COMPARISON PROCEDURE OF FOUTZ
C (BIOMETRICS 1985)
C
      DIMENSION X(20,70),SIM(1000),INDEX(100),W(10)
      DIMENSION XI(70),YI(70),SIMAT(20,20),NVEC(10)
      DIMENSION SIM2(1000,6),BET(20),XLMBDA(20,20)
      DIMENSION NWIN(20),NBET(20)
      DATA XI/70*0.0/,YI/70*0.0/,XLMBDA/400*1.0/
      DATA X/1400*0.0/,SIM/1000*0.0/
C INPUT DATA
C NGRP IS THE NUMBER OF GROUPS, FACTOR LEVELS ETC
C NCLAS IS THE NUMBER OF CLASSES, VARIABLES OR SPECIES
C NTEST IS THE NUMBER OF DATA SETS TO BE ANALYZED
      READ(5,21) NGRP,NCLAS,NTEST
 21    FORMAT(3I2)
C NVEC IS THE VECTOR CONTAINING THE NUMBER OF REPS PER GROUP
C      READ(5,23) (NVEC(I),I=1,NGRP)
      NVEC(1)=4
      NVEC(2)=4
      NVEC(3)=4
      NVEC(4)=4
 23    FORMAT(10(I2,1X))
      DO 500 MK=1,NTEST
      WRITE(6,501) MK
 501   FORMAT(1H1,/,2X,'SUMMARY FOR TEST NUMBER ',I3)

      NTOT=0
      DO 12 I=1,NGRP
 12    NTOT=NTOT+NVEC(I)
C THIS DATA SET CONTAINS ONE SPECIES PER LINE
      DO 14 I=1,NTOT
C        DO 16 J=1,NCLAS
C        READ(5,502) X(I,J)
C        X(I,J)=X(I,J)/1000.0
C16       CONTINUE
C502      FORMAT(16X,F10.0)
C14    CONTINUE
C ALTERNATE WITH NCLAS OBSERVATIONS PER LINE
 14    READ(5,*) A,B,C,(X(I,J),J=1,NCLAS)
      WRITE(6,24) NGRP,NCLAS,(NVEC(K),K=1,NGRP)
 24    FORMAT(//,2X,'NUMBER OF GROUPS OR TRTS ',I2,
     1        /,2X,'NUMBER OF SPECIES ',I3,
     2        /,2X,'NUMBER OF REPS PER GROUP ',10(I2,1X))
      WRITE(6,25)
 25    FORMAT(2X,'DATA MATRIX AS INPUT',//)
      CALL OUTMAT(X,NTOT,NCLAS,20,70)

       DO 2 I=1,NTOT
         DO 8 K=1,NCLAS
         YI(K)=X(I,K)
 8        CONTINUE
```

```
          DO 4 J=1,NTOT
            DO 6 K=1,NCLAS
            XI(K)=X(J,K)
 6           CONTINUE
          CALL SIMI(XI,YI,NCLAS,D)
          SIMAT(I,J)=D
 4         CONTINUE
 2        CONTINUE
      WRITE(6,42)
 42    FORMAT(/,2X,'MATRIX OF SIMILARITY COEFFICIENTS')
C
C OUTPUT SIMILARITY MATRIX
C 20,20 ARE THE DIMENSIONS OF THE MATRIX IN OUTMA2 - CHANGE
C AS NECESSARY

      CALL OUTMA2(SIMAT,NTOT,NTOT,20,20)

C ---- COMPUTE WITHIN AND TOTAL SIMILARITY
      TOT=0.0
      DO 30 I=1,NTOT
        IP1=I+1
        DO 32 J=IP1,NTOT
        TOT=TOT+SIMAT(I,J)
 32      CONTINUE
 30    CONTINUE
      DO 50 I=1,NTOT
 50    INDEX(I)=I
      CALL WB(SIMAT,NGRP,INDEX,NVEC,D,W,BET)
      BSIM=TOT-D
C
C       COMPUTE THE PAIRWISE LAMBDA MEASURES OF GOOD (1982)
        IA=0
        TWIN=0.0
        DO 36 IG=1,NGRP
          NWIN(IG)=(NVEC(IG)*NVEC(IG)-NVEC(IG))/2
          TWIN=TWIN + NWIN(IG)
 36      CONTINUE
        DO 37 IG=1,NGRP
          DO 38 JG=IG,NGRP
          IF(IG.EQ.JG) GO TO 38
          IA=IA+1
          NBET(IA)=NVEC(IG)*NVEC(JG)
          BET(IA)=BET(IA)/FLOAT(NBET(IA))
          WBAR=(W(IG)+W(JG))/(NWIN(IG)+NWIN(JG))
          XLMBDA(IG,JG) = BET(IA)/WBAR
 38        CONTINUE
 37     CONTINUE

          DO 391 IG=1,NGRP
 391        W(IG)=W(IG)/FLOAT(NWIN(IG))
          TOTN= NTOT*(NTOT-1)/2
          TOTBAR=TOT/TOTN
          BBAR=BSIM/(TOTN-TWIN)
          WBAR=D/TWIN
      WRITE(6,33) TOT,TOTBAR,BSIM,BBAR,D,WBAR
   2          ,(W(IM),IM=1,NGRP)
   1          ,(BET(IC),IC=1,IA)
 33    FORMAT(/,2X,'TOTAL SIM,  MEAN =  ',E12.4,2X,E12.4,
```

```
     1       /,2X,'BETWEEN SIM,  MEAN  = ',E12.4,2X,E12.4,
     2       /,2X,'WITHIN SIM,  MEAN  =  ',E12.4,2X,E12.4,
     3       /,2X,'WITHIN BY GROUP = ',4(E12.4,1X)
     4       /,2X,'BETWEEN BY COMB = ',8(E12.4,1X))
       WRITE(6,35)
 35    FORMAT(/,3X,'LAMBDA VALUES FOR PAIRWISE COMPARISONS')
       DO 47 IG=1,NGRP
 47    WRITE(6,39) (XLMBDA(IG,JG),JG=IG,NGRP)
 39    FORMAT(5X,8(F8.4,2X))
       WSIM=D/TWIN
C START SIMULATION PROCESS
C NSIM IS THE NUMBER OF PERMUTATIONS
C RPERM CARRIES OUT THE SWITCHES
C
      NSIM=1000
      DO 10 I=1,NSIM
      CALL RPERM(NTOT,INDEX)
      CALL WB(SIMAT,NGRP,INDEX,NVEC,D,W,BET)
      SIM(I)=D/TWIN
C NOTE THAT THE WITHIN TOTAL SIMILARITY IS D AND IS USED TO
C TEST FOR DIFFERENCES

C OBTAIN THE MULTIPLE COMPARISONS USING THE BETWEEN OR LAMBDA VALUES
      IV=0
       DO 100 I1=1,NGRP
         DO 102 I2=I1,NGRP
         IF(I2.EQ.I1) GO TO 102
         IV=IV+1
         BBAR=  BET(IV)/FLOAT(NBET(IV))
         SIM2(I,IV)=  BBAR

C  -- THIS APPROACH USES THE LAMBDA VALUES AS THE CRITICAL VALUES
C         WBAR=(W(I1) + W(I2))/(NWIN(I1)+NWIN(I2))
C         SIM2(I,IV)=  BBAR/WBAR

 102      CONTINUE
 100     CONTINUE
C      OPTIONAL OUTPUT TO FIND THE COMBINATIONS EXEEDING THE
C                  ORIGINAL VALUE OF THE WITHIN SIMILARITY
C      IF(D.GE.WSIM) WRITE(6,13) D,(INDEX(IK),IK=1,9),(W(JK)
C    1  ,JK=1,NGRP)
 13    FORMAT(2X,'DISTANCE ',F8.5,3X,9(I2,2X),
     1 '  WITHIN MEAN SIMI ',6(F8.3,2X))
 10    CONTINUE
C
C ADUP IS AN SUMMARY ROUTINE FOR MIN,MAX,MEAN,VARIANCE
C      OF THE WITHIN TOTALS FOR THE NSIM PERMUTATIONS
C      THESE VALUES MAY BE USED TO OBTAIN AN IDEA AS TO HOW
C      FAR AWAY THE OBSERVED VALUE OF WITHIN SIMILARITY IS
C      FROM THE PERMUTED VALUES

      CALL ADUP(SIM,NSIM,XM,XV,XMIN,XMAX)
      WRITE(6,11) XM,XV,XMIN,XMAX
 11    FORMAT(/,'  MEAN,VAR,MIN,MAX ',4(F8.5,2X))
C
C FIND THE CRITICAL VALUE FOR A 95 PCT TEST AND THE P VALUE
      NS1=0.95*NSIM  +1
      CALL ORDERS(SIM,1,NSIM,NS1)
```

```
      WRITE(6,101) NS1,SIM(NS1)
 101   FORMAT(/,2X,'INDEX OF CRITICAL VALUE ',I3,'- CRITICAL',
     1 ' VALUE ',E12.4)
      CALL PVAL(SIM,NSIM,XPV,WSIM)
      WRITE(6,480) XPV
 480   FORMAT(2X,'NUMBER GT. OR EQ. TO OBSERVED WITHIN ',F8.1)
C
C - OBTAIN THE SORTED MULTIPLE COMPARISONS
C  THE EXP WISE ERROR HERE IS SET TO ALPHA=0.06 = 60/1000
      CALL SORTA(SIM2,10,6,NSIM)
C
C CALL ROUTINE TO OBTAIN THE INFLUENCE OF THE SPECIES
C ON THE BETWEEN SIMILARITIES
C
     CALL REMOVE(X,NTOT,NCLAS,BSIM,NGRP,NVEC)

 500   CONTINUE
      STOP
      END


      SUBROUTINE SORTA(A,K,IR,N)
C A IS THE MATRIX (N BY R) OF SCORES
C IR IS THE NUMBER OF COMPARISONS
C KR IS THE NUMBER OF TESTS THAT WILL BE SORTED TO OBTAIN THE
C    ALPHA LEVELS - EXP WISE ERROR = KR/NSIM
C N IS THE TOTAL NUMBER OF SIMULATED TESTS
C THE DIMENSION ON A AND TEST ARE THE NUMBER OF CPMPARISONS

      DIMENSION A(1000,6),INDEX(100),TEST(6)
      KR=K*IR
C CHANGE THE INDEX VALUES FROM 1 T KR TO K SETS OF 1 TO R
      CALL RPERM(KR,INDEX)

      DO 10 I=1,KR
  10   INDEX(I)=MOD(INDEX(I),IR) + 1

C START THE SORT SEQUENCE
      DO 20 J=1,KR
      K1=INDEX(J)
      INDX2=J
      W=A(J,K1)
C  SELECT THE COLUMN TO SORT ON AND FIND THE SMALLEST VALUE

        DO 30 J2=J,N
        IF(J2.EQ.J) GO TO 30
        IF(A(J2,K1) .GT. W) GO TO 30
        W=A(J2,K1)
C RECORD INDEX OF SMALLEST VALUE
        INDX2=J2
 30      CONTINUE

C NOW SWITCH THE ROW WITH THE SMALLEST VALUE WITH THE ROW J

        DO 40 J3=1,IR
        Z=A(INDX2,J3)
        A(INDX2,J3)=A(J,J3)
        A(J,J3)=Z
```

```
 40       CONTINUE
C
          DO 50 I=1,IR
          IF(I.EQ.K1) TEST(I)=A(J,I)
          IF(I.NE.K1) TEST(I)=0.0
 50        CONTINUE
C
C PRINT OUT THE COMPLETE SET OF COMPARISONS.  THE LAST VALUE
C FOR EACH COMPARISON NUMBER GIVES THE CRITICAL VALUE FOR
C THE COMPARISON

         WRITE(6,11) K1,J,(TEST(I),I=1,IR)
 11       FORMAT(2X,'COMPARISON ',I3,2X,I3,2X,10(F8.5,1X))

 20       CONTINUE
          WRITE(6,111)
 111        FORMAT(/,2X,'THE CRITICAL VALUES FOR THE MULTIPLE',
    1  /,2X,'COMPARISONS ON THE LAMBDA VALUES ARE OBTAINED ',
    2  /,2X,'BY FINDING THE LAST OCCURRENCE OF THE COMPARISON',
    3  /,2X,'REJECT FOR SMALL VALUES OF THE BETWEEN SIMILARITY',
    4  /,2X,'OR SMALL VALUES OF LAMBDA ')
          RETURN
          END
C
C
C  COMPUTE THE SIMI INDEX AS THE DISTANCE MEASURE
C
      SUBROUTINE SIMI(XI,YI,NCLAS,D)
      DIMENSION XI(71),YI(71),P(71),Q(71)

      T1=0.0
      T2=0.0
      XY=0.0
      XS=0.0
      YS=0.0
      XF=0.0
      XC=0.0
      XEE=0.0
      XEC=0.0
      XE=0.0
      XA=0.0
      XB=0.0
      XAB=0.0
      XCC=0.0
      XCMA=0.0
      XCMB=0.0
      XCMC=0.0
      PX=0.0
      PY=0.0
      PXX=0.0
      PYY=0.0
      PER=0.0

      DO 101 I=1,NCLAS
      T1=T1+XI(I)
      T2=T2+YI(I)
       PX  =  PX+(LOG10((XI(I)+1)))
       PY  =  PY+(LOG10((YI(I)+1)))
```

```
 101  CONTINUE

      DO 10 I=1,NCLAS
      PXX=XI(I)/T1
      PYY=YI(I)/T2
      PER=PER+MIN(PXX,PYY)
C     XY = XY + ABS(XI(I)-YI(I))
C     XS = XS + (XI(I)+YI(I))
C     XC = (XI(I)+YI(I))/2
C     XE = (XI(I)-XC)
C     IF(XC.LE.0.0) XC=1.0
C     XEE = XE*XE
C     XEC = XEE/XC
C     XF = XF+XEC
C     XY = XY + XI(I)*YI(I)
C     XS = XS + XI(I)**2
C     YS = YS + YI(I)**2
C     XA = MIN(XI(I),YI(I))
C     XB = MAX(XI(I),YI(I))
C     IF(XA.LE.0.0) XA=1.0
C     IF(XB.LE.0.0) XB=1.0
C     XAB = XA/XB
C     XCC =XCC+XAB
C     XCMA=ABS(XI(I)-YI(I))
C     XCMB=(XI(I)+YI(I))
C     IF(XCMB.LE.0.0) XCMB=1.0
C     XCMC=XCMC+(XCMA/XCMB)
C  FORMULAS BELOW ARE FOR SQUARE ROOT  OR LOG TRANS ON ORIGINAL DATA
C     XY = XY + ABS(LOG10((XI(I)+1))-LOG10((YI(I)+1)))
C     XS = XS + (SQRT(XI(I))+SQRT(YI(I)))
C     XC = (SQRT(XI(I))+SQRT(YI(I)))/2
C     XE = (SQRT(XI(I))-XC)
C     IF(XC.LE.0.0) XC=1.0
C     XEE = XE*XE
C     XEC = XEE/XC
C     XF = XF+XEC
C     XY = XY + (SQRT(XI(I))*SQRT(YI(I)))
C     XS = XS + (XI(I))
C     YS = YS + (YI(I))
C     XA = MIN(XI(I),YI(I))
C     XB = MAX(XI(I),YI(I))
C     IF(XA.LE.0.0) XA=1.0
C     IF(XB.LE.0.0) XB=1.0
C     XAB = XA/XB
C     XCC =XCC+XAB
C     XCMA=ABS(XI(I)-YI(I))
C     XCMB=(XI(I)+YI(I))
C     IF(XCMB.LE.0.0) XCMB=1.0
C     XCMC=XCMC+(XCMA/XCMB)
C     PXX= LOG10(XI(I)+1.0)/PX
C     PYY= LOG10(YI(I)+1.0)/PY
C     PER=PER+MIN(PXX,PYY)
 10   CONTINUE
C     D = XY/SQRT(XS*YS)
C     D = 1-(XY/XS)
C     D = 1-(2*(XF/XS))
C     D = XCC/15
C     D = 1-(XCMC/15)
```

```
      D = PER
      RETURN
      END
C
C  COMPUTE THE P VALUE FOR THE TEST
C
      SUBROUTINE PVAL(S,N,X,XOBS)
      DIMENSION S(1000)
      X=0.0
      DO 10 I=1,N
      IF(S(I).GE.XOBS) X=X+1
 10    CONTINUE
      RETURN
      END
C-------------------------------
      SUBROUTINE ADUP(X,N,XM,XV,XMIN,XMAX)
C--1ADUP-------------------------
C
C    COMPUTE MEAN VARIANCE MIN AND MAX  OF VECTOR X
C
      DIMENSION X(1000)
      XM=0.0
      XV=0.0
      XMIN=X(1)
      XMAX=X(1)
        DO 10 I=1,N
        IF(X(I).GT.XMAX) XMAX=X(I)
        IF(X(I).LT.XMIN) XMIN=X(I)
        XM=XM+X(I)
        XV=XV+X(I)**2
 10      CONTINUE
        XN=FLOAT(N)
        XV=XV-(XM**2)/XN
        XV=XV/(XN-1.0)
        XM=XM/XN
        RETURN
        END
C
        SUBROUTINE RPERM(NTOT,INDEX)
C
C ---- OBTAIN A RANDOM PERMUTATION OF 1 TO NTOT
C     THIS ROUTINE COMES FROM NIJENHUS AND WILF - COMBINATORICAL
C     ALGORITHMS 1975 - QA164 N54
        DOUBLE PRECISION DSEED
        DIMENSION INDEX(100)
        DATA DSEED/12321357.D0/
        DO 10 I=1,NTOT
 10      INDEX(I)=I
        DO 40 M=1,NTOT
 30      L=M+GGUBFS(DSEED)*(NTOT+1-M)
        L1=INDEX(L)
        INDEX(L)=INDEX(M)
        INDEX(M)=L1
 40      CONTINUE
 25      CONTINUE
C       WRITE(6,11) (INDEX(J),J=1,NTOT)
 11      FORMAT(1X,'PERMS ', 10I3)
        RETURN
```

```
      END
    SUBROUTINE OUTMA2(X,NTOT,NC,N1,N2)
C
C    OUTPUT AN NR BY NC MATRIX X
C
    DIMENSION X(N1,N2)

    DO 10 I=1,NTOT
 10  WRITE(6,13) (X(I,J),J=1,NC)
 13  FORMAT(1X,12(F7.4,3X))
    RETURN
    END
    SUBROUTINE OUTMAT(X,NTOT,NC,N1,N2)
C
C    OUTPUT AN NR BY NC MATRIX X
C
    DIMENSION X(N1,N2)

    DO 10 I=1,NTOT
 10  WRITE(6,13) (X(I,J),J=1,NC)
 13  FORMAT(1X,11(F10.1,1X),/,2X,6(F11.0,1X))
    RETURN
    END
    SUBROUTINE WB(X,NGROUP,INDEX,NVEC,WTOT,W,BET)
C
C    THIS SUBROUTINE COMPUTES THE SUM OF THE WITHIN SIMILARITIES
C    FOR A MATRIX OF SIMILARITY MEASURES. NVEC CONTAINS GROUP
C    SIZES AND INDEX CONTAINS PERMUTED VALUES
C    CRITERIA IS MEAN WITHIN DISTANCE
C    W IS THE VECTOR OF WITHIN VALUES AND BET IS THE VECTOR OF BETWEENS
    DIMENSION X(20,20),INDEX(100),NVEC(10),W(10)
    DIMENSION BET(20)
    NTOT=0
    WTOT=0.0
    DO 5 I=1,NGROUP
 5  W(I)=0.0
    J1=0
    K1=0
     DO 10 II=1,NGROUP
        NV=NVEC(II)
      NV1=NV-1
      J1=NTOT
      NTOT=NTOT+NV
      DO 20 I=1,NV1
      J1=J1+1
      I1=INDEX(J1)
      IK=I+1
        DO 22 K=IK,NV
        J2=J1+K -I
        I2=INDEX(J2)
     W(II)=W(II)+X(I1,I2)
C      WRITE(6,24) I1,I2,X(I1,I2),J1,J2,K1
 24    FORMAT(2X,'I1,I2,X,J1,J2,K1 ',2I4,F8.5,3I4)
 22    CONTINUE
 20    CONTINUE
      WTOT=WTOT+W(II)
 10   CONTINUE
C
```

```
C -  COMPUTE THE BETWEEN VALUES
C
      NBET=0
      I1=0
      ISTRT=1
      IEND=0
      DO 30 I=1,NGROUP
      NV=NVEC(I)
      IEND=IEND+NV
      JSTRT=IEND+1
      JEND=IEND
        DO 40 J=I,NGROUP
        IF(I.EQ.J) GO TO 40
        NV2=NVEC(J)
        JEND=JEND+NV2
        NBET=NBET+1
        BET(NBET)=0.0
          DO 50 IA=ISTRT,IEND
          DO 60 IB=JSTRT,JEND
          IC=INDEX(IA)
          ID=INDEX(IB)
          BET(NBET)=BET(NBET) + X(IC,ID)
60         CONTINUE
50         CONTINUE

         JSTRT=JEND+1
40        CONTINUE

         ISTRT=IEND +1
30        CONTINUE



      RETURN
      END
     SUBROUTINE ORDERS(A,IB,IE,JO)
C --- FIND THE JO TH ORDERED VALUE OF AN ARRAY
C     A=ARRAY TO BE EXAMINED
C     IB=LOWER INDEX OF ARRAY
C     IE=UPPER INDEX OF ARRAY
C     JO=THE ABSOLUTE ORDER STATISTIC RELATIVE TO 1,N
     DIMENSION A(1)
     NN=IE+1
     IJ=IB-1
     K=NN
     I=IJ
  1  T=A(JO)
  11 K=K-1
     W=A(K)
     IF(W-T) 2, 3, 11
  3  IF(K.NE.JO) GO TO 11
     IF(I.EQ.K) RETURN
  2  I=I+1
     Z=A(I)
     IF(Z-T) 2,5,4
  5  IF(I.NE.JO) GO TO 2
     IF(I.EQ.K) RETURN
     A(I)=W
```

```
      A(K)=Z
      I=IJ
      NN=K
      GO TO 1
 4    IF(I.EQ.K) RETURN
      A(I)=W
      A(K)=Z
      IF(K.NE.JO) GO TO 11
      K=NN
      IJ=I
      GO TO 1
      END

      SUBROUTINE REMOVE(X,NTOT,NCLAS,B,NGRP,NVEC)
      DIMENSION SIMAT(20,20),BREM(20),W(10),BET(10)
      DIMENSION X(20,70),XI(70),YI(70),NVEC(10)
      DIMENSION INDX(100)
C
C COMPUTE THE IMPORTANCE VALUE FOR THE SPECIES
C
      DO 31 I=1,NTOT
 31   INDX(I)=I

      DO 10 IREM=1,NCLAS
        TOT=0.0
        DO 2 I=1,NTOT
          DO 8 K=1,NCLAS
          YI(K)=X(I,K)
          IF(K.EQ.IREM) YI(K)=0.0
 8        CONTINUE
          DO 4 J=1,NTOT
            DO 6 K=1,NCLAS
            XI(K)=X(J,K)
          IF(K.EQ.IREM) XI(K)=0.0
 6          CONTINUE
          CALL SIMI(XI,YI,NCLAS,D)
          SIMAT(I,J)=D
          IF(J.GT.I) TOT=TOT+D
 4        CONTINUE
 2      CONTINUE
      CALL WB(SIMAT,NGRP,INDX,NVEC,D,W,BET)
      BR=TOT-D
      BREM(IREM)=100*(BR-B )/B
      WRITE(6,35) TOT,BR,D,BREM(IREM)
 35   FORMAT(2X,'TOT, BET ,WITHIN, IMPT ',4(2X,F8.4))
 10   CONTINUE
      WRITE(6,24)
      WRITE(6,23) (I,BREM(I),I=1,NCLAS)
 24   FORMAT(//,2X,'SPECIES REMOVED   INFLUENCE',/)
 23   FORMAT(8X,I3,5X,F8.3)
      RETURN
      END
```

4

The ____ page vita has been removed from the scanned document

4

The ____ page vita has been removed from the scanned document

The 4 page vita has been removed from the scanned document

4

The ____ page vita has been removed from the scanned document