

*Ecology*, 77(3), 1996, pp. 681–685  
© 1996 by the Ecological Society of America

## BOTTLE OR BIG-SCALE STUDIES: HOW DO WE DO ECOLOGY?<sup>1</sup>

ANTHONY R. IVES, JOHANNES FOUFOPOULOS, ERIC D. KLOPPER,  
JENNIFER L. KLUG, AND TODD M. PALMER

*Department of Zoology, University of Wisconsin–Madison, Madison, Wisconsin 53706 USA*

### INTRODUCTION

While other articles in this Special Feature address the pros and cons of microcosm “bottle” experiments in ecology, the purpose of our article is to characterize the use and influence of microcosm experiments within one arbitrarily chosen field of ecology, the study of interactions among two or more species. We do this by asking four questions: (1) How common are microcosm experiments in the published literature? (2) Do microcosm experiments differ from other types of species-interaction studies either in the number of included species or in duration? (3) How influential are microcosm experiments as revealed by the number of citations they garner in the Science Citation Index? and (4) How influential are microcosm experiments as measured by their inclusion in undergraduate ecology textbooks? To answer these questions, we conducted a survey of ecological journals and textbooks. Our choices of journals and textbooks were to some degree capricious. However, even though our results may depend on these choices, the chosen journals and textbooks are certainly representative of the basic reading for most ecologists.

### METHODS

#### *Journal article survey*

We searched completely all articles published in *Ecology*, *The American Naturalist*, and the *Journal of Animal Ecology* at five-year intervals, in the years 1950, 1955, 1960, . . . , 1990, and selected those that addressed interactions among two or more species. (Actually, the 1955 volume of the *Journal of Animal Ecology* was missing from the University of Wisconsin–Madison library system, so we substituted 1954.) Although we were liberal in our decisions to include articles, studies that did not explicitly mention interactions among species were excluded even if they discussed many species. For example, we did not include descriptive studies of plant succession if they did not consider how the interactions among plants influenced plant replacements. Also, we restricted our attention to

full-length articles, because notes and comments often did not report primary results. In total, our survey included 429 articles (*Ecology*: 272, *The American Naturalist*: 77, *Journal of Animal Ecology*: 80).

We categorized studies as laboratory, field cage, field manipulation, observation/comparison, and theory. We defined as laboratory studies any experiment in which the environment was controlled (which included greenhouse as well as more obvious laboratory studies). Field cage studies included any experiment in which movement of species into or out of the study unit was barred, such as field enclosure and exclosure experiments, and “bag” experiments in lakes. Field manipulation studies included any experimental manipulations in the field that did not involve barriers to movement. Observation/comparison was a broad category for studies that did not involve any type of experimental manipulation. Finally, theoretical studies presented no original empirical data, although they may have included data from previous papers. Theoretical studies were not necessarily mathematical, and could have included verbal models or philosophical discussions of ecological problems. The majority of articles fit into only one category, although we placed 50 studies into two categories. For the analyses, we treated articles that fit into two categories as two separate studies (in other words, they were counted twice).

In addition to categorizing the type of study, we recorded the number of species considered in the study. If the authors lumped species into groups (e.g., lumped species into genera), then our tabulation referred to the number of different groups. We also recorded the characteristic time scale of the study. For example, the characteristic time scale of a study lasting 1–24 h was scored as “hours,” while a study lasting between 24 h and 1 wk was scored as “days.” If there were multiple experiments reported in the same article, we used the longest experiment to categorize the study. Finally, we recorded the duration of the study relative to the generation time of the study organisms. When species had very different generation times, the time of the experiment was categorized according to the species with the shortest generation time (thereby giving the largest

<sup>1</sup> For reprints of this Special Feature, see footnote 1, p. 663.

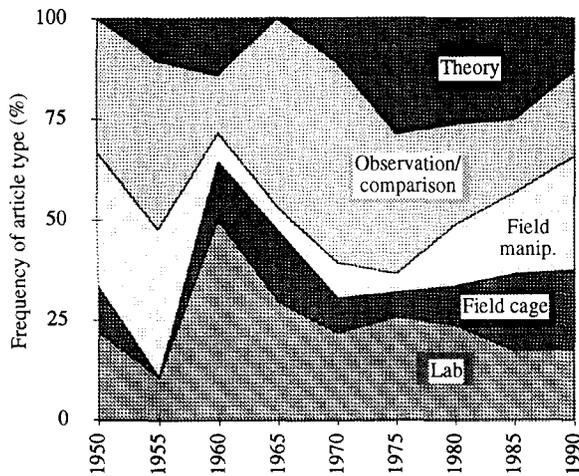


FIG. 1. The frequency of articles considering five different types of ecological studies. The articles were pooled from *Ecology*, *The American Naturalist*, and the *Journal of Animal Ecology*; we searched volumes at 5-yr intervals, 1950–1990.

number of generations in the experiment). For organisms with age-structured populations, generation time was approximated as the age of first reproduction in males. When the number of generations was not obvious, such as observation/comparison studies of static patterns, the article was not scored for duration.

#### Science Citation Index

For all articles found in the journals, we obtained the number of citations they received during 1988 and 1993 from the *Science Citation Index* (Institute for Scientific Information 1988, 1993). Number of citations was used as a measure of how influential microcosm experiments are in the current literature. *Science Citation Index* only records citations in major journals, not in books or technical reports, so we have no estimates of citations from these sources.

#### Textbooks

We chose three textbooks that are commonly used in undergraduate ecology courses: Begon et al. (1990), Ricklefs (1990), and Krebs (1994). For all three, we read the appropriate chapters to find references to articles addressing ecological interactions among two or more species. For a study to be included in the survey, the textbooks must have devoted at least four sentences or a figure/table to the article. When results presented in the textbooks were culled from a series of papers, we considered the entire set of papers a single study. Sections of the textbooks dealing with theory presented particular problems because theory was sometimes presented as common knowledge without reference to particular papers, or a large number of papers dealing with different properties of a particular mathematical model

were cited simultaneously. To be consistent with the scoring of other types of study, we grouped articles that dealt with a single theoretical idea as a single study. Grouping theoretical studies on average reduced their numbers relative to empirical studies.

#### RESULTS

##### *How common are microcosm experiments in the published literature?*

Fig. 1 shows the frequency of the different types of study over the last 40 yr. Observation/comparison and theory studies were common in the late 1960s and 1970s. However, in the 1980s field cage and field manipulation experiments increased in frequency. In 1990, studies were distributed evenly among the five types of study.

##### *Do microcosm experiments differ from other types of species-interaction studies in either the number of species included or duration?*

Fig. 2 shows the numbers of species included in the different types of study. The data are plotted as stan-

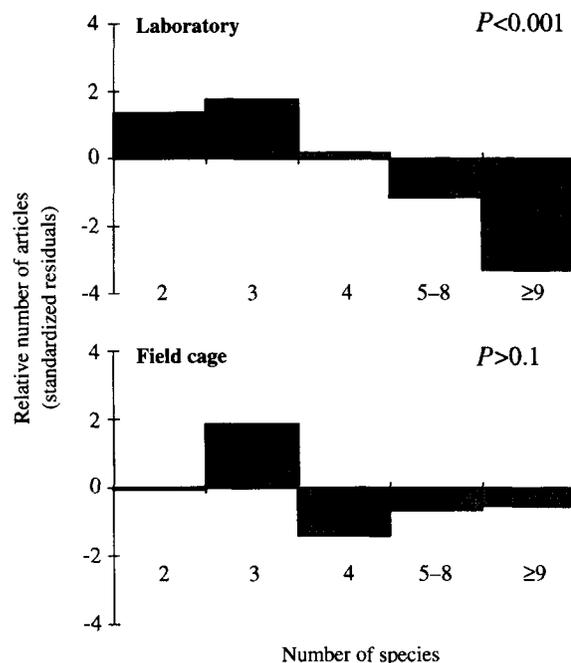


FIG. 2. The numbers of species in laboratory and field cage studies relative to other types of study. Each graph portrays a different contingency analysis for either laboratory or field cage study vs. the other types of study combined, and data are plotted as standardized residuals (Wilkinson 1988). The effect of laboratory vs. other types of study was statistically significant ( $\chi^2_4 = 21.6$ ,  $P < 0.001$ ), while the effect of field cage vs. other types of study was not ( $\chi^2_4 = 7.3$ ,  $P > 0.1$ ).

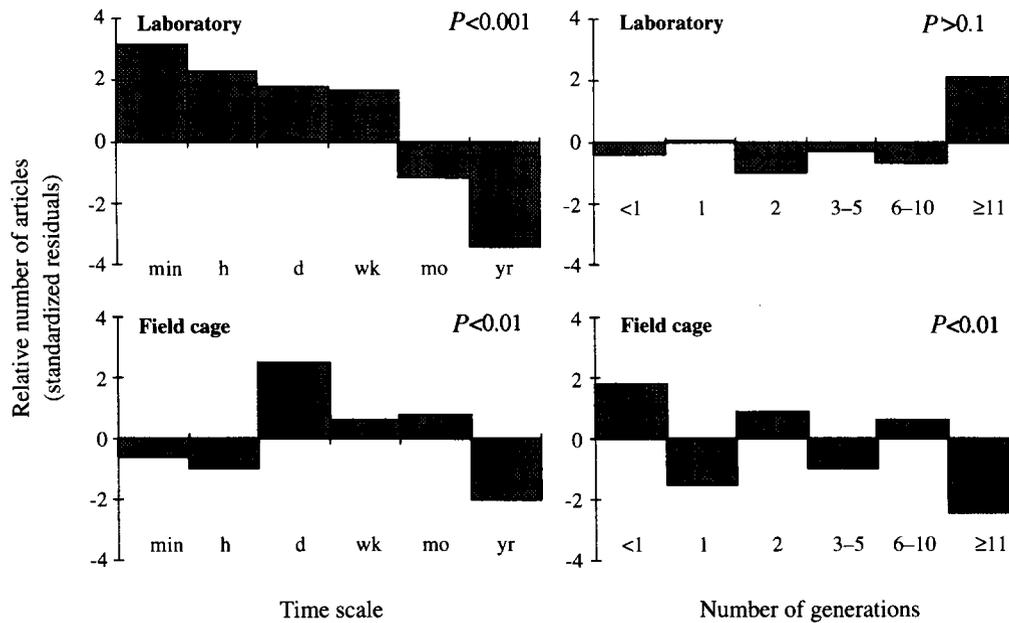


FIG. 3. The duration of laboratory and field studies relative to other types of study measured in terms of either real time or the number of generations of the study organisms. Data are given as standardized residuals for each of four separate contingency analyses. In each analysis, the specified type of study was compared to all other types of study combined. For duration measured in real time, the effects of laboratory ( $\chi^2_3 = 59.6, P < 0.001$ ) and field cage ( $\chi^2_3 = 15.6, P < 0.01$ ) were both statistically significant. For duration measured in terms of generations, field cage was statistically significant ( $\chi^2_5 = 17.2, P > 0.01$ ), while laboratory was not ( $\chi^2_5 = 8.9, P < 0.1$ ).

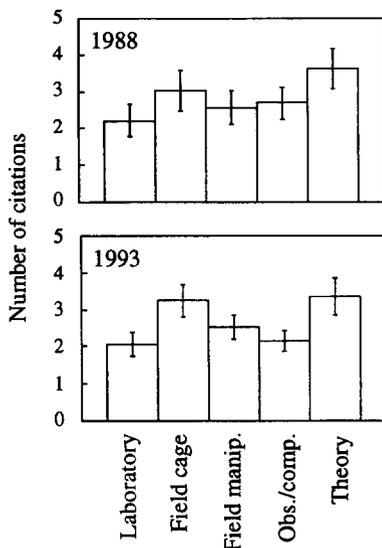


FIG. 4. The number of citations received by articles in 1988 and 1993 as a function of the type of study. An ANOVA of  $\sqrt{\text{citations}}$  against type of study showed a significant effect of type of study in 1993 ( $N = 479$  studies,  $R^2 = 0.025$ , partial  $P < 0.02$ ) but not in 1988 ( $N = 338, R^2 = 0.020$ , partial  $P > 0.1$ ). Error bars (showing  $\pm 1$  SE) are calculated from untransformed data. Obs./comp. = observation/comparison study.

standardized residuals calculated as  $(\text{observed} - \text{fitted}) / \sqrt{\text{fitted}}$ , where “observed” is the actual number of articles categorized as having 2, 3, 4, 5–8, or  $\geq 9$  species, and “fitted” is the predicted number assuming that the different types of study do not differ in the numbers of species they considered. The purpose of dividing by  $\sqrt{\text{fitted}}$  is to standardize the variances of each value. Compared to other types of study, laboratory experiments included fewer species, while the numbers of species in field cage experiments were representative of the other types of study (Fig. 2). For duration measured in real time (Fig. 3), laboratory studies were relatively short (minutes to weeks), and field cage experiments were of intermediate length (days to months). However, these patterns did not hold when time was measured relative to the number of generations of the study organisms (Fig. 3). Laboratory studies tended to include as many if not more generations than other types of study. Field cage experiments showed no clear pattern in the number of generations included.

*How influential are microcosm experiments as revealed by the number of citations they garner in the Science Citation Index?*

Fig. 4 shows the numbers of citations per article received in 1988 and 1993 for the five types of study.

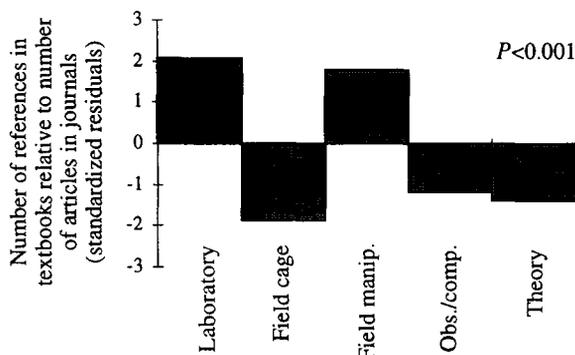


FIG. 5. Five types of ecological studies referenced in textbooks relative to journals. Differences between textbooks and journals are plotted as standardized residuals. A contingency analysis shows a significant effect of the source of the study (textbook vs. journal) and the type of study ( $\chi^2 = 19.5$ ,  $P < 0.001$ ). Obs./comp. = observation/comparison study.

In 1993, field cage and theory studies averaged more citations than other types of study ( $P < 0.02$ ), and 1988 showed a similar trend. Despite this pattern, the type of study was a poor predictor of the number of citations received per study, explaining  $<2.5\%$  of the variance in both 1988 and 1993 (ANOVA for  $\sqrt{\text{citations}}$  vs. type of study: 1988,  $R^2 = 0.025$ ; 1993,  $R^2 = 0.020$ ).

As an additional way to assess the importance of the type of study on citation rates, we compared the type of study with other variables that could explain variation in the number of citations received by an article. We performed a regression analysis with  $\sqrt{\text{citations}}$  as the dependent variable and five independent variables: journal, year of article, number of species, duration (in numbers of generations), and type of study. For both 1988 and 1993, the effects of all variables *except* type of study were statistically significant ( $P < 0.05$ ). Therefore, type of study per se was not an important contributor to the number of citations received by articles, relative to the other variables.

*How influential are microcosm experiments as measured by inclusion in undergraduate ecology textbooks?*

Fig. 5 portrays the frequency of references to the different types of study in textbooks vs. journals. Compared to journals, textbooks overrepresent the numbers of laboratory and field manipulation studies.

#### DISCUSSION

Our results show little difference between the publication rates of microcosm experiments and other types of study. Over the last 20 yr the proportion of articles reporting laboratory experiments decreased slightly, and the proportion of articles reporting field

cage experiments increased slightly; both types combined made up a constant 30–35% of the total articles on species interactions. In terms of the impact that studies have on current research as measured by citation rates, laboratory experiments received relatively fewer and field cage experiments relatively more citations than other types of study. However, the type of study explained little of the variation in citation rates and was not as important as other variables we recorded. Therefore, we conclude that the type of study plays a negligible role in determining scientific impact. The three undergraduate textbooks overrepresented laboratory studies and underrepresented field cage experiments relative to journal articles. A possible explanation for this is that simple, elegant stories are useful pedagogically, and laboratory studies produce simple stories. If this is the correct explanation, then we do not view it as a failing of the textbooks.

We also asked whether biases created by the logistics of microcosm experiments limit either the number of species studied or the length of time they are studied. Although laboratory experiments were strongly biased towards small numbers of species, we suspect that this bias will begin to erode with the growing number of large “community” laboratory experiments (e.g., Naem et al. 1994). Surprisingly, field cage experiments did not show a bias in the number of species studied. Also, although laboratory studies were shorter than other types of study when measured in real time, they did not differ in duration when measured in terms of the generations of the study organisms. This is because organisms selected for laboratory studies generally had short generation times. As with laboratory studies, field cage experiments were not longer or shorter than other types of study when measured in terms of generations. Therefore, microcosm experiments were not limited by the length of time they examined ecological interactions.

Our broad conclusion is that there are few striking differences between microcosm experiments and other types of study, either in terms of the impact they have on ecological research or in terms of biases created by logistics. There is a trend for general ecology textbooks to cite laboratory experiments more frequently than expected, but the verdict of the ecological community appears to be that, on average, microcosm experiments are neither more nor less valuable than other types of study.

#### ACKNOWLEDGMENTS

P. Kareiva and S. R. Carpenter incited this work, and C. Daehler, D. Padilla, D. Strong, M. Turner, and two anonymous reviewers provided many helpful suggestions for improving the article. V. Jagar kindly committed the data to computer. We also thank L. N. Miller for suggestions on how to categorize studies.

## LITERATURE CITED

- Begon, M., J. L. Harper, and C. R. Townsend. 1990. *Ecology: individuals, populations, and communities*. Blackwell Scientific, Oxford, England.
- Institute for Scientific Information. 1988. *Science citation index with abstracts [computer file]*—compact disc edition. Jan–Dec 1988. Institute for Scientific Information, Philadelphia, Pennsylvania, USA.
- . 1993. *Science citation index with abstracts [computer file]*—compact disc edition. Jan–Dec 1993. Institute for Scientific Information, Philadelphia, Pennsylvania, USA.
- Krebs, C. J. 1994. *Ecology: the experimental analysis of distribution and abundance*. Fourth edition. Harper & Row, New York, New York, USA.
- Naem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin. 1994. Declining biodiversity can alter the performance of ecosystems. *Nature* **368**:734–737.
- Ricklefs, R. E. 1990. *Ecology*. Third edition. W. H. Freeman, New York, New York, USA.
- Wilkinson, L. 1988. *SYSTAT: the system for statistics*. SYSTAT, Inc., Evanston, Illinois, USA.

*Ecology*, 77(3), 1996, pp. 685–690  
 © 1996 by the Ecological Society of America

## THE ROLE OF SOIL MICROCOSMS IN THE STUDY OF ECOSYSTEM PROCESSES<sup>1</sup>

HERMAN A. VERHOEF

*Department of Ecology and Ecotoxicology, Faculty of Biology, Vrije Universiteit, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands*

### INTRODUCTION

Confronted with a soil ecosystem and its complex interactions of communities of species and their associated biological, chemical, and physical processes, simplified laboratory systems have been constructed by soil scientists to bridge the gap with the complicated field situation. Collectively, these laboratory systems are called “microcosms,” and in spite of lively discussions on the reliability of these systems (e.g., Taylor and Parkinson 1988), there have been few attempts to arrive at some standardization (Van Straalen and Van Gestel 1993). However, one could wonder if it is possible, or even desirable, to strive for a universal soil microcosm methodology. The design of the experimental unit very much depends on which components of the ecosystem one wants to study: structural or functional components. For studying ecosystem structure (the composition and distribution of the biological community and the quantity and distribution of abiotic materials such as nutrients), short-term observations of simple laboratory systems are difficult to extrapolate to the field situation. However, functional components such as decomposition and nutrient mobilization have been successfully monitored in simple microcosms, and can closely follow field observations. To confirm this, microcosm studies on soil processes should include comparable studies at increasing levels of complexity and spatial scale. In this article I present some important recent results obtained from these studies and

discuss the reliability of microcosm observations for simulating and understanding natural soil processes at the ecosystem level.

### MICROCOSMS

These types of experimental systems have been shown to be very appropriate for studying the processes of litter degradation, involving measurements of CO<sub>2</sub> production, dissolved organic carbon, mineral N, P, and several cations such as K, Ca, and Mg. The microcosm design as constructed and used by Verhoef and co-workers (Fig. 1) takes the form of a cylinder with a volume of about 0.2 L. (Verhoef and De Goede 1985, Verhoef et al. 1989, Teuben and Roelofsma 1990, Verhoef and Brussaard 1990). The systems are irrigated regularly to collect leachate through the gauze bottom. This simulates the natural processes of movement of nutrients and fine humus through the soil profile in the tube. Apart from this regular collection of dissolved nutrients, extractable concentrations of nutrients in the litter are measured. CO<sub>2</sub> production is monitored by placing the microcosms in adapters through which a carbon-dioxide-free air flow is led through the soil profile. CO<sub>2</sub> concentrations are subsequently monitored in the outflow with an infrared gas analyzer. In this way all CO<sub>2</sub> produced is removed at the same rate, which is preferable to completely closed systems from which air samples are taken. Studies are performed on both coniferous and deciduous litter (Van Wensem et al. 1993). As this approach involves short-term studies (1–3 mo) and as the soil processes studied are relatively

<sup>1</sup> For reprints of this Special Feature, see footnote 1, p. 663.