The experimental paradigm and long-term population studies

CHARLES J. KREBS
Department of Zoology, University of British Columbia, Vancouver, BC, Canada V6T 2A9

Most ecologists recognize the value of long-term studies to population and community ecology, and many also subscribe to the experimental approach as the most effective way of obtaining ecological knowledge. But if we are experimentalists, do we need long-term studies? I argue that the answer to this question is *yes*, that we must combine these two approaches to solve the major ecological questions of the next century. Most of the challenging questions facing ecologists involve systems subject to long-term time trends or high environmental variability.

Because of the statistical power of many ecological methods, long-term studies are essential to measure time trends in ecosystems. Ignoring statistical power has been a major problem with short-term studies, which have predominated in the ecological literature.

Some examples of long-term studies on larch bud-moth *Zeiraphera diniana*, winter moth *Operophthera brumata* and snowshoe hares are discussed briefly to illustrate the four major desiderata of long-term projects: *spatial scale, sampling design, hypothesis testing* and *time-frame.*

Two reasons for *not* doing long-term studies are to assess density-dependence and to monitor ecosystem health. The density-dependent paradigm is bankrupt and has produced much argument and little understanding of population processes. Monitoring of populations is politically attractive but ecologically banal unless it is coupled with experimental work to understand the mechanisms behind system changes.

All ecologists favour long-term studies and in this respect differ from chemists, physicists, most other biologists, and all politicians. But, like other scientists, ecologists prefer to do experimental work. And this leads to an apparent conflict that I address in this paper: *if we are experimentalists, why do we need long-term studies?* I should like to convince you that these two goals are not in conflict, but are essential components of the 21st century.

The problems of long-term studies have been addressed in detail in the second Cary Conference in 1987 (Likens 1989) and discussed in many papers (e.g. Callahan 1984, Hinds 1984, Braham 1989). The main stimulus to this paper has been the provocative discussion by Taylor (1989) of the role of experiment in long-term research. Taylor (1989) argued that the experimental approach is antithetical to long-term studies and that the experimental philosophy can be detrimental to the future development of ecological science. I would like to argue against these pessimistic conclusions and for a broader view in which long-term studies and experimental approaches collaborate in achieving ecological understanding.

The experimental paradigm

The experimental paradigm is one of the characteristic features of science and has been discussed extensively by philosophers of science (Popper 1963, Lakatos & Musgrave 1970) and by ecologists (Strong 1980, Simberloff 1980, 1983, Hurlbert 1984, Diamond 1986, Hairston 1989). The experimental approach (also called the hypothetico-deductive method) can be summarized as six steps: (1) Find an interesting system; (2) Frame a question of a hypothesis about this system; (3) Identify your assumptions; (4) Make a set of predictions based on these assumptions; (5) Manipulate or observe your population or community; (6) Test your hypothesis by comparing your results with the predictions made previously. This procedure recycles continuously as we improve our understanding and re-define our assumptions and predictions.

There are a few minor comments I would like to add to this description since ecologists often misunderstand some features of the experimental method. No philosophical discussion is required for the first two steps listed above. Philosophers of science cannot tell you what an interesting ecological system is, nor what questions about it are interesting. These judgments are often clear only in retrospect. In highly developed areas of science, theory can dictate what questions are most critical to answer, but this is rarely the case in ecology.

The next steps (3 and 4) are primarily methodological ones and mathematical models may be useful at this stage. Ecologists often fail to recognize all the assumptions their theories make, and this has caused much altercation in the literature. To make a set of predictions from a hypothesis and a series of assumptions is partly a logical process compounded with scientific creativity.

An *experiment* is any attempt to test a hypothesis and Hurlbert (1984) clearly pointed out that experiments could be *mensurative* or *manipulative*. Thus the ability to manipulate ecological systems is not an essential feature of the experimental approach. This critical step is very difficult in ecology for methodological reasons—population and community parameters are difficult to estimate—and is discussed in detail elsewhere (Krebs 1989). To test the predictions of any hypothesis requires unbiased data collected with an appropriate experimental design and proper analysis. Although I gloss over these two steps here, they are the Achilles heel of modern ecology.

Within the experimental framework, ecologists take two different approaches to their subject. Some ecologists are system-based, relying on a single species or community for their studies. Others are concept-based and utilize whatever species or populations that are convenient for answering the question. This conference is largely system-based, and all long-term studies are system-based. There is no reason why these two approaches to ecology should generate any conflict, although some individuals seem to think concept-based science is superior to system-based science. It is superior in the sense that the circumstances are more varied. Every system-based theory should be concept-tested if it is to be of general significance.

The problem

The experimental approach is usually identified with short-term studies. There is no necessary reason for this, and Hairston (1989, p. 31) recognizes that experiments can have any duration, depending on the organism being studied. In a survey of papers published in *Ecology* from 1977 to 1987, Tilman (1989) found that only 1.7% of 749 papers reported field experiments lasting at least 5 years, and that 86% of the manipulative experiments lasted 1–3 years. These constraints are partly determined by the duration of postgraduate studies and the length of funding of research grants.

It is inevitable that ecologists will address first short-term ecological questions, and I would argue that we have by now completed the major analyses of short-term problems in ecology (Begon *et al.* 1986). What has been neglected are the long-term ecological processes. An example is the Russell Cycle in the English Channel (Southward 1980) which occurs over a 50-year span. How can we study such long-term processes?

First, we must recognize that we need descriptive studies even to *recognize* long-term processes. Ecosystems with pronounced time-trends will probably be the rule in the future, and our equilibrium-based ecological theory may not be a useful basis for understanding. We can get these descriptive data in two ways: (1) by establishing monitoring programmes to gather the key data; or (2) by establishing research teams to analyse the system experimentally. In the second case the key data are gathered on control areas, and these control data serve as a monitoring programme as well.

By combining the experimental approach with long-term studies we can achieve all the desired goals discussed by Likens (1989). Slow ecological processes like succession can be monitored and described at the same time that ecological theories about the causes and consequences of succession are tested experimentally. Long-lived species can be used as biomonitors of environmental health and simultaneously studied to test hypotheses of reproductive effort or interspecific competition.

I therefore reject the belief that if we are experimentalists we do not need long-term studies. Most of the challenging questions of modern ecology—habitat fragmentation, effects of climate change—involve systems subject to long-term trends or high levels of environmental variability that cannot be reduced to the standard 3-year study.

I also reject Taylor's (1989) view that long-term projects should be viewed as different from experimental studies. It is possible to set up long-term projects, as Taylor (1989) clearly describes, and he records historically how some of these, like the studies of Great Tits *Parus major*, became a focus for later experimental work. But I do not think this is a desirable path for ecologists to pursue. Long-term projects could be set up on a variety of ecological communities or populations, and in an ideal world we could do this and see at a later date how useful each of these studies had been to advancing ecological understanding. But this would be a waste of scarce resources. I would argue that Taylor's (1989) strategy is a good first-generation approach but that we now know the outlines of ecological theory and the central principles of importance (Cherrett 1989). Our second-generation strategy ought to combine the best of the long-term orientation with the experimental approach.

The question of statistical power

Additional support for long-term studies is now appearing from an unexpected quarter—statisticians. Ecologists have in their haste to test hypotheses forgotten about power (Gerrodette 1987, Peterman 1989). Statistical power is a measure of the probability of rejecting the null hypothesis in favour of an alternative hypothesis that is correct. In ecological studies the simplest problem is that of detecting trends in population size. Suppose that Kittiwake *Rissa tridactyla* numbers are decreasing at 5% per year. How long would you have to census this population before you could reject the null hypothesis that numbers were not changing over time? Clearly the answer to questions of this type depends on the precision of your population estimate and the magnitude of the decline. Gerrodette (1987) discusses this statistical problem in detail and shows some examples in which 10–20 years of data are required to recognize a significant time trend. Peterman (1989) gives some even more striking examples in which, for example, it may take 30 years to recognize a 50% reduction in whale abundance with conventional abundance estimates.

Ignoring statistical power has profound consequences for ecological understanding. In many cases ecological data are used to test a null hypothesis of no effect and the null hypothesis is not rejected. If variability is high or sample sizes are low, the power of this test may be very low so that Type II errors (accepting the null hypothesis when in fact it is false) occur. If this happens several times, the null hypothesis becomes an ecological generalization without a proper test. The resolution of this problem is simple—report the power of the test for a specified alternative hypothesis (Peterman 1989). A beneficial side-effect of this approach is that interest is focused on specified alternative hypotheses rather than on the null hypothesis alone (cf. Platt 1964).

Some examples of long-term studies

I would like to conclude by discussing a few examples of long-term studies that cover non-avian taxonomic groups in order to discuss some of the problems faced by these research programmes and to examine the criteria by which we can judge long-term studies.

Larch bud-moth in Switzerland

The larch bud-moth *Zeiraphera diniana* is a defoliating tortricid moth that has been periodically attacking large areas of subalpine larch–cembran pine forests in the European Alps for over 200 years. The population dynamics of this forest insect pest have been studied for 34 years (1949–1983) and the results are summarized in Baltensweiler & Fischlin (1990). The most regular outbreaks of the larch bud-moth occur in the Engadine Valley of southeastern Switzerland, where 16 outbreaks have been recorded in detail since 1855, and the outbreak cycles have a period of 8.5 ± 0.3 (SE) years.

I do not wish to discuss the biology of the larch bud-moth system in detail here. Briefly, population fluctuations in the larch bud-moth are believed to be driven by an interaction between the food plants (trees) and the moth larvae. Defoliation causes the larch trees to produce short needles the following year with high fibre content and larvae feeding on these needles grow poorly and reproduce less well. Most of the cyclic dynamics of this system seem to be explained by this plant–herbivore hypothesis. A colour polymorphism occurs in larch bud-moth, and the frequency of these genotypes changes systematically with density. It is not known whether this genetic polymorphism is a necessary component of the cycles or merely a side-effect of the basic plant–herbivore interactions.

Which features of this 34-year study were well-planned and which were not? I ask this question not as a criticism of these studies but as one way of guiding our thinking about current and future long-term studies.

(i) *Spatial scale:* for 34 years the Engadine Valley was the major focus, but four additional outbreak areas were sampled for 21 years. This study provided very good spatial coverage.

(ii) *Sampling programme:* the population census technique was clearly worked out at the beginning of the study and carried out in a consistent manner throughout the 34 years.

(iii) *Hypothesis testing:* at the end of the first ten years of study an epizootic granulosis disease was recognized as the major cause of the decline and, if the study had been terminated at that time, disease would be thought to be the major cause of these cycles. But granulosis disease has never been seen again in 11 more cycles followed on different areas and in different years!

In general, hypothesis testing in the larch bud-moth studies has been based on mensurative experiments. No manipulative experiments have been carried out. If there is a general criticism of the larch bud moth research programme, it is that a clear set of alternative hypotheses was not identified and analysed as a set. The approach used was more inductive, or *a posteriori*, rather than deductive.

(iv) *Time-frame:* most of the major studies on the larch bud-moth have now ceased, since one conclusion of this research was that the bud-moth is not a major destructive insect pest. The 34 years of this study have been sufficient to provide a general model for the population fluctuations, but not long enough to test all the interactive hypotheses such as the food quality-genetic polymorphism model or the assortative mating-genetic polymorphism model (Baltensweiler & Fischlin 1990).

Winter moth

The winter moth *Operophthera brumata* has been the subject of a classic series of population studies begun by G. C. Varley in 1949 at Oxford. Nineteen years of data on the Oxford population have formed the basis of many models of insect population regulation (Varley & Gradwell 1968, Hassell 1980, den Boer 1988). The winter moth story is unusual because this moth is a pest species, and subsequent introductions to Nova Scotia (Embree 1966) and British Columbia (Roland 1986) have allowed a test of our understanding of population processes in this insect.

The original conclusions of Varley & Gradwell (1973) were that winter moth numbers in England were regulated by pupal predation in the soil during winter. Parasites of the winter moth were not density-dependent and, if they were perceived to have any effect, it was to cause instability in winter moth numbers. But when the parasite *Cyzenis albicans* was introduced into Nova Scotia for biological control in 1954 and into British Columbia in 1980, the numbers of winter moth decreased dramatically and these cases have become a classic example of biological control. The outcome of the biological control was not predicted by the English data on the winter moth.

Which features of the winter moth study were well-planned and which were not?

(i) *Spatial scale*: the original Oxford study was limited to five oak trees in Wytham Woods, and these were studied for 19 years. This is a classic example of pseudoreplication at only one study site. This research programme was deficient in not considering spatial scale.

(ii) *Sampling programme:* the sampling design was clear and comprehensive throughout the study. In this study, as in any other, the choice of what long-term variables to measure is critical to success in understanding mechanisms of population changes.

(iii) *Hypothesis testing:* the Oxford study was oriented toward life tables and *K*-factor analysis (Varley & Gradwell 1968) and thus limited to mensurative experiments. No manipulative experiments of even the simplest kind were carried out (Roland 1986). There was no clear statement of alternative hypotheses. Perhaps it is significant that this observational approach has led to continuous controversy 20 years after the study was ended (Den Boer 1988).

(iv) *Time-frame:* winter moth numbers at Oxford fluctuated in a quasi-cyclic manner with highs about 8 years apart. If these fluctuations are common, 19 years is too short a time-frame for studying these populations (cf. larch budmoth).

The original winter moth studies are deficient therefore in both spatial scale and in hypothesis-testing and duration of study. This is not a good example of a long-term study.

Snowshoe hare

In the boreal forests of Canada snowshoe hares *Lepus americanus* fluctuate cyclically with a period of 9–10 years (Keith 1963). These cycles are recorded qualitatively in the fur trading statistics of the Hudson Bay Company since 1672 and have been analysed extensively by mathematicians (Finerty 1980). There is an abundance of general natural-history information on these cycles throughout Canada.

Three attempts have been made to study the causation of snowshoe hare cycles. Green & Evans (1940) studied a population in central Minnesota from 1932 to 1939 and came to the conclusion that a metabolic disorder, 'shock disease', was a major cause of the decline. Chitty (1959) showed that this conclusion was incorrect. The main studies on snowshoe hares were done by Lloyd Keith and his associates at Rochester, Alberta, from 1962 to 1977. We have also been studying snowshoe hare cycles in the southwestern Yukon since 1976.

The dominant theme that has emerged from the studies of Keith and co-workers is that a combination of winter food shortage and predation acts in concert to drive the snowshoe hare cycle (Keith *et al.* 1984). How well have these studies addressed the cycles question?

(i) *Spatial scale:* the spatial scale of both the Green & Evans (1940) and the Keith (1983) studies has been large. Because hares cycle in synchrony over large regions, the spatial scale is less significant. What is needed is replication in different regions of the boreal forest. There is little to fault here.

(ii) *Sampling programme:* the methods used to estimate numbers and densities have been ad-hoc in many of Keith's studies. One could argue that the changes are so large that even poor methods will pick them up. But, judging by today's criteria, the hare studies have been methodologically relatively poor for estimation of density and survival rates. Other sampling techniques used (e.g. reproductive estimates) are more standard and there is little problem (Cary & Keith 1979).

(iii) *Hypothesis testing:* the Green & Evans (1940) research and the Keith research programme were not primarily hypothesis-driven, and it is still not possible to find a written statement of what alternative hypotheses are being considered or what predictions follow from the original hypothesis (Keith *et al.* 1984). The approach has been mainly (but not entirely) inductive and can be described as mensurative. We have tried in our Yukon project to state and test explicit hypotheses about the mechanism of the hare cycle (Krebs *et al.* 1985, 1986, Smith *et al.* 1988, Sinclair *et al.* 1988). It is certainly correct, as Taylor (1989) points out, that we need preliminary work before we can frame useful hypotheses, but for hare cycles this point was reached 30 years ago (Keith 1963).

(iv) *Time-frame:* the studies by Keith and co-workers are among the longest yet done on a mammalian population, yet they are by any standard far too short to be complete. This is not a criticism of Keith but rather to point out that such studies must be set in a long-term framework that recognizes $n = 1$ means 10 years.

I think we should consider all long-term studies within a general framework of the sort I have used here in order to specify and rectify if necessary weak points.

Why not to do a long-term study?

There are at least two reasons I can think of that should *not* be the sole justification of a long-term study. The first is to assess density-dependence in a population. Long-term data sets can certainly be used in a *K*-factor analysis to measure density-dependence. That this approach has been a bankrupt paradigm should be apparent to all who have followed the endless controversies over population regulation (den Boer & Reddingius 1989). The winter moth study discussed above illustrates the problem all too well. The density-dependent paradigm is bankrupt because it is descriptive and *a posteriori*. It does not lead to understanding because no mechanisms are specified. It is an equilibrium-based concept that sheds no light on a non-equilibrial world. It assumes that the causes of death can be measured uniquely and that these operate independently and additively. Finally, statistical tests for density-dependent regulation are not possible in most species (Reddingius & Den Boer 1989).

Many ecologists seem wedded to the density-dependent paradigm because they see no alternative. I have proposed the experimental paradigm as one possible alternative in which interest is focused on limiting factors and the testing of specific mechanisms of regulation (Krebs 1978, 1985, p. 341, Watson & Moss 1970).

A second reason not to do long-term studies is in order to monitor ecosystem health. This is a politically attractive way for ecologists to convince governments to spend money on long-term projects, but it achieves this goal by prostituting the science to a trivial activity. Monitoring is essential to long-term studies because it is the 'control' treatment, but it must be coupled with experimental testing to provide useful scientific understanding. If we were to be provided with a list of population sizes for all species on earth on 1 January each year, we would not have increased ecological knowledge one iota. Only if we understand why populations increase to become pests, remain fairly constant, or decline to extinction will we increase our knowledge of population dynamics.

I realize that in arguing against simple monitoring programmes, I am running against the ecological tide (Likens 1983, Hinds 1984). My feeling is that we sell our science short if we agree with the politicians that our job is only to monitor ecological change. We need the ecological understanding of these changes so that effective management actions can be recommended with confidence. We have much to do and far to go.

REFERENCES

Baltensweiler, W. & Fischlin, A. 1990. The Larch Budmoth in the Alps. *In* Berryman, A. (ed.) Dynamics of Forest Insect Populations: 331–351. New York: Plenum Publishing Corp.

Begon, M., Harper, J.L. & Townsend, C.R. 1986. Ecology. Individuals, Populations, and Communities. Sunderland, Mass.: Sinauer Associates.

Braham, H.W. 1989. Long-term research at the National Marine Mammal Laboratory. Bull. Ecol. Soc. Am. 70: 21–25.

Callahan, J.T. 1984. Long-term ecological research. Bioscience 34: 363–367.

Cary, J.R. & Keith, L.B. 1979. Reproductive change in the 10-year cycle of snowshoe hares. Can. J. Zool. 57: 375–390.

Cherrett, J.M. (ed.) 1989. Ecological Concepts. Oxford: Blackwell Scientific Publications.

Chitty, D. 1969. A note on shock disease. Ecology 40: 728–731.

Den Boer, P.J. 1988. Density dependence and the stabilization of animal numbers. 3. The winter moth reconsidered. Oecologia 75: 161–168.

Den Boer, P.J. & Reddingius, J. 1989. On the stabilization of animal numbers. Problems of testing. 2. Confrontation with data from the field. Oecologia 79: 143–149.

Diamond, J. 1986. Overview: laboratory experiments, field experiments, and natural experiments. *In* Diamond, J. and Case, T.J.(eds) Community Ecology: 3–22. New York: Harper and Row.

Embree, D.G. 1966. The role of introduced parasites in the control of the winter moth in Nova Scotia. Can. Entomol. 98: 1159–1168.

Finerty, J.P. 1980. The Population Ecology of Cycles in Small Mammals. New Haven: Yale University Press.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology 68: 1364–1372.

Green, R.G. & Evans, C.A. 1940. Studies on a population cycle of snowshoe hares on the Lake Alexander area. J. Wildl. Mgmt. 4: 347–358.

Hairston, N.G. 1989. Ecological Experiments. Purpose, Design, and Execution. Cambridge: Cambridge University Press.

Hassell, M.P. 1980. Foraging strategies, population models and biological control: a case study. J. Anim. Ecol. 49: 603–628.

Hinds, W.T. 1984. Towards monitoring of long-term trends in terrestrial ecosystems. Environ. Conserv. 11: 11–18.

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

Keith, L.B. 1963. Wildlife's Ten-year Cycle. Madison: University of Wisconsin Press.

Keith, L.B. 1983. Role of food in hare population cycles. Oikos 40: 385–395.

Keith, L.B., Cary, J.R., Rongstad, O.J. & Brittingham, M.C. 1984. Demography and ecology of a declining snowshoe hare population. Wildl. Monogr. 90: 1–43.

Krebs, C.J. 1978. A review of Chitty's hypothesis of population regulation. Can. J. Zool. 56: 2463–2480.

Krebs, C.J. 1985. Ecology: the experimental analysis of distribution and abundance, 3rd edn. New York: Harper and Row.

Krebs, C.J. 1988. The experimental approach to rodent population dynamics. Oikos 52: 143–149.

Krebs, C.J. 1989. Ecological Methodology. New York: Harper and Row.

Krebs, C.J., Boutin, S. & Gilbert, B.S. 1985. A natural feeding experiment on a declining snowshoe hare population. Oecologia 70: 194–197.

Krebs, C.J., Gilbert, B.S., Boutin, S., Sinclair, A.R.E. & Smith, J.N.M. 1986. Population biology of snowshoe hares. I. Demography of food-supplemented populations in the southern Yukon, 1976–84. J. Anim. Ecol. 55: 963–982.

Lakatos, I. & Musgrave, A. 1970. Criticism and the Growth of Knowledge. London: Cambridge University Press.

Likens, G.E. 1983. A priority for ecological research. Bull. Ecol. Soc. Am. 64: 234–243.

Likens, G.E. 1989. Long-term Studies in Ecology: approaches and alternatives. New York: Springer-Verlag.

Peterman, R.M. 1989. Statistical power analysis can improve fisheries research and management. Can. J. Fish. Aquatic Sci. 47: 2–15.

Platt, J.R. 1964. Strong inference. Science 146: 347–353.

Popper, K.R. 1963. Conjectures and Refutations: the growth of scientific knowledge. London: Routledge and Kegan Paul.

Reddingius, J. & Den Boer, P.J. 1989. On the stabilization of animal numbers. Problems of testing. 1. Power estimates and estimation errors. Oecologia 78: 1–8.

Roland, J. 1986. Success and failure of *Cyzenis albicans* in controlling its host the winter moth. PhD thesis, University of British Columbia.

Simberloff, D. 1980. A succession of paradigms in ecology: essentialism to materialism and probabilism. Synthese 43: 3–39.

Simberloff, D. 1983. Competition theory, hypothesis testing, and other community ecological buzzwords. Am. Nat. 122: 626–635.

Sinclair, A.R.E., Krebs, C.J., Smith, J.N.M. & Boutin, S. 1988. Population biology of snowshoe hares. III. Nutrition, plant secondary compounds and food limitation. J. Anim. Ecol. 57: 787–806.

Smith, J.N.M., Krebs, C.J., Sinclair, A.R.E. & Boonstra, R. 1988.

Population biology of snowshoe hares. II. Interactions with winter food plants. J. Anim. Ecol. 57: 269–286.
Southward, A.J. 1980. The western English channel—an inconstant ecosystem? Nature 285: 361–366.
Strong, D.R. 1980. Null hypotheses in ecology. Synthese 43: 271–285.
Taylor, L.R. 1989. Objective and experiment in long-term research. *In* Likens, G.E. (ed.), Long-Term Studies in Ecology: 20–70. New York: Springer-Verlag.
Tilman, D. 1989. Ecological experimentation; strengths and conceptual problems. *In* Likens, G.E. (ed.), Long-Term Studies in Ecology: 136–157. New York: Springer-Verlag.
Varley, G.C. & Gradwell, G.R. 1968. Population models for the winter moth. *In* Southwood, T.R.E. (ed.), Insect Abundance: 132–142. Oxford: Blackwell Scientific Publications.
Varley, G.C., Gradwell, G.R. & Hassell, M.P. 1973. Insect Population Ecology: an analytical approach. Oxford: Blackwell Scientific Publications.
Watson, A. & Moss, R. 1970. Dominance, spacing behavior and aggression in relation to population limitation in vertebrates. *In* Watson, A. (ed.), Animal Populations in Relation to Their Food Resources: 167–220. Oxford: Blackwell Scientific Publications.