

THE LOGIC OF ECOLOGICAL EXPERIMENTS: A CASE HISTORY FROM STUDIES OF THE DISTRIBUTION OF MACRO-ALGAE ON ROCKY INTERTIDAL SHORES

A.J. UNDERWOOD

Institute of Marine Ecology, Zoology Building, University of Sydney, NSW 2006, Australia

The various steps involved in experimental investigations of the causes of observed ecological patterns are examined. A sequence of different steps (observation, model, hypothesis, null hypothesis, test) is identifiable. The logical necessity for distinguishing and evaluating each of these steps is discussed. The potential usefulness of explicit recognition of the logical relationships between the various parts of the procedure is evaluated by reference to experimental studies to determine the causes of boundaries of vertical distribution of macro-algae in rocky intertidal habitats. Pitfalls of experimental procedures can sometimes be recognized as failures in the logical progress from one part to the next in the sequence of steps necessary in any study. Particularly important is the early recognition of the existence of many, potentially conflicting, but sometimes interactive, models to account for any observed pattern. Greater attention to the relationships between the components of the logical structure of an experimental study might aid in reduction of errors in the design, analysis and interpretation of ecological experimental data.

INTRODUCTION

Ecologists generally, and marine ecologists in particular, do not spend sufficient time, at least according to the available literature, in introspection about the nature of the science that they do. There have been several recent contributions in an attempt to understand the history of ecology (Jackson, 1981; McIntosh, 1985; Worster, 1977) and some that have been devoted to understanding the logical basis of ecological (Connor & Simberloff, 1986; Fagerstrom, 1987; Loehle, 1987; Mentis, 1988) and other experimentation (in particular, Hacking, 1983). In general, however, there are too many problems in the way that marine ecologists function for complacency about our understanding of the difficulties of the problems we are trying to solve (*e.g.* Underwood, 1981a, 1986; Hurlbert, 1984; Underwood & Denley, 1984). Appeals to logicians or philosophers from outside the discipline will not necessarily be of any help (Edson *et al.*, 1981; Roughgarden, 1983). Sometimes, there is evidence to suggest that such people will have a lot to offer (Simberloff, 1980; Loehle, 1987). In other cases, however, it is clear that most of the philosophers of science have spent too much time on the easy sciences like physics and have never had to grapple with the difficulties of field experimentation in complex and variable worlds (Platt, 1964; Underwood, 1990).

This sort of introspection should be helpful, if only to try to avoid some of the worst mistakes that seem to be made routinely in ecological field experiments (Hurlbert, 1984; Underwood, 1986). It is the intention of the present contribution to consider some of the logical bases of experimentation in a restricted area of marine ecology. This will serve as a case study in order to determine whether or not procedures that are available will turn out to be useful as aids to thought in future studies. Much of the groundwork of philosophy and logic that underlies the way we do experiments has recently been reviewed (Underwood, 1990). In addition, much of the material that forms the examples has been reviewed in its ecological context (Lubchenco & Gaines, 1981; Hawkins & Hartnoll, 1983; Underwood, 1985). Here, these two notions are brought together in an attempt to use the review of the latter to understand the potential usefulness that has emerged from the review of the former.

The patterns of distribution of conspicuous macroscopic algae on rocky shores have been investigated in a number of parts of the world. In several studies, particularly recently, there have been experimental approaches to try to investigate the processes causing the limitations of vertical distribution of these organisms. Parts of this study have a respectably long history (Hatton, 1932, 1938). Nevertheless, as is common with much of modern ecology, the use of controlled, manipulative field experiments is relatively new. There is, indeed, still debate about whether experiments are necessary to examine any of the propositions that ecologists regularly use (Diamond, 1986). Here, it is argued that experiments are, in fact, always necessary. The problem is to do experiments that are logically constructed and soundly executed. Some experiments on the upper and lower limits of distribution of macro-algae are considered as a set of examples that will indicate the main features of the problems, some of the solutions to the problems and the general framework of logic in which experimental studies of this type are done. Much of the material comes from work in New South Wales simply because of the great familiarity I have with those studies and is not meant to suggest that these are somehow superior to those in other parts of the world. This is not intended as a complete review but is an attempt to identify the steps necessary to complete a study of this sort and to determine whether the logical procedure outlined in the previous review (Underwood, 1990) is actually effective when it is used in a specific example. The emphasis is on an explicit recognition of the parts of the process of planning the study, doing it in the field and interpreting its results.

THE LOGIC OF REFUTATION

As with everything else in ecology, there can be considerable disagreement about the sort of evidence that is required to support or refute models or theories for the distribution and abundance of populations, or the structure of assemblages in nature. There is even disagreement about whether these theories and ideas should actually be put to the test or should simply be accepted. Here, the tradition of refutationist logic most closely associated with the work of Popper (1968, 1969) is advocated. This is necessary because, to explain any particular series of observations that one has made, there will usually be more than one model for possible explanation (see below). Thus, some

procedure is always necessary to allow distinction between various, often contradictory, alternative explanations. This, I believe, makes it mandatory that the models or ideas are always subjected to test. Only where some of the alternatives (hopefully as many as have been currently thought of) have been subjected to some critical examination and have failed, and are therefore rejected, will the remaining one (or small number) be acceptable as a valid explanation. The whole notion is that the current explanations are the survivors that have not yet failed to pass a critical evaluation. They are not necessarily true, but they are currently the best and their embellishment and their challenge by new ideas and by new models is the constant process of trying to increase our understanding of the natural world.

Many authors have made it abundantly clear why it is that proof is not a logical proposition. Therefore disproof becomes the focus for attention in any attempt to evaluate ecological models or theories. The argument is very simple; it reflects the fact that to prove something requires that all possible observations about it are made. A well-known example is the proposition that all swans are white. Billions of observations of swans throughout Europe indicated that this was a generality, it could even be called a law of nature. Millions more corroboratory observations of white swans still have not made this particular proposition true. Only when all swans, everywhere, have actually been examined, is it possible to know that all current swans are white (Chalmers, 1979). The single (first) observation of a black swan in Australia certainly demonstrated that the inductive attempt to prove a law of nature by repeatedly gathering confirmatory evidence was futile. The general problem of induction is well exemplified here. No amount of confirmation will ever establish that something is true (despite analyses to the contrary by Carnap, 1962, and Hempel, 1965). Other examples of this sort of process exist and have been widely discussed (Hume 1779; Whewell, 1847; Russell, 1912; Popper, 1968).

Of course this over-simplifies everything and does not make a very good case for the complexity and variability of ecological cases. Nevertheless, these can also be shown to have the same problems associated with them in any attempt to prove something. Thus, disproof becomes the best procedure for distinguishing between alternative explanations for some phenomena. Here, refutationist tests are the only ones of interest.

A SCHEME FOR THE LOGICAL COMPONENTS OF A REFUTATIONIST TEST

One scheme which seems to fill the necessary roles of different phases of a study, both in a historical context of trying to understand why certain investigations have been successful and in the prescriptive sense of trying to understand what might be a useful scheme for future studies, has been distilled from the literature in a previous review (Underwood, 1990). It is briefly summarised here, without justification for identifying the separate steps. Full details of that justification are available in the cited review.

There are several steps that should be clearly distinguished in any study of an ecological phenomenon. Almost invariably such studies or research programmes (Lakatos, 1974) start with a series of observations, or recognition of some pattern in the natural world which requires explanation. This is the first step identified in Figure 1. Observa-

tions, or patterns, or whatever terminology is appropriate, are easy to make. Their explanation is the purpose of much modern ecology. The most convincing explanations are probably also the most useful in terms of the wider, more applied aspects of ecology for conservation, management of resources, detection of environmental perturbations and subsequent repair of habitats. As will be seen later, an objection to the use of observations as the beginning of a study can be overcome by use of falsification procedures. It is not necessary to become bogged down in disagreements about the 'theory-ladenness' of observations (Hanson, 1959; Koyre, 1968; Feyerabend, 1975). This problem can simply be dealt with by application of the same logic as required in the main study itself.

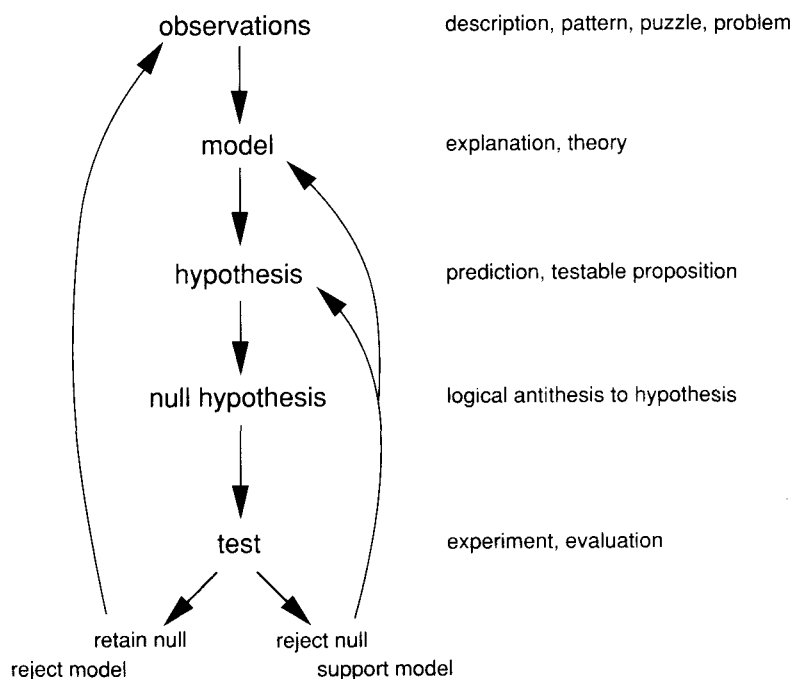


Figure 1. The sequential stages of logical construction leading to an experimental test (modified after Underwood, 1990; see text for details).

The second step, therefore, is to provide some explanation for the observations. This is hereafter referred to as a model. The only rules about a model are that it should be a plausible explanation of the original observation (Nagel, 1961). Other words that have been used for this step in the process are (see Trusted, 1979) explanations, theories, laws and sometimes hypotheses - but that is reserved for something else here.

There are almost invariably competing models that could explain the given observations (Chamberlin, 1965, reprint of 1880 paper). It is therefore necessary to have some procedure for distinguishing between them. Some will clearly be incorrect if subjected to an appropriate test. Some may be incorrect, but currently possible tests will fail to reject them. One or other explanation may, in fact, be true - but that is not necessarily the end-point of an ecological study. The chore is to distinguish between models, eliminating

those that are patently false. This will leave a sub-set (ideally only one if sufficient tests have been completed) that has, to date, survived. Complacency must not then set in, however, because this model may also be wrong (but not yet sufficiently tested) or new models may be proposed in the future (Popper, 1969). Thus, any existing model needs constant probing and re-evaluation in order to determine whether it is still doing well. It is precisely this procedure that eventually caused Newtonian mechanics to cease to be considered a correct explanation for the patterns of motion (even though they had apparently been correct for several hundred years before they failed to pass certain critical tests (Chalmers, 1979)).

The procedure advocated here is to use each model as a starting basis for construction of a prediction, hereafter called an hypothesis (Figure 1). The rules for construction of these are very simple. It matters not what form it takes, as long as the prediction is about some observation not yet made (and usually therefore under some circumstances that have not yet been found or created). The general form of an hypothesis is: "if the model is correct then under the following circumstances we can predict ..." (see particularly Hempel, 1966). There are problems inherent in the construction of predictive hypotheses from a model and it is always necessary to put a great deal of thought into this phase of the procedure. It is always potentially possible to make a prediction that would also be made by another, as yet unthought-of, model (see Underwood, 1990 for examples). Where alternative models are already articulated (see below), it is a useful procedure to try to identify hypotheses from each of the models such that the hypotheses are clearly contrasting. It would be most useful in these ecological experimental studies to be able to focus attention on the hypotheses that predict the outcomes which are most different under some circumstances and to use these in tests in an attempt to eliminate competing models as explanations for some observations. In this, there is a departure from the classic procedure outlined originally in its formal sense by Popper (1969). He advocated that the most useful hypotheses are those that are most easily refuted (if wrong) by tests. In his case, such hypotheses came from models with the most information content. Here, it is argued that models can best be distinguished by using those hypotheses which are most clearly distinctive. Thus, it may not be necessary to make the most 'bold conjectures' (Popper, 1969). It may be more relevant to identify those hypotheses which predict quite different outcomes under the same experimental circumstances.

Having arrived at an hypothesis, ideally one that is logistically feasible to examine and is in appropriate spatial and time scales (see later), it is necessary to subject it to test. An experiment is therefore the creation of the circumstances required in the hypothesis, *i.e.* the circumstances under which some prediction has been made. At this point, the use of refutation as opposed to proof becomes important. It is not now possible to examine the hypothesis in a test which attempts to prove that the hypothesis has come true once the relevant circumstances have been created. It is much more useful and practical to be able to disprove something (see above). Thus, instead of using the test to examine the hypothesis itself, an alternative, null hypothesis is created which consists of all logical alternatives to the hypothesis (Figure 1). Specific examples of this will be demonstrated below. In general, what is needed is a clear statement that, if the hypothesis is a prediction that under circumstances (a), event (b) will occur, then a null hypothesis will need to be

that under those circumstances anything other than (b) will occur. This is much simpler to do in practice. For example, if there is an hypothesis (based on some model that I have not identified here) that removal of predators will cause an increase in the numbers of some species of barnacle, in contrast to areas where predators have not been removed, then the appropriate null hypothesis is that removal of predators will cause no increase, or may be associated with a decrease in the number of barnacles. If this null hypothesis can be refuted in an experimental removal of predators, with appropriate controls, then clearly the only remaining possibility is the hypothesis. This well-known piece of jiggery-pokery by-passes the whole problem of attempting to prove things. Disproof of the alternatives is a simpler operation (and logically sound - see Lemmon, 1971; Hocutt, 1979).

Finally, at the end of a whole long sequence of steps, comes the experimental test. An experiment (or test) is simply an evaluation of the circumstances arising once the conditions required by the hypothesis have been fulfilled. In the previous example, the test is to remove the predators from certain areas, leaving untouched areas with natural predation going on in them as controls. If, after an appropriately defined period of time (this may itself be a problem), the number of barnacles in the two types of area is examined, the test is complete. If there are more barnacles in the controls than in the experimentals, this is not consistent with the null hypothesis (which required that there would be similar numbers or fewer in the experimental areas where predators were removed). Under these circumstances, the null hypothesis is rejected and this provides evidence to support the hypothesis and therefore (assuming that the logical relationship between the model and the hypothesis is sound) to support the model. This is illustrated in Figure 1.

The only alternative outcome of a fair test is failure to refute or reject the null hypothesis. There are some statistical problems associated with this (*i.e.* considerations of power will usually come into this part of the procedure; Winer, 1971; Cohen, 1977; Underwood, 1981a; Toft & Shea, 1983). In general, however, provided that statistical problems have not been the cause of failing to reject the null hypothesis, it should be retained. Retention of a null hypothesis would demonstrate quite categorically that the model is wrong because its predictions are not borne out. Thus, as demonstrated in Figure 1, under circumstances in which all parts of this procedure can be done without new problems, there is an unambiguous, refutationist test either leading to rejection of the model itself (*i.e.* the null hypothesis was retained) or to rejection of the null hypothesis and thereby corroboration or support for the hypothesis and the model.

It would be naive to imagine that everything was as simple as outlined here. Nevertheless, this is a framework which has had some success in unravelling the complexities and variability of ecological studies. As a result, this is the framework which will be evaluated here for the specific purpose of determining whether it is useful, and/or used in the study of causes of boundaries of distribution of macroalgal species. There are, nevertheless, many situations in which this framework cannot be used. Some of these are very widespread. They usually relate to investigations of patterns (by statistical 'goodness-of-fit' tests) or when models have become very complex and very precise in their predictions leading to difficulties in separating the null hypotheses used in the logical

basis for doing an experiment, and the null hypothesis required for statistical testing. This whole area is fraught with difficulty, as discussed in Heath (1970) and Underwood (1990).

The logical bases of forming null hypotheses in ecological studies (*e.g.* of causes of limits to distribution of a species) are not identical with the needs for null hypotheses as used in the statistical sense. A statistical null hypothesis, as opposed to a logical one, is a statement which identifies the parameters of the frequency distribution of an appropriate test statistic under stated conditions (*e.g.* Mather, 1967). There are many difficulties of how to proceed when the only possible statistical null hypothesis is identical to the actual hypothesis deduced from the model. In this case, there is no way of inverting the hypothesis to form a logical null hypothesis. Consequently, there are enormous problems for determining the power of such tests, and the reliability of conclusions reached from them (see Underwood, 1990; for details). This will be considered only briefly here - even though it may be of major importance in many areas of ecology.

Thus, the sequence outlined in Figure 1 is that against which studies on the vertical patterns of distribution of intertidal macroalgae will be examined.

THE ESTABLISHMENT OF PATTERNS

Reliability of observations

One of the first problems to be dealt with before a series of ecological experiments can proceed is that of identifying with some reliability any patterns in nature. There have been many cases in which widespread patterns have been claimed, although there has been little evidence to support their existence (see particularly Whittaker, 1956; Underwood, 1978a; Hurlbert, 1981; Murray, 1982; Connor & Simberloff, 1984). Some procedure is clearly necessary to ensure that the patterns being explained really do exist. In some instances, this may seem a trivial task. In others, it is clearly problematic. One of the major reasons why it is necessary to have objective procedures for determining the existence of pattern is the very problem of 'theory-ladenness' of observation statements. It has been argued in many places (Hanson, 1959; Koyre, 1968; Popper, 1968; Feyerabend, 1975) that observations are made only in response to prevailing attitudes, theories, or, more recently, paradigms (Kuhn, 1970). The argument is that the only observations that get made are those that the observer is trained, willing or able to make, as a result of previous experience, predilections, biases and history of association with the ideas of others. This is a serious problem in many areas of ecology, leading to many observations being made that turn out to be not quite as originally described. Notably, oversimplification of patterns is usually the order of the day. Publishing descriptive accounts of very variable systems is usually very difficult. Editors and referees are more willing to give space to more simplistic accounts of the way biological systems operate. Thus, it is usually simpler to get an experimental study published in which there is a limited account of the distribution or abundance of the organisms that is being interpreted in the experiments (Andrew & Mapstone, 1987). Rarely is it possible to publish large-scale surveys describing the distribution and abundance in the first place (Connor & Simberloff, 1979). This is very important.

Foster (1990) has clearly demonstrated the great worth of describing from many sites the patterns of distribution and abundance of organisms. In his intriguing study, he was able to demonstrate that certain widely held and widely respected views about the structure of coastal assemblages on the west coast of the United States are not generally true, because the outcomes of these processes are not found in numerous sites along a coast. This is quite a different impression from that in the literature on the experimental studies in which the processes are described. Many of those papers have suggested that the processes are indeed widespread (*e.g.* Paine, 1974). It would be profitable to get observations and descriptions about the system from widespread localities before embarking on the small-scale studies that we mostly use to explain phenomena in nature.

Because observations are notoriously unreliable, without some objective verification, there can be dispute about the very beginnings of an ecological study.

Use of refutationist procedures in observation statements

Take a simple observation. It has been observed that, generally speaking, on the rocky shores of New South Wales there are far more abundant fleshy, foliose macro-algae lower down on the shore than at higher levels (Figure 2A). Often there is a fairly abrupt boundary above which foliose plants tend to be absent. This casual observation can be made by anyone almost anywhere at almost any time although there are some exceptions.

I believe that the observations are made because, on this coastline, there is a general pattern for foliose plants to be found abundantly below a certain limit and to be sparse above it. This model ("it is seen because it is so") must be distinguished by appropriate test from other models, such as "this observation is made because foliose plants above a certain limit on the shore are cryptic, difficult to see against the background, only visible at certain stages of the tide (they tend to lie flat during low tide)", *etc.* If any of these models were true, it would still be possible to make the casual observation outlined above, but it would not be the case that the observation was made because the world was really like that. Alternatively, the observation could be made simply because I and co-workers in this part of the world wish to impress this pattern on the rest of the universe because we are wilfully misleading everyone. In order to refute this last model, and in an attempt to refute the former ones which could also be true, it is necessary to examine the predictions that one would make from them.

First, it is necessary to identify an objective procedure for determining the existence and abundance of foliose plants. There can be debate about appropriate methods for doing this. To make things relatively simple, assume that agreement can be reached about how one identifies a foliose macro-alga. Second, presume that it is agreed that objective observers using some sort of randomized, representative sampling at different heights on a shore, will be able to determine the existence and to count the abundance of such algae (*e.g.* as percentage cover in quadrats). Then the distribution and abundance of foliose plants in the quadrats can be ascertained and therefore unambiguous objective measures could be available. Ideally, it might be best if this were done by people other than those who made the original casual observation. The model proposed is that the

original observations were made because the world really does consist of areas with abundant foliose plants low on the shore and none at higher levels. This model leads directly to the hypothesis that, if objective sampling and quantification of foliose plants are done at different heights on the shore, there will be considerably greater abundance at low than at high levels and, indeed, at high levels the abundance will be zero. This hypothesis leads directly to the null hypothesis that there will be the same average abundance at all levels on the shore, or that there will be a greater abundance at high levels than at low levels. The null hypothesis can be directly examined by sampling. The sampling is the test.

Such quantitative testing of this null hypothesis has been done for some shores in New South Wales (Underwood, 1981b). An example is indicated in Figure 2B. At low levels on the shore, the percentage cover of foliose plants is very large, in most areas reaching about one hundred percent. As one moves upshore, the cover declines to zero somewhere below mid-tide level. Coincidentally in sampling, other observations are made, notably that encrusting species of algae predominate at high levels (see data in Underwood, 1981b). Furthermore, grazing gastropods (notably limpets, topshells and neritid snails) are abundant above the level of foliose plants (Figure 2A).

In this case, the null hypothesis can clearly be rejected, thus providing corroboration for the hypothesis, and therefore the model, that the observations originally made (there are more foliose plants at low than at higher levels) were made because the world really was like that. This test, however simple, is not trivial. It was necessary to eliminate the possibility of delusion, wilful disregard of foliose plants at higher levels, *etc.* Models

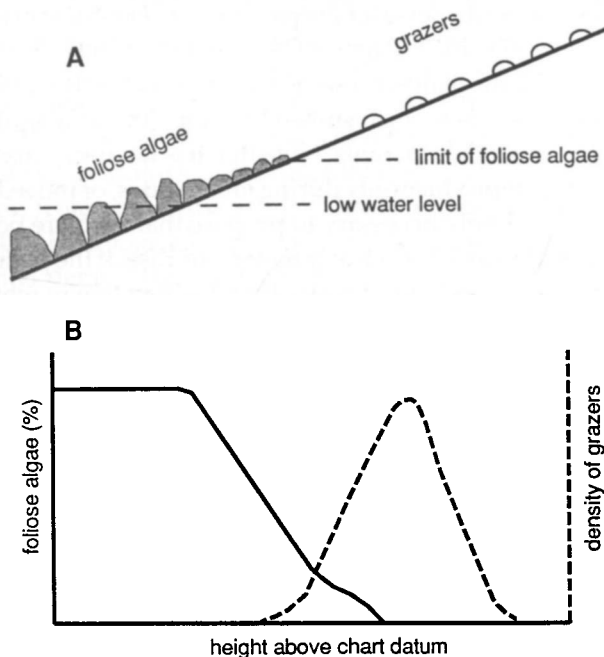


Figure 2. General patterns of distribution of foliose macro-algae on rocky shores in New South Wales. (A) Diagram indicating the major components of the system; (B) Summarized graph of abundances of foliose algae and grazers at different heights on a shore (see Underwood, 1981b for details).

based on these notions would have led to the hypothesis that there would be no difference in the abundance of foliose plants at low or high levels (or that there would be none above the boundary). These are, of course, the null hypothesis of the test.

Why bother with this rigmarole? The reason is that this is the only sure way that one can take one's original, initial observations to reflect the real world (*e.g.* Whittaker, 1956; Underwood, 1978a). On many occasions, original observations have turned out not to be based on a proper perception of the way the world is (formally, they were seen because of self-delusion).

MULTIPLE MODELS TO EXPLAIN PATTERNS

For a very long time it has been realized (Chamberlin, 1965, reprinted from 1880) that there are often many different and interactive ways in which observations might be explained. Thus, ecologists should usually be very careful to think about very broad classes of models and to consider many of them simultaneously in the early stages of an investigation. The continued attempt to identify the causes of limitations to patterns of distribution of a species, where only a single cause is investigated, is not likely to be very progressive. To date, many studies have demonstrated that an interaction of different processes is usually necessary to understand these types of patterns.

There are several models that might explain why there is an abrupt boundary to the limit of foliose plants on rocky shores in New South Wales (see also Chapman, 1973; Lubchenco, 1980; Underwood, 1980, 1985). First, is the possibility that the plants are limited by being unable to colonize the higher levels of the shore. There is evidence for several species of algae that the dispersal of propagules is over very short distances from the parent plant (Dayton, 1971, 1973; Paine, 1979; Lubchenco, 1980; Foster, 1990). If, in addition to limited dispersal, the plants, for some reason, tend to release their propagules only during falling tides, there is no mechanism by which the propagules could reach levels on the shore higher than their parents. Note that it is not necessary at this stage to demonstrate that the plants reproduce only during falling tides, or indeed that they have only limited dispersal. It is simply necessary to propose that such are possibilities. The rules for the construction of a model are clearly met in this case. If the plants have limited dispersal, coupled with some mechanism restricting the direction in which dispersal is made, then there would not be any foliose plants at high levels on the shore. Thus, the model can explain the observations. It is not necessary for a model explaining why an upper boundary of plants is observed now (*i.e.* is currently maintained) also to be an explanation of how such a pattern became established in the first place.

This is only one of the possible explanations. A second, often proposed in the literature, is that these species of plants tend not to extend to higher levels on the shore because of physiological incapacities. The notion is that the plants cannot tolerate excessive periods of emersion, during low tide. Thus, at higher levels on the shore, they would be subjected to progressively longer periods of potentially desiccatory aerial conditions. Under these conditions, they would die and that is why they are not found above a certain level. This is a very widespread explanation for the upper boundary of distribution of intertidal species (*e.g.* Connell, 1961a,b; Schonbeck & Norton, 1978). It is so widely

accepted that it has achieved the status of a paradigm and is, in many instances, no longer investigated. Thus, it is invoked to explain the upper boundaries of distribution of numerous intertidal species, even where it has never been tested (see examples in Underwood & Denley, 1984; see also the important paper by Wolcott, 1973). The model is a very clear one, and would explain the lack of foliose species above the upper boundary of physiological tolerance to aerial conditions.

A third alternative is that grazing by the animals at higher levels is responsible for the upper boundary of distribution of the plants (Hay, 1979; Underwood, 1980). The model requires that the animals are capable of removing all propagules of the lower-shore species of plants as they arrive. Note that this model is quite different from the previous two. It involves very different processes and the three models are, at least so far, unrelated.

Yet another, fourth, model is the possibility that the foliose plants from lower levels on the shore are somehow prevented from colonizing or becoming established successfully at higher levels by the existence of the encrusting species found there (see, for example, Lubchenco, 1980; Underwood, 1980). This model is again different from the previous ones and involves competitive interactions rather than the intrinsic properties of the plants (models 1 and 2 above) or the actions of animals (model 3 above). Note again that under this model it is not yet necessary to propose how exactly an encrusting species might prevent the establishment of the foliose plants.

There may be other models. Inevitably, at some point in the future there will be. For now, however, these four classes of models illustrate the plurality of potential explanations. Each, by quite different processes, provides a valid account of why the foliose plants may not be able to extend to higher levels on the shore.

Finally, of course, it may be necessary to invoke combinations of these models. Thus, it may be that the most realistic (or possibly most correct) model to explain the pattern of distribution of foliose plants on these shores is some combination of physiological intolerance to aerial conditions and susceptibility to grazing at high levels. Alternatively, plants suffering from some competitive disadvantage in the presence of encrusting species might be physiologically incompetent to tolerate periods of prolonged emersion (an experimental demonstration of this sort of process was described by Peterson & Black, 1988). There are numerous possible combinations of processes that form alternative models to the singular ones proposed above.

TESTABLE HYPOTHESES FROM THESE MODELS

Several hypotheses might be derived from the first model described above (upshore dispersal does not happen). One is that if some suitable, agreed, objective procedure (see, for example, the review by MacLulich, 1986) were used to sample representative areas of the shore above and below the upper boundary of foliose algae, there will be no propagules of the plants in the higher samples when there are some in the lower areas. The latter part is necessary to ensure that the plants are actually reproducing or recruiting, otherwise the hypothesis cannot be tested.

There are different ways in which null hypotheses might be phrased. The simplest is that there will be at least as many propagules in the higher samples as in the lower ones. This is, however, insufficient. Even if there were many (and statistically significantly) more propagules lower than higher on the shore, that could not explain the *absence* of plants at high levels. This seems obvious, but, as discussed later, has not been recognized in some studies on limits of distribution of species (*e.g.* Connell, 1961a; Schonbeck & Norton, 1978; see also Underwood & Denley, 1984).

An appropriate null hypothesis is that there will be some (*i.e.* more than zero) propagules in the higher samples when there is a non-zero number in the lower samples. This null hypothesis is logically problematic because of the need for statistical, probabilistic analysis of the samples. This is discussed in full elsewhere (Underwood, 1990). Unfortunately, it is not possible to use this statement as a null hypothesis because it would be impossible to construct a statistical test to determine the outcome of the test. If the samples from above the upper limit of the algae do contain some propagules, the null hypothesis would be retained and the hypothesis (and model) would be rejected. This is not problematic. The problem arises if there are some propagules on the shore above the upper limit of the foliose plants, but the mean density is small. Under these circumstances inadequate sampling (*i.e.* insufficient replication) may result in there being no propagules in the sample.

The null hypothesis would then be retained erroneously because of a Type II or Beta-error (Winer, 1971; Cohen, 1977; Underwood, 1981a, 1990; Toft & Shea, 1983). To avoid such errors, it is necessary to do a test of large power, but the calculation of power requires construction of an alternative hypothesis to the null (Cohen 1977; Underwood 1981a, 1990). This is not possible when the only prediction made from the model is the statistical null hypothesis (see Underwood, 1990; for details). Where the statistical null hypothesis is identical to the logical hypothesis, then a logically sound conclusion can be reached only if the null hypothesis (and the hypothesis and therefore the model) is rejected. If it is retained, no disproof of the hypothesis or model has been done and the conclusion that the model is correct is a fallacy ("confirming the consequent", Lemmon, 1971; Hocutt, 1979; Underwood, 1990). It would be better to avoid such hypotheses. A logically sound hypothesis is that if plants are absent only from the higher levels because they cannot get there, they should survive and grow if transplanted there as propagules (see below).

Two major classes of hypothesis have been deduced from the second model (that plants do not extend upshore because of physiological limitations). First is the proposition that, if algae are transplanted above their natural upper limit and others are translocated back into areas below that limit, the former will die of physiological stress at a rate faster than the translocated controls. The null hypothesis is that the two sets of plants will die at the same rate or that the ones transplanted will survive better than the lower controls.

This hypothesis has been tested in experiments by Schonbeck & Norton (1978). There are, however, logical problems in the use of this hypothesis as a basis for testing the stated model (see also Foster, 1982; Underwood & Denley, 1984). If, for example, the transplanted plants happened to survive well at the higher level, leading to retention of the

null hypothesis and therefore falsification of the model, this could be erroneous. It is not possible for the model to propose that *adult* plants would have to be the ones with physiological limitations. This would not explain the original observations; under this scenario, there would be juvenile plants at all levels on the shore and therefore no upper limit to the distribution of the plants could have been observed in the first place. Thus, survival of transplanted algae at higher levels could not be a logical basis for refuting the model as originally proposed.

In contrast, the transplanted algae could indeed die at a faster rate than the controls, thus refuting the null hypothesis and supporting the hypothesis and model. It would not be rational to conclude that the model was correct. Well-established plants low on the shore might be adapted, or habituated, to the conditions there (see Foster, 1982). If moved to a new habitat higher on the shore, they might die of some stress that would not kill them had they originally grown at the higher level. Inevitably, the model has been examined inappropriately, even if young, but well-established, algae are the ones transplanted (as done by Schonbeck & Norton, 1978).

The only stage of the life-cycle of the plants that should be transplanted are the earliest stages - the propagules that are the stages that actually recruit into the adults' habitat. Transplanting these and discovering they do not survive because of physiological stress (and not some other cause, such as being eaten) would reject the null hypothesis (that they will survive as well as, or better than, lower-level controls) and provide support for the model. It would also falsify the first model (from which it can be predicted that the propagules would survive if they could somehow reach the areas above the distribution of the adults). This demonstrates one advantage of considering more than one model at a time in field experiments. Transplantations of these early stages are likely to be more difficult than moving juvenile or adult plants. Nevertheless, it is possible to do such manipulations (*e.g.* Reed, 1990).

The alternative form of hypothesis is less subject to potential misinterpretation. This hypothesis requires that conditions above the upper limit of distribution are altered so as to ameliorate the physical harshness during low tide to match those prevailing lower on the shore. The plants should then become established at higher levels, because the causes of the upper limit (according to the model) have been removed. The hypothesis therefore predicts the establishment of plants in experimental plots and not in untreated control plots at the higher level. It would also be desirable to monitor some plots at the lower levels to demonstrate that plants continued to thrive there. The null hypothesis is then that there will be no difference in numbers or cover of foliose plants in the experimental and control plots at higher levels, or that more plants will become established in the controls.

Several methods have been used to reduce physical harshness of intertidal environments, including slowly-draining pools (Frank, 1965) and provision of shade (*e.g.* Underwood, 1980). Provided that these methods create conditions like those prevailing lower on the shore and that no uncontrolled artefacts (see particularly Dayton & Oliver, 1980) are introduced, the hypothesis can be tested.

The third model leads directly to the hypothesis that removal of grazers from areas above the boundary of distribution of the algae will lead to the appearance of plants at

higher levels. Again, control areas with grazers are necessary to demonstrate that there was no large-scale change on the shore leading, for example, to algae growing at higher levels than normal regardless of grazing. This hypothesis has been tested in New South Wales (Underwood, 1980, 1985; Underwood & Jernakoff, 1984). The null hypothesis is exactly that derived for the previous model: there will be no difference between numbers and cover of algae in experimental and control plots, or more in the latter than the former.

Finally, an obvious testable hypothesis from the fourth model (foliose plants are prevented from colonizing higher areas because of encrusting algae) is that if the encrusting algae are removed from areas above the upper limit of foliose algae, the foliose species will become established, thus extending to higher levels on the shore. As for the previous two models, unmanipulated control areas are needed. Several studies have included tests of such hypotheses about competition between algae as a process limiting distributions (Lubchenco, 1980; Schonbeck & Norton, 1980; Underwood, 1980; Foster, 1982; Hawkins & Hartnoll, 1985).

ADVANTAGES OF SIMULTANEOUS TESTS OF CONFLICTING HYPOTHESES

When, as is often the case, there are several quite different models, from which alternative hypotheses can be generated, it is advisable to choose the hypotheses that are most easily contrasted (as discussed earlier). It may also be practically and logistically advisable to choose the various possible hypotheses derived from each model so that they can be evaluated in the same type of, or the same, experiment. This would simplify and reduce the planning and work of a series of experiments to contrast the success with which the conflicting array of models led to accurate predictions under the experimental, test conditions. For the array of hypotheses discussed above, there have been experiments (Underwood, 1980) using combinations of the requisite treatments. This has the huge advantage that the controls for one hypothesis were identical to the controls for some of the others, reducing the total work done to a minimum. Appropriate unmanipulated controls for the hypotheses from the second, third and fourth models are essentially the same (*i.e.* controls are unmanipulated areas above the upper limit of foliose plants, from which no grazers or encrusting algae have been removed, and in which there is no deliberate or accidental alteration of the physiological conditions during low tide).

There is, however, a much more cogent reason why the various alternatives should be examined simultaneously. Only where it is possible to identify the effects, for example, of encrusting algae on the distribution of foliose plants (if competitive interactions between them are occurring) in conjunction with the effects of grazers (if they are also relevant) would it be possible to examine a model that the distribution of the algae is a product of the interaction between these two processes.

Lubchenco (1980) demonstrated this very clearly when she examined the causes of the lower limit of distribution of fucoid algae on a shore dominated by encrusting holdfasts of *Chondrus* at the lower levels below the fucoids. In her work, three different models, essentially those discussed as two, three and four above, can be identified. Thus, she considered the possibilities that the upper-shore plants were limited in their downward

extension by physiological inadequacies, preventing them from occupying lower levels on the shore, or because grazing by snails or urchins removed the fucoids from lower levels, or because the encrusting plants lower down prevented their establishment. Similar hypotheses to those discussed above (but inverted to take into account that the lower and not the upper limit was being studied) were examined in experiments.

Lubchenco (1980) cleared areas low on the shore of *Chondrus* and, in some of these, removed the grazers. As a result, she was able to demonstrate, by rejection of the appropriate null hypotheses, that *Fucus* would colonize areas lower on the shore than they were normally found, but only where *Chondrus* was removed. This rejected the null hypothesis that removal of the putative competitor would make no difference and simultaneously refuted the hypothesis that the plants were incapable of living at lower levels because of some physiological limitation.

She further showed (Lubchenco, 1980) that where grazing snails (*Littorina*) were present, they could remove all of the *Fucus* at lower levels, except those individuals that became established in crevices. In the presence of *Chondrus*, nevertheless, there was no role for grazers, because the *Fucus* were not able to settle or survive. In these factorial experiments, where grazers were excluded from areas with and without crusts, the interpretation is very straightforward, and the results revealed the interaction between grazing and competition from encrusting species. Even in the absence of *Chondrus*, *Fucus* should not become established at low levels on the shore because of grazers, except where there are physical refuges from grazing (such as crevices).

Similarly, the results of multifactorial experiments described in Underwood (1980) revealed that grazers were responsible for the absence of foliose species of algae at levels on the shore above their normal upper limit. Nevertheless, the experimental reduction of physical stress, in the absence of grazers, allowed plants to become adult. Again, there was an interaction between the two processes of physical stress and grazing that influenced the sizes and abundances of the plants.

In some studies, because several alternative mechanisms were not investigated simultaneously, there has been confusion about the eventual interpretation of results. For example, Schonbeck & Norton (1980) demonstrated that the high-shore species, *Pelvetia canaliculata*, could become established at lower levels on the shore where the canopy of *Fucus spiralis* was removed. *Pelvetia* also survived under an established canopy of *Fucus*, so the lower limit of distribution of the plant was not explicable by competition (although the reduced density of *Pelvetia* where *Fucus* was present was certainly a result of competition). Schonbeck & Norton (1980) were forced to speculate about the probable causes of the lower limit of distribution of the species and invoked some uninvestigated aspect of grazing by *Littorina* as the ultimate cause of disappearance of those germlings of *Pelvetia* that manage to survive under the canopy of the superior competitor.

There are other cases of experimental studies in which the relationship between the outcome of experiments to test a single model, *i.e.* about a single process, have not been entirely satisfactory and led immediately to other possible models being discussed, but dismissed (reviewed by Underwood & Denley, 1984). It seems that attention to alternative models as explanations for some observed phenomenon would be a profitable prerequisite for any study. The logical procedure discussed at length here makes this

imperative - the reason for doing experiments in the first place is to eliminate those models that are incorrect as explanations. This is a far cry from experimental studies aimed at confirming a singular model, which, when hypotheses are not borne out, provide little guidance as to what should happen next.

TWO DIVERSIONS

The logic of replication

Another manner in which close attention to the requirements of the logic of experimental proceedings may help to avoid problems is in the area of experimental design. The principles of factorial designs of experiments and the needs for adequate provision of replication are long-established and supposedly widely known. For example, 60 years ago optimists were writing that no longer should there be any studies without replication (Fisher & Wishart, 1930). The optimists have been sadly astray (see the reviews in Underwood, 1981a, 1986; Hurlbert, 1984). There are serious deficiencies in many ecological experiments, but the lack of replication is one of the most serious because it leads directly to irrational conclusions from experiments.

Hurlbert (1984) discussed at length what he described as 'pseudoreplication', apparent replication at the wrong spatial or temporal scales or of inappropriate experimental units. The error of inadequate or no replication is part of a class of problems known as confounding. There is a very good logical reason why replication is necessary, apart from considerations of statistical procedures and the need to provide some assessment of variances of magnitudes or rates of processes operating in experimental plots. If a null hypothesis is tested by use of a single experimental plot and a single control plot, there is no doubt that the null hypothesis should be rejected if, after the experiment, there is some difference between the two plots, but the difference may not be due to the experimental treatment.

There are too many studies in which there is no replication of any of the experimental treatments (Hurlbert, 1981; Underwood, 1986). Examples from recent studies of causes of limits to distribution of intertidal algae include those of Hawkins & Hartnoll (1985).

If, for example, one plot has a species removed and a control plot does not and each plot is sampled by several replicated representative quadrats, statistical tests are appropriate to determine how likely it is that the two plots differ by chance alone. Statisticians have no problem with advising that such tests are appropriate even though there is only one plot of each type. The problem is that there is no method available to determine whether the experimental treatment is itself responsible for the difference (as required by the hypothesis) or whether the difference is due to some other, uncontrolled, variable. Statisticians cannot bear the responsibility of the illogical conclusions about *causes* of differences between the two plots that are then made by ecologists. The illogical conclusions reached by ecologists are their own fault.

The logic of controls

Further confounding of experimental studies occurs wherever there is no proper control for the experimental manipulations themselves. In the context of intertidal

experimental studies, several cautionary tales can be learned from the review by Dayton & Oliver (1980). Little needs to be said to underline the importance of attempting to control for any artefacts of the experimental manipulation. For example, if experimental fences are used to keep grazing animals out of areas, as part of tests of some null hypothesis, it must be possible to demonstrate that the subsequent growth of algae was due to the removal of the grazers and not some aspect of interference with the physical environment due to the fences themselves. There has been appropriate discussion of this point in some papers (see Jernakoff, 1983, as one example) and Kennelly (1983) recommended a useful procedure for examining potential artefacts in grazing experiments.

Again, the needs of the logical procedure require adequate controls for experimental manipulations. It is illogical to construct an experimental test of some carefully articulated logical null hypothesis and then to confound the interpretation of the results of the test by not considering alternative explanations for the results obtained. The recommended procedure requires that alternative hypotheses based on alternative models be considered. Thus, if it can be proposed or even imagined that cages used to remove grazers will themselves cause increased settlement or survival of algae (because of artificial alterations of the physical environment), this must be considered as an alternative model in the experiment. Its hypothesis, that the presence of cages will cause an increase in cover of plants compared with uncaged controls regardless of the removal of the grazers, is identical in wording to that proposed from the model that grazers are responsible for the lack of algae in some area. Thus, any experiment that cannot distinguish between the two is not a valid experimental test of the hypothesis about grazers.

There seem to be several areas of intertidal ecology where this sort of confounding, due to artefacts in experiments, is widespread (Dayton & Oliver, 1980). This is particularly true in studies involving transplantation of organisms from one habitat to another to test hypotheses about the cause of upper or lower limits of distribution of organisms and on studies about the behaviour of organisms at different heights on a shore. The topic has been reviewed extensively elsewhere (Chapman, 1986; Underwood, 1986; Chapman & Underwood, 1990) and therefore does not need to be amplified further here.

OUTCOMES OF EXPERIMENTS: WHAT HAPPENS TO TESTED MODELS?

Continued re-evaluation

After any experimental test is completed, if it fulfils the requirements of an experiment as defined here, there are only two possible outcomes. Either the null hypothesis is retained and the hypothesis and model from which it was deduced must be rejected, or its null hypothesis is rejected and the hypothesis and model are retained (Figure 1). In the former case, the original observations and those gained during the experiment then form a basis for starting to define a suitable new explanatory model. This part of the proceedings is straightforward.

In contrast, however, there is a temptation to be complacent about the outcome when a model appears to have been supported by an experimental evaluation of its predictions.

This is a dangerous procedure according to the logic of experiments. First, some other model that has not yet been thought of may also explain the original observations, but may also lead to the same hypothesis and the same results in such an experiment. Under these circumstances, the model is wrong, but the experiment apparently supported it. This is one of the errors that can be made even when using a logical procedure carefully (see Underwood, 1990, for details of other errors and combinations of errors in ecological experiments).

For example, two species of algae may live at adjacent heights on a shore in such a pattern that competitive interactions between them, with one species winning at one height and the other winning at the other, could explain their disjunct distributions. One hypothesis derivable from this model is that removal of the lower species should result in the downward extension of the upper species, in comparison to control plots where the lower species is left intact (see Lubchenco, 1980; Schonbeck & Norton, 1980; Foster, 1982; Hawkins & Hartnoll, 1985; among many others). The appropriate null hypothesis could then be tested by removing the lower species from some plots. Often in such experiments, the upper species is then found to extend to lower than normal levels on the shore.

Is it clear from this type of experiment that the cause of the normal restriction of distribution is actually competition (for light or space on which to attach)? All that is known so far is that the lower species is implicated in the restriction of the upper species to the higher levels. The particulars of the resource for which competition might be occurring, or the mechanisms by which the species are competing have not been discussed (see reviews in Birch, 1979; Connell, 1983; Underwood, 1984, 1986; Schoener, 1983). Suppose that the lower species carries a disease or harbours a grazer that affects the upper species. Under either of these quite different models, removal of the lower species should be predicted to lead to a downward extension of the upper species. Thus, there is no way in such a study so far, to be sure that competitive interactions between the two species are actually occurring.

Whenever a model has withstood tests of null hypotheses derived from it, the next step is to think of more ingenious methods for trying to demonstrate that the model is wrong. For example, apart from invoking alternative mechanisms (as in the cases of competitive interactions just discussed), there are other possible major developments of the model. First, it could be expanded, enlarged, and so forth, to make it more general and therefore more in need of testing. A more general model should also explain phenomena about other species, habitats or processes. Second, it could be particularized, *i.e.* made to be more precise about the phenomena it explains and therefore the hypotheses derived from it. Or, a model can be re-evaluated by modifying it to explain other relevant observations about the system being studied. Alternatively, a model can be subjected to increasingly stringent tests by deducing new hypotheses from it about related processes. These alternatives are illustrated below.

Making a model more general

Models for the causes of distributional limits of intertidal algae can be made more robust by attempts to generalize to other species, to other types of shores (*e.g.* under

conditions of wave-exposure), or to different times of the year. If grazing is the cause of upward limitation of species (Underwood, 1980), then it should be consistent throughout different seasons of the year and under different conditions of wave-action. If it is not consistent, then predictions made about it for different times of the year and different degrees of exposure to wave-shock will not be correct and null hypotheses based on them will be retained. This would require modification of the model to ensure that it includes other factors to explain why the outcome of grazing cannot always be predicted and then subsequent, independent testing of the modified model.

One of the most elegant attempts to do this for intertidal organisms and theories is that by Foster (1990). He examined a number of models that had already withstood small-scale testing in detailed studies on one or a few shores (particularly, Foster, 1982). From these, he made predictions about how widespread certain patterns should be on shores along the Californian coast. For example, earlier studies had suggested that competition between mid- to high-shore algae, *Endocladia muricata* and *Mastocarpus papillatus* and a lower species *Iridaea flaccida* should cause a definable boundary between the upper species and *Iridaea*. If the competitive model is of any general worth, it leads to the prediction that the distribution of the two species of algae on many randomly-chosen shores should always conform to this pattern. If the model be inadequate or incomplete, this prediction will not be borne out because other processes not included in the model will also come into play. Thus, the model can be re-evaluated against the general alternative model that other processes need to be included. Foster (1990) examined the predictions of these two types of models and found that the competitive model was unable to explain much of the pattern of vertical distribution of the species of algae along the coast of California. In several sites, *I. flaccida* extended up into, and often occupied much space in, the higher levels where it was not supposed to overlap with the higher-shore species.

Foster's (1990) study is exemplary of the sort of general quantitative examination of the predictions made from restricted numbers and types of localities which are crucial before generalizations (*i.e.* general, large-scale models) are accepted (Underwood & Denley, 1984).

Making a model explain more types of observations

One of the ways in which a model can be continuously evaluated is to use it to try to explain related observations that were not made at the start of the study. For example, one interesting question in an intertidal habitat where grazers can eliminate plants from higher levels on a shore is why the grazers do not extend to lower levels and remove algae all the way down. Any model that explains why the plants have an upper boundary should also explain why the boundary is where it is. This has been explored for the New South Wales coastline in experiments by Underwood & Jernakoff (1981). A number of possible models (essentially the mirror image of those used for the upper limits of the plants) could be proposed to explain why there is a lower limit to the grazers. Underwood & Jernakoff's (1981) study demonstrated that the grazers were unable to keep areas of the shore free from algae unless they were in very large, and artificially maintained numbers.

At the low levels on the shore, algae grew so quickly that the grazers were completely swamped, had no free space over which to graze on microalgae, and therefore kept dying of starvation. Only in experimental areas where densities of grazers were maintained by addition of new individuals at regular intervals, were the limpets able to keep the substrata free of macro-algae. Alternative models that predation removed the grazers from low levels, or that the grazers could not withstand the physiological conditions prevailing there, were eliminated.

Thus, the original model about the causes of upper limits of foliose plants being determined by grazers now has the added structure that the grazers cannot extend to lower levels because of the presence of the plants.

Making a model more precise

The second issue that then should be explained from this model is why the boundary is at the particular location it is. Given that there are sensible models to explain the upper boundary of the plants and a different model to explain the lower boundary of the grazers, it should be possible to establish why those boundaries happen to be at a particular height on the shore. The explanation certainly seems to be that the boundary will be at that height on the shore at which the rapid growth rate of plants following their escape from grazing is such that they grow too large to be eaten by micro-algal grazers within a very short space of time. Thus, if propagules of the algae are not consumed very quickly, they grow too large to be eaten at a fast enough rate and will then be safe from micro-algal grazers. Under these circumstances, the boundary should be at that height on the shore at which the growth rate of the plants coupled with the probability of them escaping from grazing for a short period of time is sufficient to enable them to grow and escape from their consumers. This is a reflection of the model developed by Lubchenco & Gaines (1981). This model and its application to the situation on the coast of New South Wales has not yet been tested fully.

A further point from the above is that there should be grazers of the types found above the upper limit of foliose algae in any place where some process eliminates the plants themselves. Under the model that the grazers are limited in their downward extension by competition for space from the plants, anywhere where the plants are removed should be a suitable habitat for these grazers to live. This has been demonstrated for some intertidal grazers in New South Wales which also occur below low tide. *Cellana tramoserica*, one of the abundant limpets at mid-levels above foliose macro-algae on rocky shores in New South Wales, also extends to sublittoral regions (Underwood & Jernakoff, 1981; Andrew & Underwood, 1989; Fletcher, 1987). The model that accounts for the lower limit of the intertidal populations of limpets should also be able to explain why they can live below low tide. Hypotheses from this model have been tested and it has been demonstrated experimentally that, in areas where urchins are present, the urchins remove large macroscopic foliose plants, making space available for the limpets (Fletcher, 1987; Fletcher & Underwood, 1987). This is a parallel situation to that found in a number of other parts of the world (e.g. Ayling, 1981; Duggins, 1981; Choat & Schiel, 1982; Dethier & Duggins, 1984).

And so the model needs to be extended even further. In this particular case, further extensions should revolve around making the model explain why there are any foliose algae at all. If the microalgal grazers high on the shore and macroalgal grazers low on the shore can both eliminate the plants, why is there a zone occupied by foliose macro-algae in between the limits of these two grazers? One model that could account for this is that urchins are unable to function effectively as foragers in regions where they are subjected to wave action at low levels on the shore. Thus, they are confined to areas below low tide, leaving a margin of relatively ungrazed habitat between the microalgal grazers above and the areas where urchins are effective below. This model has not yet been tested.

Novel predictions from a model

Another area in which the model can be subjected to increased refinement of test, is to make it predict things about other processes that were not the ones that it was originally trying to explain. For example, under the model that upper limits of foliose plants on the shore are caused by grazing by microalgal grazers, it must be predicted that any areas in those high regions that are free from grazers should develop foliose macroalgae. This has been examined under two different circumstances for the shores of New South Wales. Creese & Underwood (1982) described macroalgal grazing limpets (*Siphonaria virgulata* and *S. denticulata*) grazing on algae growing on the backs of other limpets. The surfaces on the shells of limpets such as *Cellana tramoserica* are not grazed (see Branch & Branch, 1980). As a result, algae develop there. Similarly, McGuinness & Underwood (1986) found more algae growing on the tops of experimental boulders where grazers were absent, than on those boulders where grazers were in enhanced densities because of the provision of experimental crevices. Thus, the model was able to make quite new predictions about parts of the system that had not previously been examined.

Other predictions that have been successfully explored for this model are that any area higher on the shore, or any conditions higher on the shore, that enhance the growth of the algae should increase their abundance because the boundary will move upshore. Two different circumstances are known. One is that there are seasonal shifts in the distribution of the plants (Underwood, 1981b) because, presumably, the rate of growth of the algae is altered (although it is possible that the grazers are less effective during the cooler periods of the year - this remains to be investigated). The second is the demonstration that in areas where shallow experimental rock pools were made at higher levels on the shore, algae would grow profusely, because the larger grazers were not active and the algae were able to grow faster in these pools (Underwood & Jernakoff, 1984).

Finally, there is another prediction that should be made from this model. If grazers are capable of eliminating all propagules of foliose plants at high levels, they must be completely utilizing the resources of food. Otherwise, some of the propagules would escape, would eventually grow, and there would be no boundary to the upper limit of their distribution on the shore. This predicts that the densities at which grazers are routinely operating should be close to their carrying capacity. Competition between the grazers might therefore be expected to be intense anywhere where their densities are increased. This prediction has been borne out in several experimental studies (Creese,

1980; Fletcher, 1984; Underwood, 1976, 1978b, 1984).

Thus, a model can grow and be enhanced by further derived predictions and their subsequent tests. Whatever process is used, models must always be probed by renewed and more rigorous experimentation.

CONCLUSIONS

A case has been made that more attention to the logical steps and processes that are part of the experimental procedures would enhance the experimenters' capacity to interpret the system being studied. By careful identification and separation of the different stages of the logical process, it is apparent that some problems of experimental studies might be avoided. First, following this scheme (or some other scheme that is equally clarified and justified) requires that alternative explanations be given proper consideration before experimental studies are started. The elimination of alternative models is the *raison d'être* of the experimental study. This is made clear in this procedure, because it is recognised that there is never likely to be only one explanation for observed phenomena. Consideration for alternative mechanisms and alternative processes is a very important component in any ecological study (see Underwood & Denley, 1984, for examples and Hilborn & Stearns, 1986, for some theoretical and practical problems). This scheme of logic has alternative models as one of its primary foci.

Second, the scheme suggests ways in which better understanding of interactions between processes might be demonstrated. First, articulation of alternative models also carries with it the need to articulate explicitly models that involve more than one factor operating in conjunction, or in opposition. The logic of the scheme would not allow one model (e.g. grazers cause a boundary) and another model (e.g. physiological intolerance causes a boundary) to be considered only in isolation. It is only logical that a third model (some combination of grazing and physiological intolerance causes the boundary) should also be evaluated. This would lead to better consideration of multifactorial experimental studies. Such studies have proven of great value in trying to understand the complexities of some ecological situations (e.g. Underwood *et al.*, 1983; Dungan, 1986; Lively & Raimondi, 1987).

A third feature of this procedure that would assist in many areas of ecological experimentation is the requirement that alternative processes, including nuisance ones, that cause confounding in experimental studies, must be considered very carefully. This should lead to better consideration of experimental designs. This is clearly an area in which, in many studies, ecologists have demonstrated inefficiency or incompetence (Hurlbert, 1984; Underwood, 1981a, 1986; Dayton & Oliver, 1980; Underwood & Denley, 1984; Chapman, 1986). Anything that improves considerations of experimental designs in complex, very variable, very interactive ecological studies would be a good thing.

Finally, for good or ill, whatever the logic is underlying any given series of experimental studies, it should always be clearly articulated so that others can evaluate it. It is not necessary that the scheme followed here will be the only one used. It is, however, necessary that the logical basis and relationships between the components of experimental studies are clear. Clarification before the study would obviously help. Clarification

during the published account of the study is absolutely mandatory. Whatever alternative schemes are used, are successful, and are developed, they must be identifiable. In this brief review of one small area of marine ecology, some studies do not have such a close relationship between the original intentions and the final outcomes as was hoped. With any luck, a review such as this may serve to focus on the relevant issues at each part of such studies.

The preparation of this manuscript was supported by grants from the Australian Research Council, the University of Sydney Research Grant and the Institute of Marine Ecology. I am grateful for discussion over many years with, among others, L.C. Birch, M.G. Chapman, J.H. Connell, P.G. Fairweather, R.T. Hinde, P. Jernakoff, K.A. McGuinness, D. Simberloff, W.P. Sousa and P.A. Underwood which clarified my, probably still confused, thoughts. Useful comments on the manuscript were made by K. Astles and M.G. Chapman.

REFERENCES

- Andrew, N.L. & Mapstone, B.D., 1987. Sampling and the description of spatial pattern in marine ecology. *Oceanography and Marine Biology, Annual Review*, **25**, 39-90.
- Andrew, N.L. & Underwood, A.J., 1989. Patterns of abundance of the sea urchin *Centrostephanus rodgersii* (Agassiz) on the central coast of New South Wales, Australia. *Journal of Experimental Marine Biology and Ecology*, **131**, 61-80.
- Ayling, A.M., 1981. The role of biological disturbance in temperate subtidal encrusting communities. *Ecology*, **62**, 830-847.
- Birch, L.C., 1979. The effect of species of animals which share common resources on one another's distribution and abundance. *Fortschritte der Zoologie*, **25**, 197-221.
- Branch, G.M. & Branch, M.L., 1980. Competition between *Cellana tramoserica* (Sowerby) (Gastropoda) and *Patiriella exigua* (Lamarck) (Asteroidea) and their influence on algal standing stocks. *Journal of Experimental Marine Biology and Ecology*, **48**, 35-49.
- Carnap, R., 1962. *Logical Foundations of Probability*. Chicago: University of Chicago Press.
- Chalmers, A.F., 1979. *What is This Thing Called Science?* Brisbane: Queensland University Press.
- Chamberlin, T.C., 1965. The method of multiple working hypotheses. *Science, New York*, **148**, 754-759.
- Chapman, A.R.O., 1973. A critique of prevailing attitudes towards the control of seaweed zonation on the sea shore. *Botanica Marina*, **16**, 80-82.
- Chapman, M.G., 1986. Assessment of some controls in experimental transplants of intertidal gastropods. *Journal of Experimental Marine Biology and Ecology*, **103**, 181-201.
- Chapman, M.G. & Underwood, A.J., in press. Experimental designs for analyses of movements by gastropods. *Journal of Molluscan Studies*.
- Choat, J.H. & Schiel, D.R., 1982. Patterns of distribution and abundance of large brown algae and invertebrate herbivores in subtidal regions of northern New Zealand. *Journal of Experimental Marine Biology and Ecology*, **60**, 129-162.
- Cohen, J., 1977. *Statistical Power Analysis for the Behavioural Sciences*. New York: Academic Press.
- Connell, J.H., 1961a. The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. *Ecology*, **42**, 710-723.
- Connell, J.H., 1961b. Effects of competition, predation by *Thais lapillus*, and other factors on natural populations of the barnacle *Balanus balanoides*. *Ecological Monographs*, **31**, 61-104.
- Connell, J.H., 1983. On the prevalence and relative importance of interspecific competition: evidence from field experiments. *American Naturalist*, **122**, 661-696.
- Connor, E.F. & Simberloff, D., 1979. You can't falsify ecological hypotheses without data. *Bulletin of the Ecological Society of America*, **60**, 154-155.
- Connor E.F. & Simberloff D., 1984. Neutral models of species' co-occurrence patterns. In *Ecological*

- Communities: Conceptual Issues and the Evidence* (ed. D.R. Strong, Jr, et al.), pp. 316-331. New Jersey: Princeton University Press.
- Connor, E.F. & Simberloff, D., 1986. Competition, scientific method and null models in ecology. *American Scientist*, **75**, 155-162.
- Creese, R.G., 1980. An analysis of distribution and abundance of populations of the high-shore limpet *Notoacmea petterdi* (Tenison-Woods). *Oecologia*, **45**, 252-260.
- Creese, R.G. & Underwood, A.J., 1982. Analysis of inter- and intra-specific competition amongst limpets with different methods of feeding. *Oecologia*, **53**, 337-346.
- Dayton, P.K., 1971. Competition, disturbance and community organization: the provision and subsequent utilization of space in a rocky intertidal community. *Ecological Monographs*, **41**, 351-389.
- Dayton, P.K., 1973. Dispersion, dispersal, and persistence of the annual intertidal alga, *Postelsia palmaeformis* Ruprecht. *Ecology*, **54**, 433-438.
- Dayton P.K. & Oliver J.S., 1980. An evaluation of experimental analyses of population and community patterns in benthic marine environments. In *Marine Benthic Dynamics* (ed. K.R. Tenore and B.C. Coull), pp. 93-120. Columbia: University of South Carolina Press.
- Dethier, M.N. & Duggins, D.O., 1984. An 'indirect commensalism' between marine herbivores and the importance of competitive hierarchies. *American Naturalist*, **124**, 205-219.
- Diamond J.M., 1986. Overview: laboratory experiments, field experiments and natural experiments. In *Community Ecology* (ed. J.M. Diamond and T.J. Case), pp. 3-22. New York: Harper & Row.
- Duggins, D.O., 1981. Interspecific facilitation in a guild of benthic marine herbivores. *Oecologia*, **48**, 157-163.
- Dungan, M.L., 1986. Three-way interactions: barnacles, limpets, and algae in a Sonoran desert rocky intertidal zone. *American Naturalist*, **127**, 292-316.
- Edson, M.M., Foin, T.C. & Knapp, C.M., 1981. 'Emergent properties' and ecological research. *American Naturalist*, **118**, 593-596.
- Fagerström, T., 1987. On theory, data and mathematics in ecology. *Oikos*, **50**, 258-261.
- Feyerabend, P.K., 1975. *Against Method*. London: New Left Books.
- Fisher, R.A. & Wishart, J., 1930. The arrangement of field experiments and the statistical reduction of results. *Technical Communications of the Imperial Bureau of Soil Science*, **10**, 1-23.
- Fletcher, W.J., 1984. Intraspecific variation in the population dynamics and growth of the limpet, *Cellana tramoserica*. *Oecologia*, **63**, 110-121.
- Fletcher, W.J., 1987. Interactions among subtidal Australian sea urchins, gastropods and algae: effects of experimental removals. *Ecological Monographs*, **57**, 89-109.
- Fletcher, W.J. & Underwood, A.J., 1987. Interspecific competition among subtidal limpets: effect of substratum heterogeneity. *Ecology*, **68**, 387-400.
- Foster, M.S., 1982. Factors controlling the intertidal zonation of *Iridaea flaccida* (Rhodophyta). *Journal of Phycology*, **18**, 285-294.
- Foster, M.S., 1990. Organization of macroalgal assemblages in the Northeast Pacific: the assumption of homogeneity and the illusion of generality. *Hydrobiologia*, **192**, 21-33.
- Frank, P.W., 1965. The biogeography of an intertidal snail population. *Ecology*, **46**, 831-844.
- Hacking, I., 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.
- Hanson, N.R., 1959. *Patterns of Discovery*. Cambridge: Cambridge University Press.
- Hatton, H., 1932. Quelques observations sur le peuplement en *Fucus vesiculosus* des surfaces rochers dénudés. *Bulletins de la Laboratoire Marine du Musée d'Histoire Naturelle*, **9**, 1-6.
- Hatton, H., 1938. Essais de bionomie explicative sur quelques espèces intercotidales d'algues et d'animaux. *Annales de l'Institut Oceanographie, Monaco*, **17**, 241-348.
- Hawkins, S.J. & Hartnoll, R.G., 1983. Grazing of intertidal algae by marine invertebrates. *Annual Review of Oceanography and Marine Biology*, **21**, 195-282.
- Hawkins, S.J. & Hartnoll, R.G., 1985. Factors determining the upper limits of intertidal canopy-forming algae. *Marine Ecology Progress Series*, **20**, 265-271.
- Hay, C., 1979. Some factors affecting the upper limit of the southern bull kelp *Durvillaea antarctica* (Chamisso) Hariot on two New Zealand shores. *Journal of the Royal Society of New Zealand*, **9**, 279-289.
- Heath, O.V.S., 1970. *Investigation by Experiment*. London: Edward Arnold.
- Hempel, C.G., 1965. *Aspects of Scientific Explanation*. New York: Free Press.

- Hempel, C.G., 1966. *Philosophy of Natural Science*. New Jersey: Prentice-Hall.
- Hilborn, R. & Stearns, S.C., 1982. On inference in ecology and evolutionary biology: the problem of multiple causes. *Acta Biotheoretica*, **31**, 145-164.
- Hocutt, M., 1979. *The Elements of Logical Analysis and Inference*. Cambridge: Winthrop.
- Hume, D., 1779. *Dialogues Concerning Natural Religion* 1779. 2nd edition, 1947. London: Nelson.
- Hurlbert, S.H., 1981. A gentle depilation of the niche: dicean resource sets in resource hyperspace. *Evolutionary Theory*, **5**, 177-184.
- Hurlbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, **54**, 187-211.
- Jackson, J.B.C., 1981. Interspecific competition and species' distributions: the ghost of theories and data past. *American Zoologist*, **21**, 889-901.
- Jernakoff, P., 1983. Factors affecting the recruitment of algae in a midshore region dominated by barnacles. *Journal of Experimental Marine Biology and Ecology*, **67**, 17-31.
- Kennelly, S.J., 1983. An experimental approach to the study of factors affecting algal colonization in a sublittoral kelp forest. *Journal of Experimental Marine Biology and Ecology*, **68**, 257-276.
- Koyre, A., 1968. *Metaphysics and Measurement*. London: Chapman & Hall.
- Kuhn, T., 1970. *The Structure of Scientific Revolutions*. 2nd edition. Chicago: University of Chicago Press.
- Lakatos I., 1974. Falsification and the methodology of scientific research programmes. In *Criticism and the Growth of Knowledge* (ed. I. Lakatos and A.E. Musgrave), pp. 91-96. Cambridge: Cambridge University Press.
- Lehmon, E.J., 1971. *Beginning Logic*. Surrey: Nelson.
- Lively, C.M. & Raimondi, P.T., 1987. Desiccation, predation, and mussel-barnacle interactions in the northern Gulf of California. *Oecologia*, **74**, 304-309.
- Loehle C. J., 1987. Hypothesis testing in ecology: psychological aspects and the importance of theory maturation. *Quarterly Review of Biology*, **62**, 397-409.
- Lubchenco, J., 1980. Algal zonation in the New England rocky intertidal community: an experimental analysis. *Ecology*, **61**, 333-344.
- Lubchenco, J. & Gaines, S.D., 1981. A unified approach to marine plant-herbivore interactions. I. Population and communities. *Annual Review of Ecology and Systematics*, **12**, 405-437.
- MacLulich, J.H., 1986. Experimental evaluation of methods for sampling and assaying intertidal epilithic microalgae. *Marine Ecology Progress Series*, **34**, 275-280.
- Mather, K., 1967. *The Elements of Biometry*. London: Methuen.
- McGuinness, K.A. & Underwood, A.J., 1986. Habitat structure and the nature of communities on intertidal boulders. *Journal of Experimental Marine Biology and Ecology*, **104**, 97-123.
- McIntosh, R.P., 1985. *The Background of Ecology*. Cambridge: Cambridge University Press.
- Mentis, M.T., 1988. Hypothetico-deductive and inductive approaches in ecology. *Functional Ecology*, **12**, 1-5.
- Murray, B.G., Jr, 1982. On the meaning of density-dependence. *Oecologia*, **53**, 370-373.
- Nagel, E., 1961. *The Structure of Science*. London: Harcourt, Brace, World.
- Paine, R.T., 1974. Intertidal community structure. Experimental studies on the relationship between a dominant competitor and its principal predator. *Oecologia*, **15**, 93-120.
- Paine, R.T., 1979. Disaster, catastrophe, and local persistence of the sea palm *Postelsia palmaeformis*. *Science, New York*, **205**, 685-687.
- Peterson, C.H. & Black, R., 1988. Density-dependent mortality caused by physical stress interacting with biotic history. *American Naturalist*, **131**, 257-270.
- Platt, J.R., 1964. Strong inference. *Science, New York*, **146**, 347-353.
- Popper, K.R., 1968. *The Logic of Scientific Discovery*. London: Hutchinson.
- Popper, K.R., 1969. *Conjectures and Refutations*. London: Routledge & Kegan Paul.
- Reed, D.C., 1990. The effects of variable settlement and early competition on patterns of kelp recruitment. *Ecology*, **71**, 776-787.
- Roughgarden, J., 1983. Competition and theory in community ecology. *American Naturalist*, **122**, 583-601.

- Russell, B., 1912. *Problems of Philosophy*. London: Willian & Norgate.
- Schoener, T.W., 1983. Field experiments on intraspecific competition. *American Naturalist*, **122**, 240-285.
- Schonbeck, M. & Norton, T.A., 1978. Factors controlling the upper limits of fucoid algae on the shore. *Journal of Experimental Marine Biology and Ecology*, **31**, 303-313.
- Schonbeck, M.W. & Norton, T.A., 1980. Factors controlling the lower limits of fucoid algae on the shore. *Journal of Experimental Marine Biology and Ecology*, **43**, 131-150.
- Simberloff D., 1980. A succession of paradigms in ecology: essentialism, materialism and probabilism. In *Conceptual Issues in Ecology* (ed. E. Saarinen), pp. 63-99. Dordrecht: Reidel.
- Toft, C.A. & Shea, P.J., 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. *American Naturalist*, **122**, 618-625.
- Trusted, J., 1979. *The Logic of Scientific Inference*. London: Macmillan.
- Underwood, A.J., 1976. Food competition between age-classes in the intertidal neritacean *Nerita atramentosa* Reeve (Gastropoda: Prosobranchia). *Journal of Experimental Marine Biology and Ecology*, **23**, 145-154.
- Underwood, A.J., 1978a. A refutation of critical tidal levels as determinants of the structure of intertidal communities on British shores. *Journal of Experimental Marine Biology and Ecology*, **33**, 261-276.
- Underwood, A.J., 1978b. An experimental evaluation of competition between three species of intertidal prosobranch gastropods. *Oecologia*, **33**, 185-202.
- Underwood, A.J., 1980. The effects of grazing by gastropods and physical factors on the upper limits of distribution of intertidal macroalgae. *Oecologia*, **46**, 201-213.
- Underwood, A.J., 1981a. Techniques of analysis of variance in experimental marine biology and ecology. *Oceanography and Marine Biology, Annual Review*, **19**, 513-605.
- Underwood, A.J., 1981b. Structure of a rocky intertidal community in New South Wales: patterns of vertical distribution and seasonal changes. *Journal of Experimental Marine Biology and Ecology*, **51**, 57-85.
- Underwood, A.J., 1984. Vertical and seasonal patterns in competition for microalgae between intertidal gastropods. *Oecologia*, **64**, 211-222.
- Underwood A.J., 1985. Physical factors and biological interactions: the necessity and nature of ecological experiments. In *The Ecology of Rocky Coasts* (ed. P.G. Moore and R. Seed), pp. 372-390. London: Hodder & Stoughton.
- Underwood A.J., 1986. The analysis of competition by field experiments. In *Community Ecology: Pattern and Process* (ed. J. Kikkawa and D.J. Anderson), pp. 240-268. Melbourne: Blackwells.
- Underwood, A.J., in press. Experiments in ecology and management: their logics, functions and interpretations. *Australian Journal of Ecology*.
- Underwood A.J. & Denley E.J., 1984. Paradigms, explanations and generalizations in models for the structure of intertidal communities on rocky shores. In *Ecological Communities: Conceptual Issues and the Evidence* (ed. D.R. Strong, Jr, et al.), pp. 151-180. New Jersey: Princeton University Press.
- Underwood, A.J., Denley, E.J. & Moran, M.J., 1983. Experimental analyses of the structure and dynamics of mid-shore rocky intertidal communities in New South Wales. *Oecologia*, **56**, 202-219.
- Underwood, A.J. & Jernakoff, P., 1981. Effects of interactions between algae and grazing gastropods on the structure of a low-shore intertidal algal community. *Oecologia*, **48**, 221-233.
- Underwood, A.J. & Jernakoff, P., 1984. The effects of tidal height, wave-exposure, seasonality and rock-pools on grazing and the distribution of intertidal macroalgae in New South Wales. *Journal of Experimental Marine Biology and Ecology*, **75**, 71-96.
- Whewell, W., 1847. *The Philosophy of the Inductivist Sciences*, Volume 2. London: Parker.
- Whittaker, R.H., 1956. Vegetation of the Great Smoky mountains. *Ecological Monographs*, **26**, 1-80.
- Winer, B.J., 1971. *Statistical Principles in Experimental Design*, 2nd Edition. Tokyo: McGraw-Hill Kogakusha.
- Wolcott, T.G., 1973. Physiological ecology and intertidal zonation in limpets (*Acmaea*): a critical look at 'limiting factors'. *Biological Bulletin. Marine Biological Laboratory, Woods Hole*, **145**, 389-422.
- Worster, D., 1977. *Nature's Economy: the Roots of Ecology*. San Francisco: Sierra Book Clubs.