# NOTE

# Explaining patterns in abundances of organisms on boulders: the failure of 'natural experiments'

Keith A. McGuinness

Institute of Marine Ecology, Zoology Building A 08, University of Sydney, Sydney 2006, New South Wales, Australia

ABSTRACT· In studies of sessile organisms on intertidal boulders on 2 rock platforms near Sydney, Australia, the green alga *Ulva lactuca* and the tubeworm *Galeolaria caespitosa* were usually more abundant on one shore (Cape Banks). A likely explanation for this pattern was the difference in rocktype between the 2 shores – boulders were sandstone at one shore (Cape Banks) and shale or laterite at the other (Long Reef). Experiments showed this explanation to be correct, though not complete, for *U. lactuca* but incorrect for *G. caespitosa*. Further, a common spirorbid, normally equally abundant on both shores, settled and/or survived better on shale/laterite. These results illustrate the dangers of drawing conclusions about processes based simply on observations of patterns.

The aims of ecology are often broadly expressed as the explanation of the distribution and abundance of species and the structure of natural communities (e.g. Andrewartha & Birch 1954, Collier et al. 1973). One method commonly used to test the effects of a factor on a species or community is the so-called 'natural experiment' (Connell 1972, Diamond 1986). In such 'experiments', the abundance of the species, or the diversity of the community, is compared among places or times which differ in the intensity of the factor of interest. In the simplest cases only 2 places or times are compared. For example, the effects of water depth on the macroinvertebrates living on algae could be 'tested' by comparing abundances on algae from deep and shallow sites (e.g. Edgar 1983), or the effects of sand on intertidal algae by estimating the percentage cover of the algae before and after burial (e.g. Seapy & Littler 1982).

Several authors have pointed out the weaknesses of this approach, noting that its validity rests on the tenuous assumption that the places or times differ *only* in the intensity of the selected factor (Connell 1972, Paine 1977, Underwood 1984). Further, it is always assumed that an observed difference between places in a factor (e.g. sand) causes, or exaggerates, differences between the places in the structure of the community (e.g.

abundance of algae). The possibility that complex interactions among factors confuse or obscure patterns, and reduce differences in the community among places, is rarely considered. Though these criticisms have often been made, few studies have shown them to be important in practice, and this may be the reason that some ecologists still place great reliance on the results of 'natural experiments' (e.g. Diamond 1986, Miller 1986). The aim of this note is simply to demonstrate how such 'experiments' can lead to quite erroneous conclusions.

**Patterns in the abundance of species.** Abundances of sessile species on intertidal boulders low on the shore (0.25 m above Mean Low Water) on 2 rock platforms on the east-coast of Australia, near Sydney, were estimated as part of an extensive study of the ecology of these communities (McGuinness & Underwood 1986, McGuinness 1987a, b, 1988). One of the commonest algae was the green, ephemeral *Ulva lactuca* Linn., which occupied the entire upper-surface of some boulders. Four animals common on the undersides of rocks were the barnacle *Tetraclitella purpurascens* (Wood), the tube-building polychaete *Galeolaria caespitosa* Savigny, and the spirorbids *Janua pseudocorrugata* Bush and *Pileolaria pseudomilitaris* Thiriot-Quievreux (the latter 2 species could not be distinguished reliably in the field and are referred to below as 'spirorbids').

Recruitment and subsequent abundance of these species were examined by placing initially bare boulders of different sizes at 3 sites on each shore (details in McGuinness 1987a). Ten replicate boulders were placed in each site in August 1981; another 5 replicates were added 6 mo later in February 1982. At several later times, the percentage cover of *Ulva lactuca* on the tops of rocks, and the densities of *Galeolaria caespitosa*, the spirorbids and *Tetraclitella purpurascens* on the undersides, were estimated (percentage cover using random points and densities using quadrats;
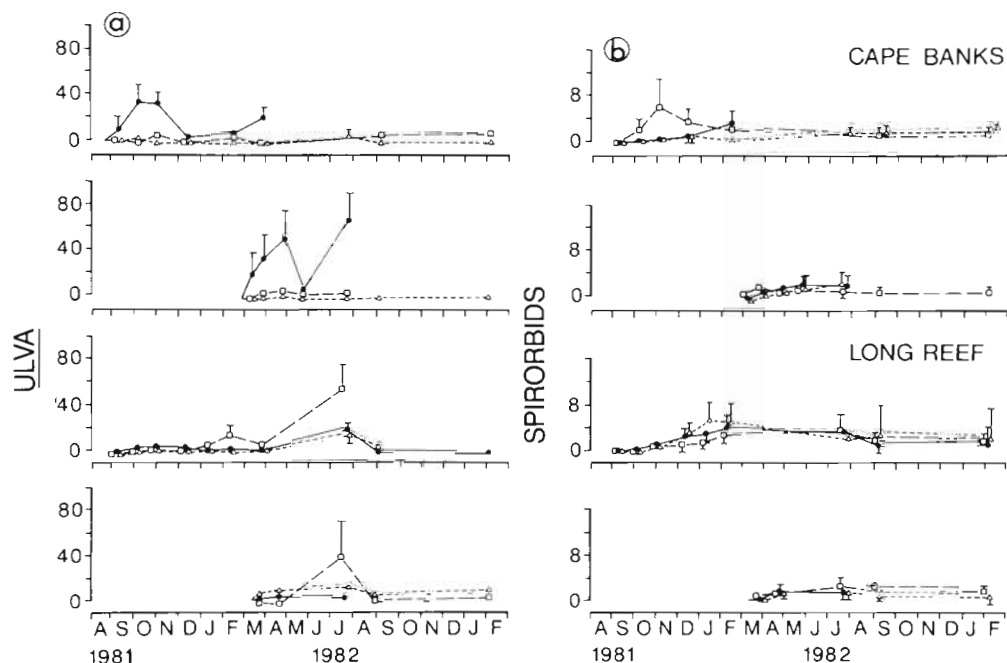
Fig. 1. Abundances of *Ulva lactuca* (a) and the 2 spirorbids (b) on intertidal boulders. Values for *U. lactuca* are mean percentage cover on the tops of the rocks. Abundances of the spirorbids are densities on average-sized (300 cm$^2$) rocks calculated from regressions of density on rock-size at each place and time. Upper graph for each taxa gives results for rocks put out in August 1981; lower, for rocks put out in February 1982. Bar: plus and minus 1 SE; uncapped bar: standard error that extends beyond the confines of the axis. Circle, square and triangle: results for the 3 different sites on each shore

Table 1. Results of analyses of the abundance of species on boulders on 2 shores (Factor A, fixed, df = 1) and at 3 sites (Factor B, fixed, df = 4) on each shore (ANOVA for *Ulva lactuca*; ANCOVA for others). Residual degrees of freedom (df) varied depending upon the number of marked rocks which could be located. The CB/LR column indicates whether abundance on boulders at Cape Banks was significantly greater (>; at $p = 0.05$), significantly less than (<) or equal to (=) that on boulders at Long Reef. The site column indicates whether there were (•), or were not (ns), significant differences among the sites on each shore (at $p = 0.05$). A dash indicates missing data

| Sample time (d) | df | Ulva | | Spirorbids | | Galeolaria | | Tetraclitella | |
|---|---|---|---|---|---|---|---|---|---|
| | | CB/LR | Site | CB/LR | Site | CB/LR | Site | CB/LR | Site |
| **Aug 1981 rocks** | | | | | | | | | |
| 20 | 54 | > | ns | > | ns | = | ns | = | ns |
| 47 | 54 | > | ns | > | • | > | ns | = | ns |
| 74 | 48 | = | • | > | • | > | ns | = | • |
| 115 | 48 | < | • | = | • | > | ns | = | ns |
| 179 | 42 | = | • | = | ns | > | ns | = | ns |
| 224 | 42 | = | • | ¬ | | – | | – | |
| 251 | 18 | > | • | ¬ | | – | | – | |
| 340 | 18 | > | ns | = | – | > | – | = | – |
| 473 | 18 | = | – | = | – | > | – | = | – |
| **Feb 1982 rocks** | | | | | | | | | |
| 223 | 18 | > | • | > | ns | = | ns | = | ns |
| 251 | 12 | > | • | = | • | = | ns | = | ns |
| 340 | 6 | = | • | = | • | > | ns | = | ns |
| 473 | 6 | = | – | = | – | > | – | = | ns |

methods in McGuinness & Underwood 1986). Animal densities were usually positively correlated with rock size (estimated as surface-area). Densities were, therefore, compared among shores and sites using analyses of

covariance (ANCOVA) to take the relation between density and rock-size into account (Huitema 1980). The percentage cover of *U. lactuca* was unrelated to rock-size, so there was no need to use ANCOVA, and
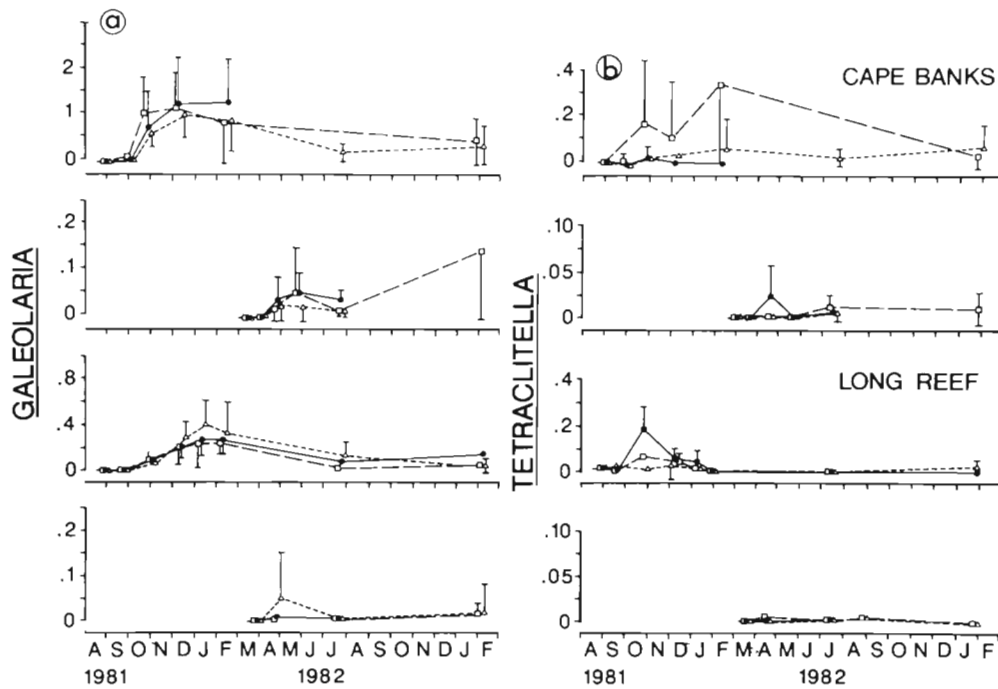
Fig. 2. Densities of *Galeolaria caespitosa* (a) and *Tetraclitella purpurascens* (b) on average-sized (300 cm²) rocks calculated from regressions of density on rock-size at each place and time. Format as for Fig. 1

Table 2. Patterns in the abundance of species, hypothesised effect of rock-type and actual effect of rock-type. The pattern in long-term sampling was generalised from Table 1. 'Hypothesis' is the hypothesised effect of rock-type on the settlement and/or survival of the species. 'Experimental pattern' compares abundances observed during the experiment on sandstone at Cape Banks with shale/laterite at Long Reef (same comparison as in the first column but made 2 yr later). 'Actual effect' was the experimentally determined effect of rock-type on the settlement and/or survival of the species

|              | Long-term pattern | Hypothesis    | Experimental pattern | Actual effect  |
|--------------|-------------------|---------------|----------------------|----------------|
| *Ulva*       | CB > LR           | Sand > shale  | CB = LR              | Sand > shale   |
| Spirorbids   | CB = LR           | Sand = shale  | CB = LR              | Sand < shale   |
| *Galeolaria* | CB > LR           | Sand > shale  | CB > LR              | Sand = shale   |
| *Tetraclitella* | CB = LR        | Sand = shale  | CB = LR              | Sand = shale   |

analyses of variance (ANOVA) were used (Winer 1971). Separate analyses were done for each sampling time.

The abundance of *Ulva lactuca* varied greatly over time (Fig. 1a). In general, percentage cover on boulders on the 2 shores was similar when small, but when mean cover was great it was much greater at Cape Banks (Fig. 1a, Table 1). In contrast, the spirorbids recruited slightly faster onto boulders at Cape Banks but reached similar densities on the 2 shores after about 100 d (Fig. 1b, Table 1). *Galeolaria caespitosa* showed a different pattern: once there was any settlement the worm was much more abundant at Cape Banks (Fig. 2a, Table 1). Finally, *Tetraclitella purpurascens* varied greatly in abundance with no difference between Cape Banks and Long Reef (Fig. 2b, Table 1). In summary, the spirorbids and *T. purpurascens* were generally similarly abundant on the 2 shores, *G. caespitosa* was much

more abundant at Cape Banks, and *U. lactuca*, when present, covered more space at Cape Banks (these patterns are summarised in Column 1 of Table 2).

**Hypotheses to explain patterns.** An obvious difference between the 2 shores which could account for the greater abundance of *Ulva lactuca* and *Galeolaria caespitosa* at Cape Banks was the type of rock present. The boulders at Cape Banks were, almost without exception, sandstone, whereas those at Long Reef were laterite or, more usually, shale. The sandstone was light in colour with large grains (0.1 to 2.0 mm in diameter), whereas laterite and shale were dark and had smaller grains (< 0.01 mm; Chalmers 1969). These features of the rock could influence the settlement and/or survival of *U. lactuca* and *G. caespitosa*. In laboratory experiments, the larvae of marine invertebrates do respond to surface-texture (review by Meadows & Campbell 1972,

Table 3. Mean abundance of sessile species on sandstone and shale/laterite rocks at Cape Banks and Long Reef after 86 and 131 d. Abundance of *Ulva lactuca* is the percentage cover on the tops of rocks; values for the other species are number of individuals $cm^{-2}$ on the undersides of rocks. A greater than (>) or less than (<) sign indicates a significant difference (at $p = 0.05$ by ANOVA; see McGuinness & Underwood 1986). Values in parentheses are standard errors ($n = 8$ at 86 d; $n = 7$ at 131 d). Note that boulders at Cape Banks were moved but not overturned by people between 86 and 131 d, killing some animals on the undersides: data for undersurface species at Cape Banks at 131 d are, therefore, unreliable

| | Cape Banks | | | Long Reef | | |
| | Sandstone | | Shale | Sandstone | | Shale |
| --- | --- | --- | --- | --- | --- | --- |
| *Ulva* | | | | | | |
| 86 d | 33.8 | > | 4.6 | 85.4 | > | 46.9 |
| | (15.7) | | (3.6) | (7.1) | | (13.6) |
| 131 d | 29.3 | > | 3.7 | 63.1 | > | 29.6 |
| | (11.8) | | (3.1) | (11.6) | | (14.0) |
| Spirorbids | | | | | | |
| 86 d | 0.224 | < | 0.574 | 0.186 | = | 0.224 |
| | (0.077) | | (0.133) | (0.083) | | (0.057) |
| 131 d | 0.255 | = | 0.196 | 0.153 | = | 0.466 |
| | (0.118) | | (0.072) | (0.070) | | (0.107) |
| *Galeolaria* | | | | | | |
| 86 d | 0.045 | = | 0.112 | 0.007 | = | 0.006 |
| | (0.017) | | (0.042) | (0.002) | | (0.002) |
| 131 d | 0.142 | = | 0.205 | 0.038 | = | 0.044 |
| | (0.041) | | (0.038) | (0.008) | | (0.014) |
| *Tetraclitella* | | | | | | |
| 86 d | 0.042 | = | 0.019 | 0.004 | = | 0.003 |
| | (0.025) | | (0.014) | (0.003) | | (0.002) |
| 131 d | 0.001 | = | 0.001 | 0.003 | = | 0.013 |
| | (0.006) | | (0.006) | (0.002) | | (0.006) |

Crisp 1976), although one field study indicated no such effect for an Australian barnacle (Caffey 1982). The colour of the rock could affect heat-retention, influencing the desiccation stress experienced by algae. *U. lactuca* does grow faster when less stressed (Underwood 1980, Underwood & Jernakoff 1984). The patterns in the abundance of *U. lactuca* and *G. caespitosa* could, therefore, be due to greater settlement and/or survival of these species on sandstone than on shale. The spirorbids and *Tetraclitella purpurascens* are, of course, assumed to be unaffected by rock-type (these hypotheses are summarised in Column 2 of Table 2).

**Tests of hypotheses.** These hypotheses were tested quite simply by placing 8 bare sandstone rocks (from Cape Banks) and 4 shale and 4 laterite rocks (from Long Reef) in the low shore area on each platform (the 4 shale and 4 laterite rocks together made 8 'Long Reef-type' rocks – there were no differences between these 2 rock-types in abundances of species; McGuinness & Underwood 1986). The rocks were put out on 27 July 1983 at Long Reef and on the following day at Cape Banks (at Site 2 in both cases). The results of this experiment were reported by McGuinness and Underwood (1986), but they concentrated on the effects of rock-type on each shore and did not discuss the differences *between* shores which are important here.

The effects of rock-type were, in some cases, quite contrary to those predicted. *Ulva lactuca* did settle and/or survive better on sandstone than on shale/laterite at both shores, and the difference in cover was substantial, averaging about 32 % (Table 3). The spirorbids were, however, significantly more abundant at Cape Banks on shale/laterite than on sandstone after 86 d (Table 3). A similar pattern seemed to be developing by 131 d at Long Reef. *Galeolaria caespitosa* and *Tetraclitella purpurascens* were not affected by rock-type (Table 3). Thus, the effects of rock-type on settlement and/or survival were as predicted for only 2 of the 4 species, *U. lactuca* and *T. purpurascens* (Column 4 in Table 2). *G. caespitosa* did not become more abundant on sandstone, as predicted, and, if anything, settled and/or survived better on shale/laterite (Table 3). Furthermore, an effect of rock-type on the spirorbids was observed (density more than doubled) when no such effect was predicted (Columns 2 and 4 in Table 2). In hindsight, some of these 'unpredicted' results probably resulted, at least in part, from the burial of rocks in sand, a process which was less common at Cape Banks than at Long Reef (because there was less sand; McGuinness 1987a). Other experiments showed that even slight burial reduced the recruitment of undersurface species, such as the spirorbids, *G. caespitosa* and

Table 4. Difference between Cape Banks and Long Reef in abundance of each species at 86 and 131 d. Columns compare results for sandstone at Cape Banks vs shale/laterite at Long Reef (i.e. the natural situation), and for sandstone vs sandstone and shale/laterite vs shale/laterite on the 2 shores. The 3 sets of comparisons are not independent and cannot be compared statistically (but see standard errors in Table 3). Abundance of *Ulva lactuca* was measured as percentage cover; abundances for animals are numbers $cm^{-2}$

|  | CB-sandstone minus LR-shale | CB-sandstone minus LR-sandstone | CB-shale minus LR-shale |
|---|---|---|---|
| **86 d** |  |  |  |
| *Ulva* | −13.1 | −51.6 | −42.3 |
| Spirorbids | 0.000 | 0.038 | 0.350 |
| *Galeolaria* | 0.039 | 0.038 | 0.106 |
| *Tetraclitella* | 0.039 | 0.039 | 0.016 |
| **131 d** |  |  |  |
| *Ulva* | −0.3 | −33.9 | −25.9 |
| Spirorbids | −0.211 | −0.102 | −0.271 |
| *Galeolaria* | 0.098 | 0.104 | 0.161 |
| *Tetraclitella* | −0.012 | −0.002 | −0.012 |

*T. purpurascens* (McGuinness 1987b). This might have counteracted the tendency of the spirorbids, and perhaps *G. caespitosa*, to settle or survive better on the type of rock naturally present at Long Reef.

It is instructive to consider what the abundances of these species on the 2 shores *might* have been if boulders on *both* were the *same* type of rock. This can be done simply by comparing the abundance of the species on sandstone at Cape Banks and shale/laterite at Long Reef (that is, the natural situation; Column 1 of Table 4) with its abundance on either sandstone at Cape Banks and Long Reef (Column 2), or shale/laterite (Column 3) at Cape Banks and Long Reef. In most cases, the difference in the abundance of the species between the shores would have been numerically *greater* if boulders on the 2 platforms were of the same rock-type (Table 4). For example, after 86 *Ulva lactuca* covered 34 % of the surface of the sandstone boulders at Cape Banks and 47 % of the shale boulders at Long Reef, a difference of 13 %. If, however, boulders on both shores were sandstone then the difference would have been 52 % (comparing cover on sandstone boulders in the experiment on the 2 shores). Likewise, if the 2 shores had shale/laterite boulders then the difference between them in cover would have been 42 %. In other words, the difference in type of rock reduced the difference in abundance between the 2 shores (though this comparison is probably not reliable for the undersurface forms at 131 d).

This is in complete contrast to the conclusions of 'natural experiments' which always assume that a difference in some factor between places causes differences between the places in the abundance of species. In the present case, the direction and magnitude of differences between Cape Banks and Long Reef probably depended on the intensity of recruitment to the 2 shores, as well as any effects of rock-type. Variations in settlement, or changes in survival for other reasons (e. g. grazing or differences in local weather), could obscure differences between shores and offset the tendency for a greater cover of *Ulva lactuca* to develop on sandstone. During the experiment, for example, the abundance of *U. lactuca* was great but similar on sandstone at Cape Banks and shale/laterite at Long Reef, in contrast to previous observations (Column 3 of Table 2; Table 3).

**Conclusions.** My results clearly demonstrate the dangers of drawing conclusions about processes based solely on interpretations of 'natural experiments': unrecognised factors can and do act to violate the underlying assumptions. It is worth noting that the importance of rock-type for the spirorbids would have gone entirely unnoticed if the experiment had not been done. There was no pattern to indicate that such an effect was present. Thus, even the lack of a correlation between, say, the abundance of a species and the intensity of some factor may not reliably indicate that the factor does not influence the abundance of the species.

This is not an isolated case. I also sampled the species-area patterns of these boulders to determine if the diversity of assemblages on rocks of different sizes varied in a fashion consistent with the effects of disturbance by overturning (which is known to be greater for small rocks; Osman 1977, Sousa 1979, McGuinness 1984). On the basis of this analysis of patterns, I concluded that disturbance by waves did not influence the number of sessile species on the undersides of rocks high on the shore at Cape Banks (McGuinness 1984), but experiments subsequently showed this inference to be false (McGuinness 1987b). A possible explanation for the discrepancy is that the disturbance may have resulted from the movement of rocks, whether or not they were overturned, and further studies showed this type of disturbance to be similar for rocks of all sizes – an unexpected result (McGuinness 1987a).

Examples of errors which might have been made, were experimental evidence not available, are becoming more common. Dayton (1973) showed that patterns consistent with 2 aspects of competition theory actually had quite different causes. Schroder & Rosenzweig (1975) tested previous ideas that 'habitat specializations, food habits, and species diversity of desert rodents . . . were molded by interspecific competition'. Their experiments on kangaroo rats revealed 'a strong element of intraspecific competition with little or no evidence of competition between the species'.

These results refute recent suggestions that 'natural experiments' are superior to manipulative experiments (Diamond 1986, Miller 1986). But, as Diamond (1986) himself notes, the inference involved in 'natural experiments' 'bears the obvious burden of demonstrating that the claimed difference rather than some other difference between the sites really caused the outcome'. He concludes that 'one can never be certain'. With manipulative experiments, however, provided that the design was appropriate and due attention was given to potential artifacts (Hulberg & Oliver 1980, Hurlbert 1984, Underwood 1986), one can be certain that the manipulation has caused the outcome. In fact, nearly all of the problems discussed by Diamond (1986) and Miller (1986) result from flaws in the design or execution of particular manipulative experiments, rather than in the logic underlying the procedure. Clearly, there are situations in which field experiments are, for one reason or another, impractical. It is, however, desirable that in such situations the tentative nature of conclusions be recognised, and due and careful consideration be given to alternative hypotheses.

## LITERATURE CITED

Andrewartha, H. G., Birch, L. C. (1954). The distribution and abundance of animals. University of Chicago Press, Chicago

Caffey, H. M. (1982). No effect of naturally-occurring rock types on settlement or survival in the intertidal barnacle, *Tesseropora rosea* (Krauss). J. exp. mar. Biol. Ecol. 53: 119–132

Chalmers, R. O. (1969). Australian rocks, minerals and gemstones. Angus & Robertson, Sydney, Australia

Collier, B. D., Cox, G. W., Johnson, A. W., Miller, P. C. (1973). Dynamic ecology. Prentice-Hall Inc., Englewood Cliffs, N. J.

Connell, J. H. (1972). Community interactions on rocky intertidal shores. Ann. Rev. Ecol. Syst. 3: 169–192

Crisp, D. J. (1976). Settlement responses in marine organisms. In: Newell, R. C. (ed.) Adaptations to environment: essays on the physiology of marine animals. Butterworths, London, p. 83–124

Dayton, P. K. (1973). Two cases of resource partitioning in an intertidal community: making the right prediction for the wrong reason. Am. Nat. 107: 662–670

Diamond, J M. (1986). Overview: laboratory experiments, field experiments, and natural experiments. In: Diamond, J. M., Case, T. J. (eds.) Community ecology. Harper & Row Publishers, New York, p. 3–22

Edgar, G. J. (1983). The ecology of south-east Tasmanian phytal animal communities. IV Factors affecting the distribution of amphithoid amphipods among algae. J. exp. mar. Biol. Ecol. 70: 205–225

Huitema, B. E. (1980). The analysis of covariance and alternatives. John Wiley and Sons, Interscience, New York, London

Hulberg, L. W., Oliver, J. S. (1980). Caging manipulations in marine soft-bottom communities: importance of animal interactions and sedentary habitat modifications. Can. J. Fish. Aquat. Sci. 37: 1130–1139

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

McGuinness, K. A. (1984). The species-area relations of the communities on intertidal boulders: testing the null hypothesis. J. Biogeogr. 11: 439–456

McGuinness, K. A. (1987a). Disturbance and organisms on boulders. I. Patterns in the environment and the community. Oecologia (Berl.) 71: 409–419

McGuinness, K. A. (1987b). Disturbance and organisms on boulders. II. Causes of patterns in diversity and abundance. Oecologia (Berl.) 71: 420–430

Mc Guinness, K. A. (1988). Short-term effects of sessile organisms on colonization of intertidal boulders. J. exp. mar. Biol. Ecol. (in press)

McGuinness, K. A., Underwood, A. J. (1986). Habitat structure and the nature of the communities on intertidal boulders. J. exp. mar. Biol. Ecol. 104: 97–123

Miller, J. C. (1986). Manipulations and interpretations in tests for competition in streams: 'controlled' vs 'natural' experiments. Oikos 47: 120–123

Meadows, P. S., Campbell, J. I. (1972). Habitat selection by aquatic invertebrates. Adv. mar. Biol. 10: 271–382

Osman, R. W. (1977). The establishment and development of a marine epifaunal community. Ecol. Monogr. 47: 37–63

Paine, R. T (1977). Controlled manipulations in the marine intertidal zone, and their contributions to ecological theory. Acad. Nat. Sci. USA, Spec. Publ. 12: 245–270

Schroder, G. D., Rosenzweig, M. L. (1975). Perturbation analysis of competition and overlap in habitat utilization between *Dipodomys ordii* and *Dipodomys merriami*. Oecologia (Berl.) 19: 9–28

Seapy, R. R., Littler, M. M. (1982). Population and species diversity fluctuations in a rocky intertidal community relative to severe aerial exposure and sediment burial. Mar. Biol. 71: 87–96

Sousa, W P. (1979). Experimental investigations of disturbance and ecological succession in a rocky intertidal algal community. Ecol. Monogr 49: 227–254

Underwood, A. J. (1980). The effects of grazing by gastropods and physical factors on the upper limits of distribution of intertidal macroalgae. Oecologia (Berl.) 46: 201–213

Underwood, A. J. (1984). The analysis of competition by field experiments. In: Anderson, D. J., Kikkawa, J. (eds.) Community ecology – pattern and process. Blackwell Scientific, Oxford, p 240–268

Underwood, A. J. (1986). Physical factors and biological interactions: the necessity and nature of ecological experiments. In: Moore, P. G., Seed, R. (eds.) The ecology of rocky coasts. Hodder and Stoughton, London, p. 372–390

Underwood, A. J., Jernakoff, P. (1984). The effects of tidal height, wave-exposure and rock pools on grazing and the distribution of intertidal macro-algae. J. exp. mar Biol. Ecol. 75: 71–96

Winer, B. J. (1971). Statistical principles in experimental design (second edition). International student edition. McGraw-Hill Kogakusha Ltd., Tokyo