bioscience
education

Research Article

Assessing the effectiveness of a computer simulation for teaching ecological experimental design

Richard Stafford[1], Anne E. Goodenough[1], Mark S. Davies[2]

[1]*Centre for Active Learning and Department of Natural and Social Sciences, University of Gloucestershire;* [2]*Faculty of Applied Sciences, University of Sunderland, Sunderland.*

Date received: 01/10/2009 *Date accepted: 08/02/2010*

Abstract

Designing manipulative ecological experiments is a complex and time-consuming process that is problematic to teach in traditional undergraduate classes. This study investigates the effectiveness of using a computer simulation — the Virtual Rocky Shore (VRS) — to facilitate rapid, student-centred learning of experimental design. We gave a series of tests to biology undergraduates to determine how well experimental design and data analysis was understood: 1) before any teaching sessions on this topic; 2) after theory sessions on experimental design; and 3) after an additional practical session using the VRS. Due to poor weather, sample sizes were small with a total of 12 students participating in all three sessions. Nevertheless, marks increased significantly between the initial and final tests (1 and 3). The variability of marks during test 2 was also significantly higher than for the other two tests. Thus some students learned experimental design effectively from the theory sessions alone, while others only understood the process after the experiential learning component of the VRS. Feedback from students on the process was mainly positive, although some students found the VRS too abstract, indicating that the use of digital learning resources may need to be supported by real experience in the field or laboratory.

Keywords: Experimental Design, Hypothesis Testing, Student-Centred Learning, Computer Simulations

Introduction

Science is founded on principles of observation, measurement, hypothesis testing and experiment. In the biological sciences, the last two are especially important to advance understanding of how organisms function and interact with one another and their environment. Training in the principles and methods of observation, measurement and the formation of hypotheses presents comparatively little difficulty and, in our experience, is achieved mostly through relatively didactic instruction, though increasingly hypothesis testing has been taught in a more interactive way (e.g. Open University, 2009). The teaching of experimental design, however, presents particular problems (Hiebert, 2007). Experimental design embodies a set of rules to be applied where independent variables are manipulated in an attempt to determine cause and effect or linkages between independent and dependent variables (Underwood, 1997). These rules are complex, varying with each particular effect and potential cause that is examined (Underwood and Denley, 1984), such that it is not possible to devise a common set of 'rules'. Much is based on common-sense principles — for example what controls, if any, are necessary — but much is based on a statistical understanding of the numbers and types of components within the study; for example, the determination of the statistical power of an experiment. Thus experimental design is a skill to be mastered rather than a body of knowledge to be taught. Accordingly, in many cases, experimental design cannot be taught in the abstract; rather its understanding is best, perhaps only, achieved through experiential learning (Steffe and Thompson, 2000; Ju *et al*., 2004).

The need for experimental design to be taught through experiential learning poses problems because such teaching, for any biological subject, is both time-consuming and costly. Even where relatively rapid results can be obtained in the laboratory, producing enough true replicates can be difficult not only because of the cost of the materials, but because of problems with inter-operator variability (often particularly high in a group of students learning skills for the first time). Thus the subject of experimental design does not lend itself to being taught in a typical university ‘practical’ teaching session of between two and four hours.

The problems with teaching experimental design in the biosciences are perhaps most obvious for ecology, where experiments are notoriously time-consuming and therefore particularly difficult to teach (Underwood, 1997). This is a pity since ecology, by virtue of its complexity, has become the subject that relies most on experimentation. Indeed, early experimental design was developed specifically to solve problems in ecology (Fisher, 1926) and ecology is still the discipline where new aspects of, and refinements to, the experimental design process are most often made.

Because of the difficulties outlined above, experimental design is often not taught in any detail, nor undertaken actively by undergraduate students in ecological disciplines, such that it is not generally encountered until postgraduate research (Easterling, 2004). Nevertheless, the processes of careful design, undertaking appropriate experiments and analysing the resultant data statistically is an important scientific skill (Easterling, 2004), which actually underpins subject-specific knowledge and improves a student’s ability to critically assess research results (Finn *et al*., 2002). There is considerable evidence that this skill is not mastered even by professional biologists: Underwood and Denley (1984) and Raffaelli and Hawkins (1996) both give examples of poor design, and resultant incorrect conclusions, in published ecological research. In most biological studies (and related research in pharmaceutical, psychological and medical disciplines), confounding factors need to be eliminated, controls and procedural controls need to be utilised, and sufficient, but not excessive, replication needs to occur to ensure results are meaningful and useful (Ruxton and Colegrave, 2006).

Rocky shore ecology has been at the forefront of ecological experimental design for many decades, largely due to the ease of access to study sites and the ease of manipulation of study organisms (Raffaelli and Hawkins, 1996; Underwood, 2000). The high rocky shore is an extremely simple ecological community — essentially consisting of snails and a photosynthetic biofilm of lichens, diatoms and bacteria on which the snails graze — and thus it lends itself well to experimental manipulation (Stafford, 2002). The movement and trophic interactions of the snails are well understood, as is their interaction with the abiotic environment in the way that they form aggregations inside crevices to reduce desiccation stress (Garrity, 1984; Stafford, 2002).

In an attempt to meet the demands of teaching experimental design, the relative simplicity of the high shore environment has been exploited in the development of the Virtual Rocky Shore (VRS), a computer simulation based on accurate peer-reviewed scientific knowledge. The simulation is derived from agent-based models of snail movement on rocky shores in Hong Kong (Stafford *et al*., 2007) and interactions between individual snails and biofilm (Stafford, 2002). As such, it is an ecologically-accurate representation of the high shore community. The simulation allows the use of two tools, one that can be used to create grazer exclusion areas by means of applying various amounts of a barrier to the shore, and one that allows the formation of crevices on the shore (identical in concept to previous experiments such as those by Williams, 1994; Davies *et al*., 1997; Mak and Williams, 1999; Stafford and Davies, 2005; Range *et al*., 2008). Data can be obtained by means of analysing virtual rock chips for chlorophyll *a,* which is used as a proxy measure for biofilm abundance (Nagakar and Williams, 1999). The tools allow students to design a range of exclusion areas and use a range of procedural controls. In this way habitat can be modified in a hypothesis-driven manner without preconceived ideas or program limitations biasing work (for example, operators receive no prompts on what size of exclusion area is most appropriate). In this way, the VRS is an open-ended learning tool designed to maximise students' independent learning (D'Avanzo, 1996). Following this pattern, the VRS also allows for a range of statistical techniques to be used to analyse the data, including one-way, repeated measure, nested and two-way ANOVA designs, as well as regression analysis.

This study evaluates the effectiveness of the open source and free to use VRS software (Stafford *et al*., 2010) in student learning of experimental design in first and second year university biosciences students in the UK.

Methods

Evaluation of the VRS took place between December 2008 and February 2009 using both first and second year undergraduate students registered on bioscience degree programmes at a single UK university. To assess knowledge and understanding of experimental design, a series of three tests (see appendix 1) on experimental design was given to each student. The order in which the students took the tests was randomised so although each student took a different test teach time, they did not all take the same test at each session, thereby eliminating the confounding factor of potential differences in the difficulty of the tests. The tests were marked (RS) and marks were verified by a second marker (AEG). Students did not identify themselves on the test papers to facilitate blind marking.

Each test was divided into three sections. Section A tested understanding of the concept of a scientific experiment in ecology. This section listed five scientific investigations and students were asked to identify which were biological experiments. The students were unaware that only two of these investigations were experiments. The section was negatively marked, giving a mark range of between -3 and 2. Section B asked students to list five improvements to a poorly designed example experiment, which, in its presented form, suffered from confounding factors and a lack of replication. Marks were awarded on a scale of 0 to 5. Section C dealt with the analysis of experimental data, and asked students how they would analyse the results of a hypothetical experiment. Marks were given between 0 and 10. Two self-evaluation questions (identical on all tests; see appendix 1) were presented at the end of each test, and each asked students to rate themselves on a scale of 0 to 10 on how well they thought they understood experimental design and how confident they thought they were at analysing results.

Tests were taken at three points in the academic year: (1) before lecture sessions on experimental design and data analysis — referred to as CTR for control group herein; (2) after these lecture sessions, referred to as EXP for after experimental theory, herein; and (3) after a practical session collecting and analysing data using the VRS referred to as VRS herein. First year or 'level 1' students had previously been taught the concepts of random sampling and sample size for ecological surveys, but had had no experience of independent practical work. They had received no instruction in statistical techniques beyond the calculation of averages and variability. In addition to the above, second year or 'level 2' students had taken a short course in statistical techniques, including one-way ANOVAs, and had attended a field course where they had designed and conducted ecological surveys (not manipulative experiments) and analysed the results.

Results were analysed using a series of two-way ANOVAs where the scores for each section of the test (as well as the total score for sections A-C) were entered as the dependent variable (4 ANOVAs in total). The students' level of study (level 1 or level 2) and the time of the test (before any relevant teaching sessions (CTR), after theory sessions on experimental design (EXP), or after using the VRS (VRS)) were both entered as fixed factors. Since the test scores used were a random subsample of the total number of those collected (it was not necessarily the same students taking the different tests on each occasion — see results), and there was no student taking the same test more than once, repeated measures analysis was not used, as, essentially, each student/test combination was independent.

Although no student feedback on the use of the VRS was explicitly collected, any relevant comments in end-of-course evaluations were noted and collated by student representatives independently of teaching staff.

Results

Due to heavy snow on one of the teaching days, only six of the level two students (out of a total group of 20) completed all taught sessions and tests. To keep the statistical analysis balanced, the marks of six tests were selected randomly from each of the other five treatment combinations (student level (1 or 2) and teaching session (CTR, EXP, or VRS) giving a total sample size = 36). Balanced ANOVA designs are considered more robust than unbalanced, even if the total sample size is smaller. This is especially true when the assumptions of homogeneity of variance are not fully met, or when post-hoc analysis is needed (Underwood, 1997), as is the case here. There was no significant difference between the level of study (first or second year) and students overall marks for sections A-C inclusive, neither was there any interaction term between level of study as the stage at which tests were taken (Table 1).

Table 1 *Summary of ANOVA and post hoc Tukey tests showing where significant differences in test marks occurred in relation to learning activities. 'CTR' before any relevant teaching sessions; 'EXP' after theory sessions on experimental design; 'VRS' after a practical session with the VRS. L=1 first year students. Significant effects are shown in bold. Cochran's test for heterogeneity of variances was non-significant for all tests. For details of sections A-C and self evaluation, see methods.*

Test	Time	Level	Interaction	Tukey post-hoc significant differences
Total of sections A-C	$F_{2,30}$ = 4.086 **p = 0.027**	$F_{1,30}$ = 0.018 p = 0.895	$F_{2,30}$ = 0.588 p = 0.562	CTR < VRS
Section A	$F_{2,30}$ = 4.181 **p = 0.025**	$F_{1,30}$ = 0.172 p = 0.681	$F_{2,30}$ = 0.043 p = 0.958	CTR < VRS
Section B	$F_{2,30}$ = 0.709 p = 0.500	$F_{1,30}$ =1.674 p = 0.206	$F_{2,30}$ = 0.314 p = 0.733	–
Section C	$F_{2,30}$ = 4.343 **p = 0.022**	$F_{1,30}$ = 1.552 p = 0.223	$F_{2,30}$ = 0.420 p = 0.661	CTR < VRS
Self-evaluation	$F_{2,30}$ = 2.862 **p = 0.072**	$F_{1,30}$ = 1.042 p = 0.316	$F_{2,30}$ = 4.503 **p = 0.020**	L=1 CTR > L=1 EXP

Significant differences in marks did occur, however, dependent on the timing of the test; marks were significantly higher after the use of the simulation (VRS) than before any relevant teaching session (CTR). There was no significant difference in marks between CTR and EXP treatments, or between EXP and VRS treatments (Table 1; Figure 1). This was due to the high variance of scores following the lecture based session alone (EXP), which was 7.8 times greater than the mean variance at the time of the CTR and VRS tests. Although Cochran's test for heterogeneity of variance was not significant (C = 0.4072, d.f. = 6, p = 0.101), thereby fulfilling the criteria for a balanced ANOVA (Underwood, 1997), the more sensitive Bartlett's test for heterogeneity of variance was highly significant (K^2 = 14.1, d.f. = 2, p = 0.001).

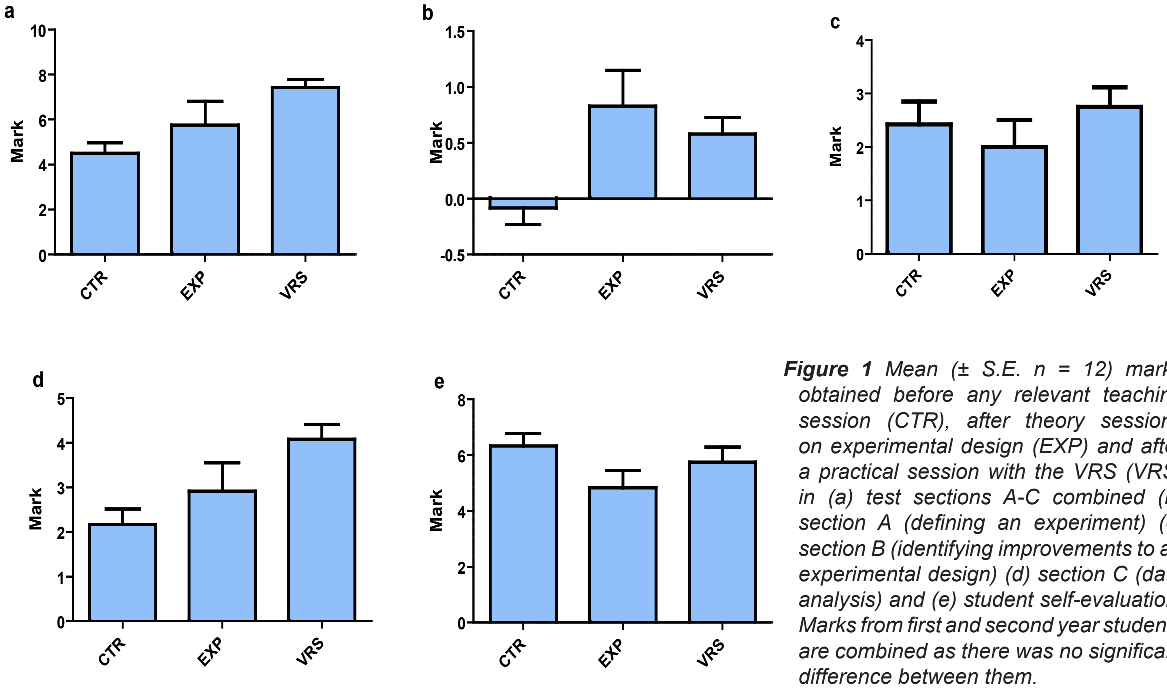

Figure 1 *Mean (± S.E. n = 12) marks obtained before any relevant teaching session (CTR), after theory sessions on experimental design (EXP) and after a practical session with the VRS (VRS) in (a) test sections A-C combined (b) section A (defining an experiment) (c) section B (identifying improvements to an experimental design) (d) section C (data analysis) and (e) student self-evaluation. Marks from first and second year students are combined as there was no significant difference between them.*

Identical patterns to that found for the combined score of sections A-C inclusive were also found in the separate analyses of test sections A and C (defining an experiment and suggesting appropriate data analysis, respectively), but no significant differences at all were found for test section B (identifying improvements to an experimental design) (Table 1; Figure 1). Although improvements in overall marks were found over time (see above), in all cases the mean marks were low, being less than 50% of the maximum mark achievable (Figure 1).

Student self-evaluations showed a significant interaction term between year of study and time period (Table 1). Tukey post-hoc comparisons showed that this was solely because level one students evaluated themselves more highly at the onset (CTR) than at after a lecture on experimental design (EXP) than did level two students.

Comments from end-of-course evaluations by second year students suggested that the VRS was useful: the representative reported, "useful, but this work should have been covered in the first year." These comments were collected from a group of 22 students by a student representative, who indicated approximately 60% of the group agreed with this statement. However, a first year student representative reporting on feedback from 18 students said over 40% considered that the "computer simulations were a bit vague and hard to see relevance, for some students."

Discussion

Owing to poor weather conditions preventing full attendance and the necessity of students attending all three sessions to take the final test, the results of this study must be considered preliminary. Nevertheless, our study clearly demonstrates that the use of digital teaching resources can aid the understanding of experimental design, at least for some students.

Interestingly, it appears that some students do not require the hands-on or experiential learning using the digital resource to grasp the concepts of experimental design, and are capable of learning this from theory-based lecture sessions. However, as indicated by the high variability of marks after only the theory-based sessions and the reduction of this variability following the practical session with the VRS, many students only gained significant understanding after using the digital resources. While many qualitative studies have focussed on the role of games and simulations in school and undergraduate student learning (e.g. Facer, 2003; Maharg and Owen, in press; Stafford, in press), this study indicates a direct link between use of digital resources and increased understanding of complex concepts in students.

The testing described here took place over three winter months, but could have been accomplished over a much shorter period, indicating just how quickly digital resources can be effective. Moreover, using real field examples would have been difficult during the time of year when this teaching had, by virtue of timetabling, to be undertaken, while the VRS — and other digital resources — effectively free teaching from the constraints imposed by weather and season.

We recognise that while digital or virtual resources can be used to good effect, they may be best, at least at the present time, regarded as supplemental to, and perhaps to be used prior to, real experimentation (Spicer and Stratford, 2001). Nonetheless, the considerable saving in time that the digital resource affords is likely to affect the depth of learning and understanding that students have about experimental design, which can increase the quality of learning achieved (Entwistle, 2009). Experiential learning of experimental design is preferable over shorter time scales, though this must be balanced with the value of students' exposure to real situations, where 'unknown unknowns' in experimental design occur, requiring students to formulate flexible experimental designs and adapt them as necessary (D'Avanzo, 1996).

Although there was no significant quantitative difference between year groups, the qualitative feedback from students does highlight some differences. Second year students, who had previously conducted field-based surveys or experiments on residential field courses, suggested they would have liked to learn the concepts earlier in their degree programme. However, some first year students, who had yet to complete field-based ecological work, could not understand the relevance of the work, stating on the group feedback that the "computer simulations were vague". This may indicate that while the use of digital teaching resources can be helpful, they cannot fully replace hands-on experience in a real field or laboratory environment. This is an important consideration given that although practical aspects of bioscience, such as field trips, are usually highly valued by students (Willmott, 2005), there has been a sustained general decrease in the amount of field experience at undergraduate level (SHEFC, 1998; Moore, 2001; Smith, 2004).

The fact that the overall marks in the tests used to assess knowledge of experimental design remained comparatively low (average mark <50%) throughout, even after post-teaching improvement, is interesting. Although the tests were based on general principles of experimental design and analysis, not directly on work taught or experienced, experiential and active learning are thought to promote deep learning and the ability to create new implications for abstract concepts (Kolb *et al.,* 2000). As such, it is surprising that the use of the VRS only improved results for some students, and the marks were consistently low through all three trials. It is possible that marks were low as the questions were challenging, and students were not prepared for, or forewarned of, the tests. However, the fact that the marks were low for all three tests does indicate that students' performance did not improve substantially as they became used to the type of questions set. However, it is important to note that improvement in marks over time could be due to familiarisation with the type of questions set in the tests as well as due to learning concepts of experimental design, and this is a clear limitation of any study that incorporates repeated student assessment.

In conclusion, in situations where there is a need for students to learn complex skills that lend themselves to experiential learning and that need is coupled with a significant resource constraint (time, money, or equipment) on the acquisition of that experience, digital teaching tools may provide effective learning. Here we demonstrated that the VRS can provide for cost-effective learning of experimental design in an experiential context, for some students. However, the sample sizes used in this study were small, and not all students appeared to learn more than just through a theory session. Further research is therefore required to fully understand the potential of digital learning resources.

Acknowledgments:

This work was supported by a grant from the Centre for Active Learning at the University of Gloucestershire and through the JISC / Higher Education Academy's Centre for Bioscience Open Educational Resources programme: The Interactive Laboratory and Fieldwork Manual for the Biosciences.

Corresponding author:

Richard Stafford, Department of Natural and Social Sciences, University of Gloucestershire, Cheltenham, GL50 4AZ Telephone: +44 (0) 1242 714681 email: rstafford@glos.ac.uk

References

D'Avanzo, C. (1996) Three ways to teach labs by inquiry: guided, open-ended, and teacher collaborative. *ESA Bulletin*, **77**, 92–93

Davies, M.S., Knowles, A.J., Edmondston, P. and Hutchinson, N. (1997) The use of a commercial insect-trapping compound to maintain grazer densities on rocky shores. *Transactions of the natural history society of Northumbria*, **57**, 185–190

Easterling, R.G. (2004) Teaching Experimental Design. *The American Statistician*, **58**, 244–252 doi:10.1198/000313004X1477

Entwistle, N. (2009) *Teaching for Understanding at University: Deep Approaches and Distinctive Ways of Thinking*. Palgrave Macmillan

Facer, K. (2003) *Computer games and learning: Why do we think it's worth talking about computer games and learning in the same breath?* NESTA Futurelab

Finn, H., Maxwell, M. and Calver, M. (2002) Why does experimentation matter in teaching ecology? *Journal of Biological Education*, **36**, 158–164

Fisher, R.A. (1926) The arrangement of field experiments. *Journal of the Ministry of Agriculture,* **33**, 503–513

Garrity, S.D. (1984) Some adaptations of gastropods to physical stress on a tropical rocky shore. *Ecology*, **65**, 559–574 doi:10.2307/1941418

Hiebert, S.M. (2007) Teaching simple experimental design to undergraduates: do your students understand the basics? *Advances in Physiological Education,* **31**, 82–92 doi:10.1152/advan.00033.2006

Ju, W. Oehlberg, L. and Leifer, L. (2004) Project-based learning for experimental design research. In *Proceedings of the International Engineering and Product Design Education Conference, 2004 Delft, The Netherlands.* http://best.berkeley.edu/~lora/Publications/IEPDE04-xPBL.pdf (accessed 28 January 2010)

Kolb, D.A., Boyatzis, R.E. and Mainemelis, C. (2000) Experiential learning theory: previous research and new directions. In *Perspectives on cognitive, learning, and thinking styles* eds Sternberg, R.J. and Zhang, L.F. pp 227–247. Mahwah, New Jersey: Lawrence Erlbaum Ass.

Maharg, P. and Owen, M. (in press) Simulations, learning and the metaverse: changing cultures in legal education. *Journal of Information, Law & Technology.*

Mak, Y.M. and Williams, G.A. (1999) Littorinids control high intertidal biofilm abundance on tropical Hong Kong rocky shores. *Journal of Experimental Marine Biology and Ecology,* **223**, 81–94 doi:10.1016/S0022-0981(98)00122-1

Moore, P.G. (2001) Developing and sharing best practice in marine-related fieldwork. *University Marine Biological Station Millport Occasional Publication No. 8*

Nagarkar, S. and Williams, G.A. (1999) Spatial and temporal variation of cyanobacteria-dominated epilithic communities on a tropical shore in Hong Kong. *Phycologia*, **38**, 385–393

Open University (2009) *SXR270 Investigative Biology, Introduction and Guide*. Open University Press

Raffaelli, D.G. and Hawkins, S.J. (1996) *Intertidal Ecology*. Chapman & Hall

Range, P., Chapman, M.G., Underwood, A.J. (2008) Field experiments with “cageless” methods to manipulate grazing gastropods on intertidal rocky shores. *Journal of Experimental Marine Biology and Ecology*, **365**, 23–30 doi:10.1016/j.jembe.2008.07.031

Ruxton, G.D. and Colegrave, N. (2006) *Experimental design for the life sciences*. Oxford University Press

SHEFC (1998) *The promotion of quality in Scottish higher education is a central concern of the Council’s quality assessment process. Quality Assessment Annual Report 1996-7*, Scottish Higher Education Funding Council

Smith, D. (2004) The issues and trends in higher education biology fieldwork. *Journal of Biological Education,* **39**, 6–10

Spicer, J.I. and Stratford, J. (2001) The virtual field trip – good but no substitute. In *Developing and sharing best practice in marine-related fieldwork* ed Moore, P.G., pp 8-14. Cumbrae, Scotland, University Marine Biological Station Millport Occasional Publication No. 8.

Stafford, R. (2002) *The role of environmental stress and physical and biological interactions on the ecology of high shore littorinids in a temperate and a tropical region.* Ph.D. Thesis. University of Sunderland

Stafford, R., Hart, A.G., Newberry, J., Catlin-Groves C.L. and Goodenough A.E. (2010) *The Virtual Rocky Shore - a tool for understanding ecological experiments*. Available at http://web.mac.com/richardstafford1/vrs (accessed 28th January 2010)

Stafford, R. (in press) Using computer simulations to teach sustainable harvesting concepts and interactions between biological populations. In *Engaging students in active learning: case studies from education, humanities and sciences*. eds Thacker, D., Healey, M. and Roberts, C. Cheltenham, UK. Cider House Press

Stafford, R. and Davies, M.S. (2005) Spatial patchiness of epilithic biofilm caused by refuge-inhabiting high shore gastropods. *Hydrobiologia,* **545**, 279–287 doi:10.1007/s10750-005-3320-5

Stafford, R., Davies, M.S. and Williams, G.A. (2007) Computer simulations of high shore littorinids predict small-scale spatial and temporal distribution patterns on real rocky shores. *Marine Ecology Progress Series*, **342**, 151–161 doi:10.3354/meps342151

Steffe, L.P. and Thompson, P. W. (2000) Teaching experiment methodology: Underlying principles and essential elements. In *Handbook of Research design in mathematics and science education,* eds Lesh R. and Kelly, A.E., pp 267- 307, Mahwah, New Jersey: Lawrence Erlbaum Ass.

Underwood, A.J. (1997) *Experiments in ecology*. Cambridge University Press

Underwood, A.J. (2000) Experimental ecology of rocky intertidal habitats: what are we learning? *Journal of Experimental Marine Biology and Ecology,* **250**, 51–76 doi:10.1016/S0022-0981(00)00179-9

Underwood, A.J. and Denley, E.J. (1984) Paradigms, explanations and generalisations in models for the structure of intertidal communities on rocky shores. In *Ecological Communities, Conceptual Issues and the Evidence*, eds Strong, D.R., Simberloff, D., Abele, L.G. and Thistle, A.B. pp. 151-180, Cambridge, MA. Princetown University Press

Williams, G.A. (1994) Grazing by high shore littorinids on a moderately exposed tropical rocky shore. In *The Malacofauna of Hong Kong and Southern China III*, ed Morton, B. pp. 379-389, Hong Kong University Press

Willmott, C. (2005) What Makes the Best Learning Experience? *Bioscience Education* **5**-c1 available at www.bioscience.heacademy.ac.uk/journal/vol5/beej-5-c1.aspx (accessed 28th January 2010)