

*Research Article***Lectures Replaced by Prescribed Reading with Frequent Assessment: Enhanced Student Performance in Animal Physiology**

Peter F. D. Chevins

*School of Life Sciences, Keele University, Staffordshire ST5 5BG**Date received: 05/08/2004**Date accepted: 13/10/2004***Abstract**

*This article describes a study of the effects of partial replacement of lectures with a system of prescribed reading, supported by weekly objective testing in a second year animal physiology module. Formative tests with feedback within 24 hours were followed a week later with summative tests on the same material, utilising a proportion of the same questions. Over a three year period, students' reported study hours during the module increased significantly over their normal study time. Summative test scores were always significantly higher than the preceding formative test scores. More study hours were devoted to the summatively assessed material. The topics taught in this way (nerve, muscle and endocrine physiology) were previously the subject of the poorest examination results, and were chosen least as examination essay topics. During the three years of the study, examination results improved for nerve physiology (marginal statistical significance) but the proportion of students choosing the different topics for their examination answers was unchanged. Students approved of the learning regime, but only when supported by some lectures.*

**Keywords:** prescribed reading, lectures, assessment, feedback**Introduction**

In spite of the range of teaching methods available, lecturing remains the most widely used. Whilst many practitioners accept that the lecture is not an ideal way of imparting knowledge or understanding (Bligh, 1998; Ramsden, 1992; Gibbs, 1992) others support the method for its efficiency (Beard and Hartley, 1984). Undoubtedly the effectiveness of lectures depends on a complex of factors, including what they are intended to achieve, and the individual's lecturing style (Steinart and Snell, 1999). The work described here resulted from reflection on my own experience in a second year Animal Physiology module which had led me to believe that a few weeks after a block of lectures students had learned little. This was based on their almost complete lack of recall of lecture content, during small-group revision tutorials, some four weeks after the lectures. Several students stated that they had not yet looked at their notes as it was not yet time for "revision". Brown *et al* (2003) illustrate the patterns of study of students on a course with no assignments as compared with 3 in-course assignments, showing that study is concentrated in weeks immediately before assessment. This module had one in-course assignment and it was evident that after preparing this in-course essay from the prescribed text (Randall *et al*, 1997) and other sources, student's recall in tutorials was much better.

Students' study of lecture material seemed to be all but absent during the course, but concentrated almost entirely in the pre-examination revision period. Plainly they were actively engaging with the text during preparation of the in-course assignment, but not with the lectures. The research showing that students actively involved in learning gain more than those who are passive (for example Butler, 1992 and Feden, 1994) is undisputed. It appeared the students had been passive, save for note-taking, during the lectures, and had not mentally engaged with the material they were noting. The strategy I decided to adopt to counter this problem was to prescribe a weekly programme of reading from the set textbook to replace lectures at least in part. The risk was that, as with their lecture notes, they would leave the reading until their revision period. To circumvent this I planned weekly objective-style tests of the prescribed reading, first in formative mode with rapid feedback, then the following week a summative test on the same material. Glover (2004) states that "assessment has an overwhelming influence on what, how, and how much students study" and that "one of the most powerful influences on student achievement is feedback".

I planned to evaluate this method, comparing it with the conventional lectures, by using the end of module examination results as measure of students' achievement. The effect of feedback on the formative tests was to be investigated by comparing summative with formative test scores. Questionnaires would also be used to discover the changes, if any, in student study patterns and their opinions of the changed learning regime.

## **Method**

The experiment was conducted for three successive years, in the autumn semesters of 2000, 2001 and 2002, in a module of selected topics in comparative animal physiology taught by three lecturers. Prior to the experiment, the teaching format was two lectures per week for 12 weeks (with a break for a reading week at week 7) together with associated practical classes. In addition, an in-course essay designed for examination practice was set in the second half of the module and returned to students during a tutorial shortly before the vacation. For the years of the experiment, weeks 1 - 6 were delivered in the usual lecture format, covering the topics of water and ion balance and excretion, cardiovascular and respiratory physiology.

## ***Prescribed reading***

The remainder of the module was presented in the form of prescribed reading from the set text (Randall *et al*, 1997) introduced in a single lecture at the end of week 6 immediately before the reading week. For this week reading of 3 chapters was set and the students were informed that in the first lecture period of week 8 they would be tested on the material in a formative objective-style test. This introductory lecture was used both to explain the teaching method, and to indicate the most important sections of the reading. Some parts of the chapters would not be tested, and each student received a handout giving these details. The most difficult concepts in the week's study material were briefly explained in the remainder of the lecture. For the rest of the module (weeks 8 - 12) the first lecture period each week was used for an

objective test, and the second – the following day – for feedback and details of the next week's reading. After the initial reading week, each week's study material was restricted to the majority of a single chapter. Topics covered in this way were nerve and muscle physiology and endocrinology.

### ***Formative and summative testing***

Tests consisted mainly of multiple response questions, together with a number of "two statement" questions: a format in which students are required to answer whether one or both of the two statements are true (and if so which) and if both are true, does the second follow logically from the first. Tests were designed to be capable of completion comfortably within the 50-minutes available.

Tests were paper-based, manually marked overnight and returned to students the following day with detailed feedback. The feedback was not personalised, but dealt with each possible correct and incorrect answer. After the first test in week 8, which was formative only, each test consisted of a summative section re-testing the previous week's work, and a formative section testing the new material. The summative section contained about 30% of the same questions as the previous week's formative test, and the students knew this would be the case. They were given a record sheet (proforma) on which to record their answers, and were invited to check their marks (which were published immediately) as soon as they received their correct answers and feedback. A few errors in marking were corrected in this way. In the penultimate week, no new work was set, and in the final week the test was summative only. Each of the three summative tests contributed 4% to the module assessment: 12% in total.

### ***Evaluation***

In the first year of the experiment the mean formative test scores were compared with the following summative test scores on a weekly basis to determine whether the feedback had any positive effect on the subsequent performance.

At the end of each test students were asked to complete a short questionnaire concerning the number of hours they had devoted to reading work for the formative and summative tests. After the final test they were asked their opinions of the new method of study, and the quantity of work required. They were also asked about their study patterns and study time during lecture-based teaching, whether they had part-time jobs, and if so, how many hours per week. In each year of the experiment, student self-reported study hours were analysed and compared with their reports of time spent reading or studying under the normal lecture regime.

The end of module examination consisted of a compulsory short-answer section based mainly on the students' laboratory work, and an essay section that assessed the theoretical (lecture or reading) material. In the essay section, students chose two questions from four, which were normally grouped into four topics: water and ion balance and excretion (1) cardiovascular and respiratory physiology (2) nerve and muscle (3) and

endocrinology (4). This grouping of topics was not explicit, but would be evident to students who studied past examination papers.

The effect of the method on examination results was evaluated in two ways. First, the marks for examination essays on the conventionally taught topics (water and ion balance and excretion and cardiovascular and respiratory physiology) were compared with those for nerve and muscle and endocrinology which had been studied using the new method. Second, the marks for these latter topics were compared for the years 1997-1999 (before the experiment) with 2000-2002 (the years of the experiment). For the first year of the experiment, the essay questions in the examination were identical with those set in 1998, thus providing a comparison free of possible variations in difficulty of the examination questions. For all year comparisons, marks were normalised to control for variations in the performance of the student cohort. This was achieved by expressing individual exam essay marks as a percentage of the mean class mark for the module in that year.

Finally, as the students had a free choice of examination essay topics, the numbers choosing to answer questions on the topics taught by the new method were compared for the years prior to and during the experiment.

## Results

### **Formative and summative test scores**

In the objective tests of the year 2000, students scored significantly better on average in each summative test than the preceding formative one (Table 1).

**Table 1** Mean objective test scores in the first year of the study (2000)

Topic	Mean formative %	Mean summative %	Paired t-test (N=24)
Nerve physiology	49.9+/-1.8 (SE)	69.1+/-3.0 (SE)	p<0.001
Muscle physiology	38.5+/-3.1(SE)	58.0+/-12.7 (SE)	p<0.001
Endocrine physiology	45.7+/-1.8(SE)	65.2+/-2.4 (SE)	p<0.001

### **Study time**

Study hours reported in 2000 are shown in Table 2. Of the 23 responders to the first test, 16 reported that they had studied for a longer period (mean 3.8 hours, range 0.5 - 11 hours) than they would normally have spent on this material (mean 2.1 hours; range 0.5 - 6.5). Twelve thought the amount of work set just right: the remaining 11 thought it too much. Five admitted to not normally looking at the lecture material "in the days immediately following the lecture". The mean time declared as normal during lecture-based teaching was 1.2 study hours per lecture (range 0.5 - 3 hours, but 15 out of 19 in the range 1-2 hours).

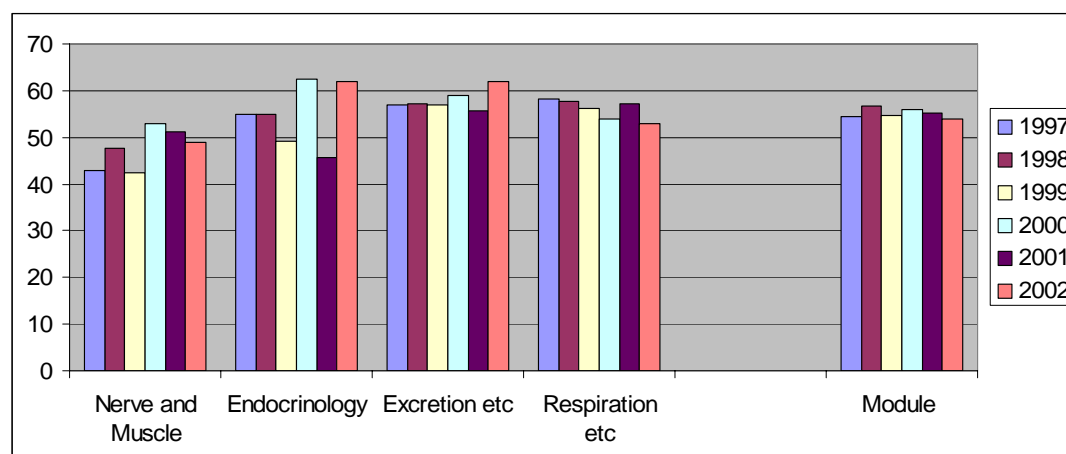
Nine of the 24 had part-time jobs averaging 10.2 hours per week (range 3 - 16 hours). Having a job did not affect study time: 5 of the 9 job-holders spent more than average time on this material and of the rest almost all normally spent about average time on lecture study; they divided 5:4 in thinking this work too much: just right.

**Table 2** Reported Study hours in year 1 of the study (2000)

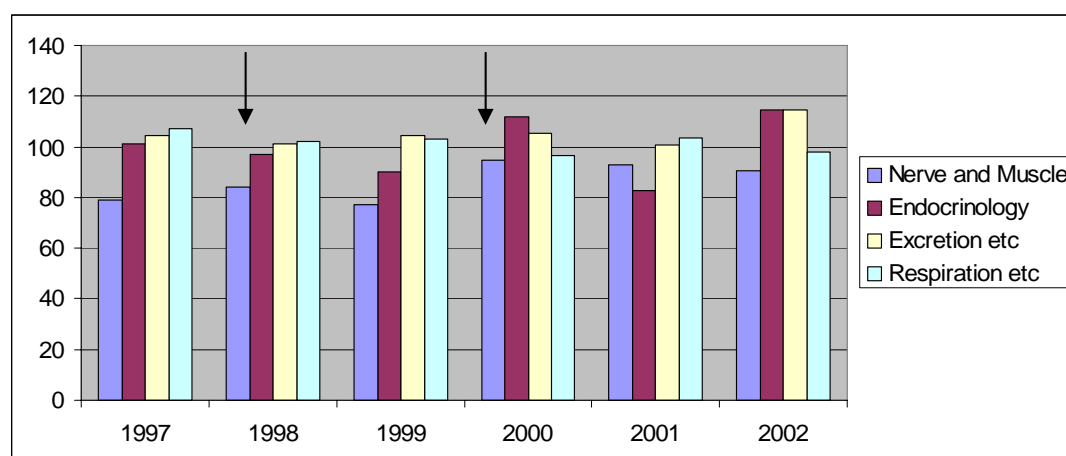
	New material for formative test (mean + range)	Feedback and revision for summative test (mean + range)	Total study hours for week (mean + range)
Week 7	3.8 (0.5 - 11.0)		3.8 (0.5 - 11.0)
Week 8	2.0 (0.5 - 5.0)	3.1 (0.5 - 5.0)	5.1 (2.3 - 15.0)
Week 9	1.3 (0 - 5.0)	3.1 (0.5 - 6.0)	4.6 (1.0 - 10.0)
Week 10		3.4 (1.5 - 10.0)	3.4 (1.5 - 10.0)
All weeks			15.3 (6.0 - 41.0)

### Examination performance

Analysis of raw examination results before and during the years of the experiment is shown in Figure 1. The same results normalised for overall module performance are given in Figure 2, which shows an improvement in performance in the nerve-muscle section of the examination closely approaching statistical significance (Student's  $t = 0.06$ ).



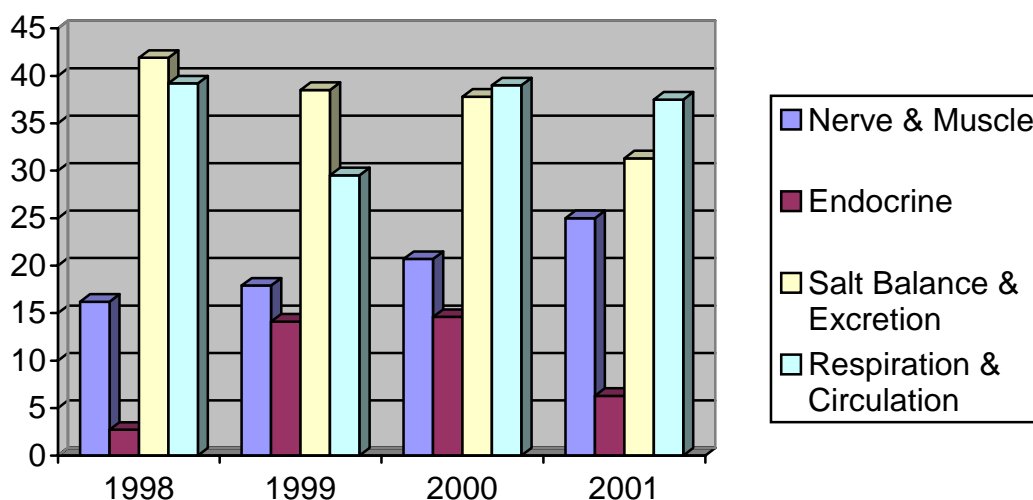
**Figure 1** Year comparison of examination marks (percent) by topic and overall module marks (1997-2002)



**Figure 2** Mean Exam Topic Marks Expressed as a Percentage of Overall Module Score. Arrows indicate Student's  $t=0.06$  for comparison of Nerve & Muscle scores, 1998 and 2000

### Choice of Examination Topics

Students had a free choice of two from four examination topics throughout the study years. Figure 3 shows the proportions choosing from the topics conventionally taught and topics studied by the new method (Nerve, Muscle and Endocrinology). There were no significant differences between the two years before compared with the first two years of the experiment (2000 and 2001).



**Figure 3** Percentage of students choosing examination question topics for two years (1998-99) before and during (2000-01) the experiment

### Student opinion

At the end of the last test, the opinion of the 2002-03 cohort was surveyed and the results of the 44 responders (55 in the class) are shown below (Table 3).

**Table 3** Results of Cohort Survey. The scale for responses was from Strongly Agree (1) to Strongly Disagree (5) with 3 as Neutral

Statement	Mean Response
I think I have learned more from this style of teaching than with a conventional lecture-based approach.	2.33
I have done more reading of the text book than I would normally have done.	2.07
I would support a plan to present this entire module in this way, with 30% of the module marks derived from the 6 objective tests.	2.61
The questions are too factual and encourage rote learning rather than deep understanding.	3.00
The more difficult questions of the two statement type should be worth twice as many marks.	2.39
Knowing that some questions from the formative tests will appear in the following summative test is an incentive to study the feedback.	1.45
Having the formative tests available on-line, with immediate feedback, and available to revisit would be an improvement.	1.86
I like this style of learning.	2.32

## Discussion

Excellence in student learning, according to Cross (1996) is promoted by three conditions: (1) high expectations, (2) student participation and involvement, and (3) assessment and feedback. This work addresses conditions (2) and (3). Evidence for improvement in student participation in the present work comes from their reported increased study time both in their early questionnaire responses on their number of study hours, and in their qualitative responses at the end of the experiment.

Students' study hours varied little week by week, in spite of week 7 being a study week, with no classes and extra reading prescribed. The experiment resulted in more time reported per week as spent studying (mean 3.8 hours) than reported for the conventional lecture presentation mode (2.1 hours per week). The students were thus reporting an increased and even spread of effort, which was one of the objects of the experiment, and described as a the "ideal" outcome by Glover (2004)

The rapid availability of feedback (Cross's third condition) is seen as a crucial element of the current work. Numerous other authors have emphasised its importance, among them Black and William (1998) who review a large number of earlier studies, concluding that "there is a body of firm evidence that formative assessment is an essential component of classroom work, and that its development can raise standards of achievement". Students' in-course objective test scores did improve significantly between the first (formative) test and the summative test of the same material a week later. Although this appears to indicate a positive effect of the feedback, it should be noted that students consistently spent more time on the work about to be summatively tested than on the new study material that would be formatively tested on the same occasion (Table 2). This should come as no surprise, given the strategic nature of student learning (Glover, 2004). The cohort of 2000 – the first year of the study, and for which detailed study hours and test results are reported – were unusually few (24) in number. However, the same patterns were evident in the subsequent two years when numbers in the class were 52 and 55 respectively. Some students wrote in their questionnaire responses that they had insufficient time to study both, so chose the summatively assessed material.

Feedback is often seen as inextricably linked with measurement of performance: Ramprasad (1983) defines it as "information about the actual level of a system parameter which is used to alter the gap in some way." Often, feedback is written specifically in response to an individual student's performance and to correct his/her errors. In this study, however, the same feedback was given to all students regardless of their performance, and in the formative tests they knew that the score did not count towards module grade. It was not tailored to their errors or correct answers, and this may be an advantage. Taras (2002) considers that "any appearance of a grade ... before the students have had the opportunity to interiorise feedback .... interferes with the assimilation and understanding of this feedback". There can of course be no guarantee that students read, much less internalise the

feedback, but a particular feature of this experiment was that students knew that they would be summatively tested on some of the material from the previous formative test. This was included as an extra incentive to learn: Taras (2002) considers that true formative feedback is not complete until “students have produced an equivalent piece of work....” (in this case the subsequent summative assessment) “...where the issues have been addressed and remedied, that is to say, until true learning has taken place”. In their questionnaire responses, students expressed the strong opinion that knowing that some questions from the formative tests would appear in the following summative test was an incentive to study the feedback – in fact this was their most emphatic response. The fact that some pointed out errors in marking is also evidence that the feedback had been studied.

Gibbs and Simpson (2003) list eleven conditions under which assessment supports learning. The first two, that the assessment tasks capture sufficient time and effort, and that they distribute the effort evenly across topics and weeks, are fulfilled by this method. So is the third, that they engage students productively and actively, though the fourth, that they communicate high standards and expectations may not be met by all the objective-style questions. The feedback was detailed, frequent and timely, fulfilling conditions five and six, and also focussed on learning rather than the score or on the students themselves (condition seven). The feedback was linked to the purpose of the assignment (the reading) and was believed to be understandable, fulfilling condition eight and probably nine. It was certainly received by the students, and the evidence of improved summative scores and their questionnaires indicates that they attended to it and acted upon it, thus the final two conditions are also fulfilled.

Analysis of examination performance on nerve physiology over the 3 years prior to the experiment (1997-1999) shows mean scores averaged 44.3% compared with 56.2% and 53.4% respectively for the conventionally taught topics of excretion and respiration. In the years of the experiment (2000 - 2002) the nerve physiology scores rose to 53.0%, 51.2% and 49% whilst their comparators remained more or less constant (Figure 1). At least two confounding variables need to be taken into account in interpreting this result; differences in the abilities of year cohorts of students, and differences in the examination questions. The former was controlled by expressing the examination marks for each topic as a percentage of the module marks (Figure 2). Prior to the experiment, the module mark comprised the examination (75%) an in-course assessment (25%). During the experiment the ratio of examination to in-course assessment remained the same, but the 25% in-course assessment was changed from being derived entirely from a single essay (before the experiment) to 12% from the summative tests and 13% from the essay.

Difficulty of the examination was eliminated as a variable in the first year of the study only (2000) by using an examination paper identical to the one used in 1998. Comparison of the examination marks as a percentage of the module mark for nerve physiology between these two years showed an increase from 84.3% to 94.6%, bordering on statistical significance (*Student's*

$t = 0.06$ ; see arrows in Figure 2). The topic of endocrinology was also presented as prescribed reading during the course of the experiment, but although these results too appear to have improved, no meaningful analysis is possible because of the low number of students choosing to answer questions on these topics (Figure 3). Perhaps the closest published work to that reported here is the comparison of the effectiveness of lecturing with programmed learning was carried out by Kromfrey and Purdom (1995) with 95 American education students. Their programmed learning consisted of a self-instructional booklet incorporating test exercises with answer keys and self-checking post-test at the end of each section. They found no difference in achievement, using multiple choice tests administered before, immediately after and four months after the study. Their three tests used identical questions, but presented in a different order. The numerous differences between their study and the present work no doubt account for the difference in result. Caution should therefore be applied in generalising the current results to other methods, subjects and students.

Biology students often find some of the concepts involved in nerve function difficult, particularly if they have not also studied physical and/or chemical sciences. This was reflected not only in the relatively poor examination marks (Figure 1) but also in the low proportion of students choosing to answer questions on this topic (Figure 3). Whilst the examination results improved as a result of the change in teaching regime, the proportion of students choosing to answer examination questions on these topics did not.

The initial idea and plan for this experiment intended the complete elimination of lectures. In the event, the need to introduce the reading material changed the intended practice. Once the reading material had been outlined (and in the opening session the method explained), the remainder of a single 50-minute lecture period was used to explain the most important or difficult points in the week's study material. However, because of illness of the investigator in the second year (2001) a replacement lecturer was employed, and she spent no face-to-face time with the students at all. In that year their questionnaire response indicated strong disapproval for this aspect of the method. Many teachers accept the idea that there may be better ways than lectures to encourage active involvement of students in learning (e.g. Steinart and Snell 1999). However, Isaacs (1994) reports that in a survey of 100 Australian university academic staff, they were overwhelmingly of the opinion that "the presence of a teacher in the classroom interacting with the students is essential to the educational experience".

Small-group teaching that could induce more active learning by providing increased interaction between students and teachers requires more time and resources for are not usually available (Schwartz, 1989) in the modern context. The current method saves little, if any time – although computerisation of the testing and feedback might make substantial savings after the initial outlay. In their questionnaire responses, the students expressed a wish to have the formative tests available on-line, with immediate feedback, and available to revisit. It is intended to introduce this as soon as the software and infrastructure is available, but this will remove the

compulsory element from the formative testing, and it remains to be seen to what extent students take advantage of the facility, and what effect this has on their examination performance and the effectiveness of the method. In its present form, the method of prescribed reading supported by minimal lecturing and frequent objective testing with rapid feedback produced an improvement in study pattern and examination performance. The “seen” element of the testing – that is the promise of some of the formative questions appearing in the summative tests – may be an important element of the method. Questionnaire responses support this view, suggesting that this appeals to the students’ inclination to apply their efforts strategically towards obtaining maximum module marks with minimum time and effort.

### **Acknowledgements**

This study would not have been possible without the help and co-operation of my lecturing colleagues, Drs. D. Hoole, G. Hamilton and J. Hulme.

The author gratefully acknowledges financial support from an LTSN Biosciences Teaching Development Grant, and from Keele University Teaching Enhancement Fund.

**Corresponding author:** Dr. Peter F. D. Chevins, School of Life Sciences, Keele University, Staffordshire ST5 5BG. Telephone: 01782 583024; Fax: 01782 583516; Email: [p.f.d.chevins@biol.keele.ac.uk](mailto:p.f.d.chevins@biol.keele.ac.uk)

## References

- Beard, R. M. and Hartley, J. (1984) *Teaching and Learning in Higher Education*, 4th Edition. London, UK: Harper and Row
- Black, P. and William, D. (1998) Inside the Black Box. Raising standards through classroom assessment. *Phi Delta Kaplan* **80**, (2) 139-144
- Bligh, D. (1998) *What's the Use of Lectures?* 5<sup>th</sup> Edition, Intellect
- Brown, E., Gibbs, G. and Glover, C. (2003) Evaluation tools for investigating the impact of assessment regimes on student learning. *Bioscience Education E-journal*, volume 2 available at <http://www.bioscience.heacademy.ac.uk/journal/vol2/beej-2-5> (accessed 29 July 2004)
- Butler, J. A. (1992) Use of teacher methods within the lecture format. *Medical Teacher* **14**, 11-25
- Cross, K. P. (1996) Improving teaching and learning through classroom assessment and classroom research, in Gibbs, G. (ed.) *Improving student learning: using research to improve student learning*. (Oxford, Oxford Centre for Staff Development), pp 3-10. cited by Taras M (2002) Using Assessment for Learning and Learning from Assessment. *Assessment & Evaluation in Higher Education* **27** (6), 501-510
- Feden, P.D. (1994) About instruction: powerful new strategies worth knowing. *Educational Horizons* **73**, 18-24
- Gibbs, G. (1992) *Lecturing to More Students*, Oxford, UK: Polytechnics and Colleges Funding Council
- Gibbs, G. and Simpson, C. (2003) Does your assessment support your students' learning? *Learning and Teaching in Higher Education* **1**,1 Available at <http://www.open.ac.uk/science/fdtl/documents/lit-review.pdf> (accessed 29 July 2004)
- Glover, C. (2004) Report of research carried out at Sheffield Hallam University for the Formative Assessment in Science Teaching Project (FAST) available at [www.open.ac.uk/science/fdtl/documents/SHUfinal\\_report.pdf](http://www.open.ac.uk/science/fdtl/documents/SHUfinal_report.pdf) (accessed 29 July 2004)
- Isaacs, G. (1994) Lecturing practices and note-taking purposes. *Studies in Higher Education* **19** (2), 203-217
- Kromfrey, J.D. and Purdom, D. M. (1995): A comparison of lecture, co-operative learning and programmed instruction at college level. *Studies in Higher Education* **20** (3), 341-349

Ramprasad A. (1983) On the definition of feedback. *Behavioural Science* **28**, 21-38

Ramsden, P. (1992) *Learning to Teach in Higher Education*. London, UK: Routledge

Randall, D., Burggren, W. and French, K. (1997) *Eckert Animal Physiology: Mechanisms and Adaptations*, 4th Edition. New York, USA: W. H. Freeman

Schwartz, P. (1989) Active small group learning with a large group in a lecture theatre: a practical example. *Medical Teacher* **11**, 81-86

Steinart, Y. and Snell, L.S. (1999) Interactive lecturing: strategies for increasing participation in large group presentations. *Medical Teacher* **21**, (1), 37-42

Taras, M. (2002) Using Assessment for Learning and Learning from Assessment. *Assessment & Evaluation in Higher Education* **27** (6), 501-510