Learning in lectures
Huxham, Mark

Postprint / Postprint
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Nutzungsbedingungen:
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under the "PEER Licence Agreement". For more Information regarding the PEER-project see: http://www.peerproject.eu This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-230927
Learning in lectures
Do ‘interactive windows’ help?

MARK HUXHAM Napier University, Edinburgh, UK

ABSTRACT Many educational development resources recommend making conventional lectures more interactive. However, there is little firm evidence supporting either the acceptability (to students) or efficacy of doing so. This research examined the use of short ‘interactive windows’ (discussions and problem-solving exercises) in first year evolution lectures delivered to between 73 and 126 students over five years. Semi-structured evaluations of the teaching, involving more than 500 responses, identified the interactive nature of the lectures as the single most popular feature of the sessions. The division of the year class into two separate groups allowed the opportunistic testing of how interactive windows influenced learning about discrete problems within each lecture. Two short problem-solving or discussion sessions were devised for each lecture; one of these sessions was taught interactively to the first student group, the second was taught interactively to the second group. Comparing test scores achieved in questions addressing these paired problems showed strong evidence for a generally weak, positive influence of the interactive windows on recall and learning.

KEYWORDS: buzz groups, evaluation, interaction, lectures, recall

Introduction

Lecturing is under attack. Abundant evidence supports the conclusion that lectures are a poor way of stimulating thought and of changing attitudes (see extensive review in Bligh, 1998: 269–89). Even the relatively trivial educational goal of imparting information is not achieved much better through lectures than through other common teaching methods, such as seminars (Bligh, 1998). Lectures are often unpopular with students, especially those in advanced years of study, who demonstrate their feelings by not attending them. Maloney and Lally (1998), for example, recorded an absentee rate of 40 percent among third-year students, and Sander et al. (2000) found formal lectures were ranked amongst the least favoured
teaching methods by their sample of psychology, medical and business studies students. Whilst research into the relationship between attendance at lectures and academic achievement usually reports a positive correlation, the relationship is often surprisingly weak and may be better explained by confounding factors such as motivation, rather than by any direct educational value of strict and regular attendance (Hammen and Kelland, 1994; St. Clair, 1999).

Given this evidence, it is not surprising that criticism of the traditional, didactic lecture as a teaching method is virtually ubiquitous in modern educational development literature. Challenging the dominance of lectures is a staple in most induction courses for new academics. Despite all this, ‘lecture’ slots still fill university timetables, and it seems that the simple transmission of information remains the predominant mode of teaching within these slots (Lammers and Murphy, 2002; Van Dijk et al., 2001). The prevalence of lectures probably reflects the formidable forces of economic efficiency, institutional inertia and personal habit. Lectures are thus likely to remain a major part of traditional Higher Education for the foreseeable future, regardless of the arguments against them. Given this, a plethora of sources are at hand to advise the lecturer on how to make the best of a bad (or at least sub-optimal) job, and to improve the learning experience within the lecture. Most of these emphasize the importance of limiting the time spent simply imparting information, and of encouraging thinking through interactive and reflective exercises. For example, Bligh (1998: 65) declares:

> It behoves lecturers to lecture less, to convince students of the intellectual aims of their courses, and to create opportunities, in lessons and outside, in which thinking can flourish.

Such opportunities for the flourishing of thinking often consist of short ‘interactive windows’, opened within the lecture to allow the fresh air of discussion and thought into the somnolent atmosphere. A frequently recommended tool is that of the small ‘buzz group’, in which students are asked to spend a few minutes discussing a question or solving a problem with a few of their peers. This flexible method can be applied in most teaching situations. It is appropriate even within large lectures in tiered lecture halls, where less structured forms of interaction, such as asking questions, may be difficult (Alien et al., 1999). It is the method used in the current study (although I will use the term ‘interactive window’ to mean any relatively short activity, within a traditional lecture, designed to encourage interaction and stimulate thought).

Whilst common sense and limited empirical evidence support the idea that opening these ‘windows’ should enhance learning, there are also arguments for keeping them shut. These include:
• Loss of teaching time – Lammers and Murphy (2002) found that lecturers who used interactive techniques had more periods within their classes when no one was involved in a learning activity, perhaps because of delays and difficulties in shifting between teaching styles.

• Reduction in content – introducing interaction is likely to reduce the total amount of factual content that can be covered in a session, perhaps dramatically. Murray and Brightman (1996) recommend ‘the holiday rule’: ‘for a successful holiday, halve your packing and double your spending money; for successful interactive lectures, halve your material and double your time’. This may be a particularly important consideration for some groups of students. For example, Merrill (2001) found that obtaining concise information for essays and examinations was a key objective of attendance at lectures for part-time students with little study time.

• Reduction in the accuracy of transmission – asking students to work on problems and derive their own solutions may increase the risk of them leaving a lecture with incorrect solutions.

• Student resistance and perception – some studies (e.g. Van Dijk et al., 2001) report a preference amongst students for ‘traditional’ lectures over those with interactive elements. Lake (2001) found that student perceptions of course and lecture effectiveness were lower for interactive compared with lecture sessions, even though student performance was higher in the former.

• Loss of control – teaching methods that reduce a lecturer’s control over a class, such as using peer discussions, may be avoided by teachers (Van Dijk et al., 1999). This could imply a lack of confidence on the part of the teacher, which could undermine his or her efficiency were it to become clear to the students.

There are thus potential disadvantages in making lectures more interactive. These, along with the documented resistance of some academics to changing lecturing style (e.g. Van Dijk et al., 1999) and the considerable institutional and personal inertia that may have to be overcome to do so, illustrate the need for more research on whether interactive windows really help learning. Research also needs to establish which of the numerous possible interactive tools is most effective. It is not sufficient for educational developers to encourage the opening of interactive windows on the basis of common sense or educational theory alone; empirical evidence is needed.

So research needs to answer (a) do interactive windows help and (b) which types, if any, are most effective? Answering the first of these is, in principle, straightforward; answering the second requires much more thought. To address the first question, studies need to compare the learning
of students in lectures with and without interactive windows. However, there are considerable logistical and statistical barriers to overcome in achieving this. Because of the many influences on learning in lectures, large sample sizes are likely to be needed to detect any effect. Note that in this case, the lectures, not the students, will usually be the replicates; because students within a given class are not independent of each other, it will generally be incorrect to regard their individual efforts as single replicates. This implies that many lectures must be given in both treatment and control groups. A single individual is unlikely to be able to give all these lectures him or herself. So different teachers may be involved, introducing further sources of variation into the putative study and therefore implying even greater replication. Answering the second question involves all these difficulties, and includes additional complications introduced by the many possible techniques and the many ways in which they can be used. Given these challenges, along with the notorious tendency of students to behave as autonomous and unpredictable human beings (rather than compliant research subjects), it is perhaps not surprising that much remains to be discovered about interaction in lectures.

The present study has less ambitious goals than those discussed above. Its purposes are twofold. First, to explore what support there is for interactive windows amongst students and second, to examine whether there is evidence for interactive windows enhancing learning about specific problems and topics. As a relatively small and opportunistic study, it does not address the question of whether interaction can enhance learning in lecture sessions as a whole.

**Methods**

Interactive windows were introduced into a first-year lecture series on evolution in 1998, and have been used every year since (giving five years in total). The module contains one lecture a week, for 11 weeks, and is delivered to a student group which has ranged in size from 73 to 126, and in age from 17 to over 40. Each lecture contains a minimum of two interactive windows. These consist of short problem-solving exercises or discussion points. 'Buzz groups' of two or three students are given three to six minutes to discuss the problem; the correct answers are explained in class immediately after the interactive session. Two typical examples of problems are given below:

Human genes can be transferred into bacteria, and the bacteria will successfully 'read' the genes and produce the protein that is coded by them. What does this tell us about the genetic code, and how does this provide evidence for evolution?
The coefficient of relatedness between mother and offspring is 0.5. What is its value between grandmother and granddaughter?

**Evaluation**

Lectures are evaluated by using a quick, semi-structured, qualitative method after the fourth lecture (this timing allows students to become accustomed to my teaching style, but also gives time to adjust the style and content in the light of feedback received). All students are asked to write anonymous, short answers to the following three questions: (a) what do you like about these lectures? (b) what do you dislike about these lectures? (c) what would you like to see changed?

**Testing the effects of interactive windows**

Most of the people studying this module are full-time students, but a group of approximately 15–20 each year are part-time students, studying one day a week on placement from industry. From 1998–2000, the two groups were taught together. In 2001, a change in timetabling separated the two groups. In 2001 and 2002, I therefore delivered each lecture twice, on different days of the same week, to the two different student groups. This made the assessment of the effects of interactive windows statistically possible. Two interactive windows in each lecture (excluding the first, introductory lecture) were refined, as far as possible, to be of the same length, difficulty and type (e.g. discussion point or numerical problem-solving exercise). One window in each lecture was randomly assigned to full-time students, and one to part-timers. During lectures, each student group was asked to participate in the window assigned to them. The window selected for the other group was ‘left closed’; that is, the question was posed on an overhead, and I explained its solution in a traditional, didactic manner without any student involvement. Hence each student group was assigned ten ‘open’ windows (one per lecture), each paired with a ‘closed’ one taught in the same lecture (see Figure 1).

Attentiveness varies during lectures, with items at the beginning and end most likely to be recalled (Holen and Oaster, 1976). It was therefore important in this study to control for the effects of timing. It was also necessary to avoid any bias created by one student group having been randomly assigned easier interactive windows. To avoid these confounders, windows were made as comparable as possible and were given, where possible, at roughly the same point in the lecture. They were also swapped between groups in 2002, with the ‘open’ windows assigned to part-timers in 2001 given to full-timers in 2002 and vice-versa.

Recall and understanding were tested in three ways. First, a quick,
anonymous multiple-choice test was given at the end of the sixth lecture in both years, with no prior warning given to the students. This consisted of short questions, each one specific to a window opened in either full or part-time sessions. Second, the examination at the end of the module (three months after the first interactive windows were introduced) includes seven compulsory short-answer questions. As far as possible (without, for example, repeating questions from a previous year or missing important learning outcomes) these were made relevant to a particular interactive window, and the appropriate paired question (referring to the other topic covered by a window in the relevant lecture) was included. Third, the final examination also includes two elective essay questions on evolution. When students attempted these, their answers were examined for the inclusion of all possible examples and principles covered by the windows. Two marks were given for each example given which was in the correct context and was correctly understood; one mark was given for an example without full explanation or evidence that it was fully understood. All marks were summed for each possible example within the full-time and part-time groups, giving two frequency distributions. Although group identity was known when marking the class tests, all examination questions were marked blind, without knowledge of which group a student belonged to.

Statistical analyses were performed using chi-squared tests. For each type of question (the multiple choice class tests, short answer and essay exam questions) frequency distributions of correct answers for full and part-time groups were compared. In each case, $2 \times 2$ contingency tables were
constructed using student type (full or part-time) as the rows, and a question along with its relevant pair as the columns; cells consisted of the summed total for all correct answers. Where more than one correct answer was given in a test (as was the case, for example, in the essay questions) all marks were summed into their relevant categories, to produce $2 \times 2$ tables; this avoided any problems with low sample size in any individual chi-square bin. For example, suppose that in an essay question the part-time group scored a summed total of 20 marks by referring to topics and examples covered by the interactive windows that were 'opened' for them during the semester, whilst they scored a total of 15 referring to topics covered by 'closed' windows. Suppose also that the full-time group scored summed totals of 80 for their 'open' topics (those that were 'closed' for the part-timers) and 60 for their closed topics (the 'open' ones for the part-timers) in the same question. The resulting contingency table would be as shown in Table 1.

Because chi-squared tests depend on proportions, results of these analyses are not confounded by any differences in ability between the two student groups. A total of 12 tests were performed; in this analysis, individual students are not regarded as replicates. Rather, each different window examined within a given year is equivalent to a repeat experiment on the same groups of students.

Results

Evaluation

A total of 515 responses were collected over the five years. Answers to the question 'what do you like about these lectures?' were coded into nine categories (see Figure 2); chosen to best fit the responses received. The category ‘interaction’ was used for comments such as:

- I like the way you set problems and walk around the class.
- Getting the students involved in discussion makes it more interesting.
'Interaction' was the most frequently reported positive aspect of the lectures, followed by 'explanation' and 'interesting content' (Figure 2). Only one student in the five years mentioned interaction in response to the question 'what do you not like about these lectures'?

Effectiveness of interactive windows

The answers of 120 full-time students were compared to those of 18 part-timers in 2001. Equivalent class sizes in 2002 were 64 and 14.

Results of the chi-squared tests are given in Table 2, which also shows whether the expected trend was found in each individual test. That is, regardless of the statistical significance of the test, was there a higher proportion of correct answers in the expected categories? Figure 3 shows an example distribution.

None of the tests in 2001 was significant (although answers to essay question 1 approached significance at the 5% level). However, all of them showed the expected trend. Two comparisons in 2002 gave significant results, with five of the six tests showing the expected trend.

A one-tailed binomial test, for goodness-of-fit to a binomial distribution
assuming that there is an equal probability for each comparison to show the expected or the unexpected trend, gave a significant result (P = 0.0032).

**Discussion**

The current study posed two questions: are interactive windows popular with students, and can they enhance recall and understanding? The results support a positive response to both of these. Semi-structured student evaluations identified ‘interaction’ as the single most highly-rated feature of the lectures. Student answers to class tests and examination questions showed a significant trend towards enhanced performance in topics covered during interactive windows.

Entwhistle et al. (2000), in their review of literature on student evaluation, identify ‘the three Es’ as being those factors most favoured by students as important for good teaching: enthusiasm, empathy and explanation. All of these categories appear in my feedback, but were mentioned less frequently than interaction. Of course, the current study provides no grounds for comparing between teachers or for evaluating ‘good teaching’ in general. Students were asked only to identify what features they liked

---

**Table 2** Results of 12 chi-squared comparisons conducted on marks from class tests, short answer (no choice) examination questions and elective essay examination questions

<table>
<thead>
<tr>
<th>Test</th>
<th>P value</th>
<th>Expected trend?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2001 results</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class test</td>
<td>0.33</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (short answer 1)</td>
<td>0.90</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (short answer 2)</td>
<td>0.18</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (short answer 3)</td>
<td>0.85</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (essay answer 1)</td>
<td>0.06</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (essay answer 2)</td>
<td>0.50</td>
<td>✓</td>
</tr>
<tr>
<td><strong>2002 results</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class test</td>
<td>0.009</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (short answer 1)</td>
<td>0.900</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (short answer 2)</td>
<td>0.800</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (short answer 3)</td>
<td>0.150</td>
<td>X</td>
</tr>
<tr>
<td>Exam (essay answer 1)</td>
<td>0.026</td>
<td>✓</td>
</tr>
<tr>
<td>Exam (essay answer 2)</td>
<td>0.137</td>
<td>✓</td>
</tr>
</tbody>
</table>

*Note.* In each case, d.f. = 1. P values are given, along with whether any differences, regardless of significance, were in the expected direction.
about the evolution lectures; had the teaching been different (for example, had it been more or less enthusiastic) then the ranking of categories may also have differed. However, it is clear that the students considered the interactive windows to have been a positive aspect of the teaching. This is important in itself, since students do not always welcome interaction. Although some studies of interaction concur with the present one in finding that students prefer more interactive methods (Goldfinch 1996; Sander et al., 2000; Tam et al., 1993), others found a negative response. For example Van Dijk et al. (1998) report students preferring traditional to ‘activating’ lectures. The possible reasons for this difference are many, and could include the different type of student body (engineering students) and different expectations on the part of students and staff. Lake (2001)

Figure 3 Example distribution of correct responses to a paired comparison

Note. ‘Part time window’ refers to the question taught interactively to the part-timers. This is paired with a topic covered in the same lecture, but taught interactively to the full-time group. In this case, although both groups of students performed better answering the topic covered interactively in the part-time lecture, the proportion of correct answers (as a percentage of all the correct answers they gave in both questions) given by part-timers to this question was greater. Thus, although the difference was not significant, the trend was in the expected direction.
compared performance and perceptions in two groups of physical therapy students given either all traditional lectures or fewer lectures supplemented by hour-long discussion groups. Despite achieving higher performances in tests, the group exposed to the more interactive teaching method believed that they had learned less, and rated their teacher less highly, compared with the group taught by lectures alone. Goodwin et al. (1991) also found that students with 50 percent discussion groups felt they had learned less compared with students in lecture-only courses, despite their comparable performance. It is possible that student resistance to interactive methods may be related to age and life experience, with mature students more confident in discussion and more likely to appreciate interactive teaching (Merrill, 2001). Lake (2001) suggests that courses early in a curriculum should limit the interactive teaching to 25 percent, and steadily increase this percentage in later years, in order to overcome the problem of students perceiving this form of teaching as less effective. Similarly, Murray and Brightman (1996) suggest that interactive teaching is likely to be more successful in later years, after a firmer grounding in a subject has been achieved. This is an argument for combining interaction with traditional methods in the same sessions in earlier years, as was done in this study. By incorporating windows into traditional lectures, no distinctions are made between ‘lectures’ and ‘discussions’ and student perceptions of what constitutes expected university teaching may be more easily met. Evaluations from both part-time and full-time students in 2001 and 2002 (when their responses were separate) rated interaction as the most important element, despite the suggestion that mature students might differ from school leavers in their perceptions of interactive methods (Merrill, 2001).

Student evaluations, on their own, do not provide sufficient grounds for changing teaching practice. The phenomenon of ‘seduction’ (deliberately courting good evaluation, regardless of teaching efficiency) is well recognized (see Bligh, 1998: 181), and what students want may not be what is pedagogically best. So evidence that a teaching practice works to enhance learning is essential, especially when it is likely to involve some disadvantages, such as the loss of time that may be associated with interactive windows. The contrasts found here, between responses to topics taught using interactive windows and traditional didactic methods, provide strong evidence of a generally weak, positive effect. Most of the P values shown in Table 1 are far from significant, reflecting the often small differences found between the groups. But the chances of finding so many differences in the expected direction, in the absence of any real effect of interaction, are less than one in 300.

There is evidence that interactive windows had different effects on performance in short answer and essay type questions in the examination; significant differences were found in essay questions in year 2, and the
differences approached significance in this kind of question in year 1. This may suggest a rather subtle effect of interaction on recall, one which could be labelled a ‘Palomar Moment’. In ‘Mr Palomar’ by Italo Calvino, the eponymous hero is a connoisseur of cheese. Despite an encyclopaedic knowledge of, and gastronome’s enthusiasm for, obscure and delicious cheeses, he inadvertently requests the blandest mass-market product when put under pressure and forced to choose quickly, because this product has been most reinforced by advertising and comes quickest to mind. It is possible that, when given a choice, students will recall the example or topic reinforced by interactive study. In contrast, when given no choice, as in the short answer questions, the effects of interactive windows may be less, since the student has either revised and understood the topic or has not.

The most significant contrast occurred in the 2002 class test. Stronger effects of interactive windows might be expected in the class tests, since they were conducted in the sixth lecture, a maximum of five weeks (and a minimum of half an hour) since teaching the topics covered in the tests. In contrast, the examination was sat three months after the start of the module. Previous work (Hollingsworth, 1995) has reported a rapid decline in the efficacy of interactive methods in enhancing recall, hence it might be expected that the examination questions would reveal a smaller difference between groups. However, the non-significant result in the 2001 class test suggests a more complex picture than one of simple decline in recall with time, although this may well be a factor.

The present study addresses the hypothesis that interactive windows of the type used here can enhance learning of those topics discussed during the windows. It cannot test hypotheses about the effects of interaction on learning in a lecture as a whole. This is because there were no control lectures without interaction, and there are good reasons to believe that interactive windows may influence learning even after they are shut. For example, Bligh (1998: 57) shows that asking questions in class can affect the heart rates (and hence presumably the attention) even of students to whom the questions are not addressed, and that the effect can last for some time after the question is asked. Similarly, Ruhl and Suritsky (1995) demonstrate significant enhancement in student note-taking and recall as a result of inserting three, two-minute pauses into a lecture. No learning activities were prescribed during these pauses – they were effectively rests. They affected the recall of the lecture as a whole, rather than any specific topics within it.

It is therefore quite possible that the interactive windows used in the present study affected the learning of topics covered in traditional ways. This would have reduced any contrasts found between the paired interactive and non-interactive topics, and the conclusion that interaction can nevertheless have positive affects is therefore a conservative one.
Previous studies on the effects of interaction in lectures have given mixed results. Van Dijk et al. (2001) studied learning in three groups of students. One was given a 90-minute lecture taught using an electronic Interactive Voting System (IVS). The second listened to the same lecture, with IVS and also with peer instruction (similar to my interactive windows). The third group acted as a control, with no interaction. They recorded a significant reduction in learning in the IVS group, compared with the control; this result serves to reinforce the importance of conducting empirical research, rather than assuming interaction – high technology or otherwise – will enhance learning. There was a small, non-significant improvement in test scores in the combined IVS/peer instruction group. Tam et al. (1993) used interactive handouts, and reported no improvement in learning compared with a control group. In contrast, Hollingsworth (1995) found a significant effect of short interactive discussions on performance in a test completed immediately after the lecture, but not in a test ten days after. These studies used comparisons between only one, or very few, sets of experimental and control lectures. If the effects of interaction are rather weak and variable, as in the present study, then significant differences are unlikely to emerge over short time periods, and mixed results are to be expected.

In conclusion, the present study found that interactive windows opened in traditional lectures were popular with students. The high rating for interaction was consistent in all years and was shown in both full-time and part-time classes, despite previous suggestions that these different groups of students may differ in their responses to interactive methods. Interactive windows had a generally small, but measurable, effect on recall and understanding. There are reasons to believe that these conclusions are conservative (because interaction may enhance learning in the lecture as a whole, and not only in the topics covered interactively, and because my sample sizes – in particular for the part-time group – were small). Many of the potential disadvantages of interactive methods, such as overcoming student expectations and loss of time and content, are mitigated by incorporating small interactive sessions into traditional lectures, rather than replacing lectures entirely. Hence the costs of incorporating interaction in this way are likely to be small. Given this, along with the enthusiastic response of students to the interactive windows in this study, these results support the inclusion of interactive windows in lectures.

Acknowledgements
This work was possible only because of the goodwill of the hundreds of students who participated. The manuscript was improved after helpful comments from Shirley Earl, Judy Goldfinch and two anonymous referees.
References


Biographical note

MARK HUXHAM is senior lecturer and teaching fellow in ecology at Napier University. He is interested in finding simple ways of enhancing student learning and the student experience within the confines of mass higher education. He specializes in marine ecology and ecosystem function, teaches on undergraduate and postgraduate courses and really believes the cliché that good research should enhance good teaching, and vice versa.

Address: School of Life Sciences, Napier University, Edinburgh EH10 5DT, UK. [email: m.huxham@napier.ac.uk]