

Article

Teaching More by Lecturing Less

Jennifer K. Knight and William B. Wood,

Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309

Received June 23, 2005; Revised September 1, 2005; Accepted September 2, 2005

Monitoring Editor: Eric Chudler

We carried out an experiment to determine whether student learning gains in a large, traditionally taught, upper-division lecture course in developmental biology could be increased by partially changing to a more interactive classroom format. In two successive semesters, we presented the same course syllabus using different teaching styles: in fall 2003, the traditional lecture format; and in spring 2004, decreased lecturing and addition of student participation and cooperative problem solving during class time, including frequent in-class assessment of understanding. We used performance on pretests and posttests, and on homework problems to estimate and compare student learning gains between the two semesters. Our results indicated significantly higher learning gains and better conceptual understanding in the more interactive course. To assess reproducibility of these effects, we repeated the interactive course in spring 2005 with similar results. Our findings parallel results of similar teaching-style comparisons made in other disciplines. On the basis of this evidence, we propose a general model for teaching large biology courses that incorporates interactive engagement and cooperative work in place of some lecturing, while retaining course content by demanding greater student responsibility for learning outside of class.

Keywords: undergraduate students, developmental biology, peer instruction, just-in-time teaching, concept maps

INTRODUCTION

Thirty years ago, the future success of biology students might have been predictable by the amount of factual knowledge they had accumulated in their college courses. Today, there is much more information to learn, but the increasingly easy accessibility of facts on the Internet is making long-term memorization of details less and less important. Students who go on to biology-related careers after college will be required to apply conceptual knowledge to problem solving rather than simply to know many facts, and they will probably be asked to work as members of a team, rather than individually. Therefore, teaching for conceptual understanding and analytical skills while encouraging collaborative activities makes increasing sense in undergraduate courses.

There is now a great deal of evidence that lecturing is a relatively ineffective pedagogical tool for promoting conceptual understanding. Some of this evidence is general,

showing that learners at all levels gain meaningful understanding of concepts primarily through active engagement with and application of new information, not by passive listening to verbal presentations (reviewed in National Research Council, 1999). More specific evidence, primarily from university courses in physics, shows that students learn substantially more from active inquiry-based activities and problem solving than from listening to lectures (Beichner and Saul, 2003; Hake, 1998; discussed further below). Nevertheless, many university faculty who are comfortable with their lecture courses remain unconvinced that more interactive teaching will lead to increased student learning, or that interactive teaching is even feasible in large classes. Colleagues we have talked with are also concerned that the time and effort required for course revision would be prohibitive, that their students would learn less content, that outcomes could not be reliably assessed in any case, and that such changes would take students and faculty alike out of their current comfort zones (see Allen and Tanner, this issue).

To address the validity of these concerns, we carried out an experiment in “scientific teaching” (Handelsman *et al.*, 2004) in a large upper-level Developmental Biology course, in

DOI: 10.1187/05-06-0082

Address correspondence to: Jennifer K. Knight (knight@colorado.edu).

which the same two instructors (the authors), teaching the same syllabus, tested the effect of two different teaching styles on student learning gains in successive semesters. The results we present here indicate that even a moderate shift toward more interactive and cooperative learning in class can result in significantly higher student learning gains than achieved using a standard lecture format. We also discuss the impact of changing the course format on student attitudes, some of the practical implications of course reform, and implications of these findings for the future design of large biology courses.

EXPERIMENTAL DESIGN

Context of the Course

In our department, Molecular, Cellular and Developmental Biology (MCDB) at the University of Colorado, Boulder, about 50% of declared majors initially self-identify as premedical. Majors are required to take a sequence of five core courses, four of which have an associated but separate laboratory course. Developmental Biology is the last of these courses, preceded by Introduction to Molecular and Cell Biology, Genetics, Cell Biology, and Molecular Biology. Therefore, students taking the Developmental Biology course are juniors or seniors who should have acquired a substantial knowledge base in these fields of biology. All the core courses are large, ranging from about 400 students in the introductory course to about 70 in Developmental Biology. Because these courses are required, they include students with varying interest levels in the material. In all the courses that precede Developmental Biology, the format is lecture based, and students are required to learn large amounts of factual information. The average MCDB student is thus accustomed to memorizing material and to working independently and competitively to achieve a desired grade. Our best students have acquired a deep grasp of the subject matter from previous courses; the majority, however, have not. Their lack of conceptual understanding becomes particularly evident during the Developmental Biology course, in which we expect them to apply what they have been taught in their previous core courses about genetics, molecular biology, and cell biology.

Structure of the Course

Developmental Biology, with an accompanying laboratory as a separate course, has been taught in a similar format for about 25 yr in MCDB. It was initially taught as three 50-min classes per week with five high-stakes (summative) exams (four "midterm" exams on current material and a cumulative final exam), which together determined 100% of a student's grade. Several years ago, we changed the classes to two 75-min lectures per week, and we added homework problem assignments worth 20% of the grade, basing the remaining 80% on three midterm exams and a cumulative final exam. The course we used in this study as our "traditional lecture" control course, taught in the fall semester of 2003 (F'03), had this basic format. Lectures, in a standard lecture hall with fixed seating, were divided evenly between the two instructors. The lectures involved no interactive in-class work, although we did occasionally pose questions to the class. Although we attempted to make students comfortable about asking their own questions, they did so only rarely. We

Table 1. Demographics of the F'03 traditional and the S'04 interactive classes

	F'03	S'04
Female	65%	58%
Seniors	60%	70%
Minority	23%	18%
Total students	73	73

did not encourage students to work together on homework problems, although we did not explicitly forbid collaboration.

The spring semester 2004 course (S'04), into which we introduced several interactive features, also met twice a week (Tuesday and Thursday) for 75 min in a standard lecture hall. The same textbook was used for assigned readings (Gilbert, 2003), and student demographics in the two courses were similar (Table 1). However, as described below, the S'04 course differed substantially in several respects, including a teaching staff expanded to include undergraduate learning assistants (LAs), various learning activities during classes in place of some lecture time, collaborative work in student groups, and substantially increased use of in-class, ongoing (formative) assessment and group discussion. Both the F'03 and S'04 courses, as well as the spring 2005 (S'05) course described below, were taught by the authors. A prerequisite for each of these courses was a concurrent Developmental Biology Laboratory course, in which sections of ~20 students could interact in a more informal setting. One of us (J.K.) taught this course in each of the three semesters, and the structure and organization of this course did not change during this time.

Teaching Staff. As in the control F'03 course, each instructor in S'04 conducted one-half of the 30 class periods, but the other was almost always present and often participated in discussions. Four graduate teaching assistants (TAs) were assigned to help with the laboratory course and as graders in the lecture course, but they were also asked to attend classes and help to facilitate group work as described below. In addition, we employed four undergraduate LAs, chosen from among students who had performed well in the fall traditional course. The LAs were supported by a National Science Foundation STEM-TP grant to several University of Colorado science departments and the School of Education, which was intended to promote the training of undergraduates interested in possible K-12 teaching careers (see Acknowledgments). The primary role of LAs was to facilitate group work, both in class and during the first hour of each laboratory session, where the students worked on weekly homework assignments.

Classroom Activities. In the S'04 course, we assigned students to groups of three or four at the beginning of the semester and asked them to sit with and work with these groups during both lecture and lab classes. Although we still lectured for 60%–70% of class time, we interspersed periods of lecturing with multiple-choice, in-class questions (ICQs), similar to the ConcepTests described by Mazur (1996) and

Table 2. Assessments and percent contributions to grade for F'03 and S'04 classes

Assessments	Contribution to grade	
	F'03	S'04
Exams	68%	40%
Homework problems	32%	30%
In-class participation and group work	0%	30%

designed to assess conceptual understanding of the lecture content. Students answered these questions individually and anonymously using an audience response system (“clickers”), which then displayed a histogram of the answers to the class (Wood, 2004). If there was significant disagreement among the answers, we asked students to discuss the alternatives within their groups and then revote. At other times, we posed a more general question to the class and asked students to discuss with their group members before we solicited responses from several groups. On four occasions, we devoted an entire class period to discussing a recent developmental biology journal article, which we had posted for students to read prior to class with a specific set of questions to answer. Finally, we asked students to spend the first hour of each period in the accompanying laboratory course working on the current homework assignment collaboratively with their group members. As described further below, we assessed student performance based on four problem sets, participation in class (clicker questions and discussion), two in-class exams, and one cumulative final exam (Table 2).

Group Organization and Group Work. As mentioned above, there was no explicitly encouraged group work in the control F'03 class. In the interactive S'04 class, groups were organized as described by Beichner and Saul (2003). We assigned students to groups on the basis of their grades in prerequisite MCDB courses (this method was not disclosed to the students), such that each group included one excellent (“A” range), two medium (“B” range), and one low (“C” range) performer. We also attempted to balance gender and ethnic ratios to minimize possible feelings of isolation among female, male, or minority students. Assuming that group work would be unfamiliar and perhaps uncomfortable for many students, we explained at the start of the semester why working in a group could be valuable both for learning purposes and as training for future job situations, and we gave pointers on managing group dynamics. As an incentive for groups to work effectively together, we added five percentage points to the exam grades of each student in a group if the group as a whole achieved an 80% average on the exam. Group work both in class and during laboratory sessions was facilitated by LAs, as described above.

Assessment

Pretests and Posttests and Homework Problems. For both classes, we administered a multiple-choice pretest during the first week and then readministered the same test as embedded questions in the final exam, to provide a rough

measure of student learning gains during the semester. All students received the same number of points for completing the pretest, which otherwise had no effect on their grade. Their outgoing performance on posttest questions comprised less than 2% of their final grade (10 out of 500 points). The test was composed of 15 multiple-choice questions designed to test understanding of core concepts in development, a few of which they could have learned from previous courses (see Appendix A). When we repeated the interactive course in spring 2005 (described below under Results), only 12 of the pretest and posttest questions were the same as in the previous two semesters. To facilitate comparisons between the three semesters, only the 12 common questions were used in the data analysis for the tables and figures presented.

The questions on the pretest and posttest were similar to many that we asked during the semester and on the midterm exams, and thus were representative of the knowledge and understanding required in the course. To determine whether students were interpreting the pretest and posttest questions correctly, we conducted student interviews in spring 2005. All students in the S'05 class were invited to interview after the course, and six students were interviewed (including at least one who had achieved each of the grades “A,” “B,” and “C” in the course). In the interview, we asked students to think out loud as they worked through each of the questions, and to verbalize why they selected or eliminated certain answers. Our intent was to verify that students understood what each question was asking. All six students correctly interpreted each of the 12 questions we used to compare normalized learning gains. The results of these sample interviews suggest that at least for most students, our pretests and posttests provided a valid assessment of conceptual understanding for the concepts tested.

In addition to the pretest and posttest, we repeated 19 of the homework problems in the F'03 and S'04 semesters, so that student performance on these questions could be compared. Results based on the problems are less comparable than those based on the pretests and posttests, because problems were, for the most part, solved individually in the F'03 course and collaboratively in the S'04 course.

In-Class Formative Assessment. In both the F'03 and S'04 classes, we asked students at several times during the semester to give us feedback on how various aspects of the class were aiding their learning. Otherwise, however, there was no formative assessment in the F'03 course beyond the midterm exams. By contrast, in the S'04 course, we carried out formative assessment during each class period using clicker responses to ICQs. Students could receive up to five points per class period for correct answers, but we also gave partial credit up to 2½ points for incorrect answers in order to encourage participation. In general, we used these questions both to gauge the understanding level of the class, and as opportunities for discussion and peer instruction if more than 30% of the class selected an incorrect answer (see Appendix B for examples of ICQs).

RESULTS

Overall Performance in the Classes

All MCDB core courses traditionally have been graded on a curve. We decided to maintain that practice for these courses

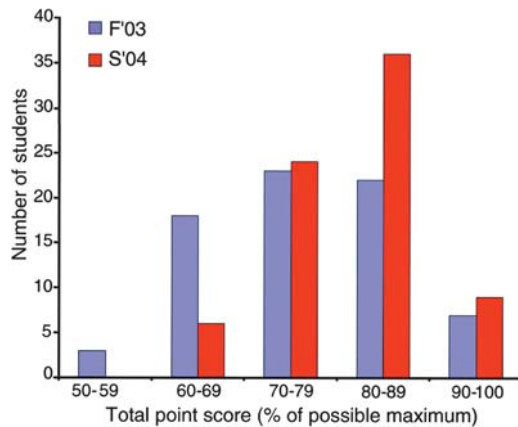


Figure 1. Final course point distributions (% of possible maximum) in traditional (F'03, blue) and interactive (S'04, red) classes. The number of students achieving a final score is shown for five ranges of scores.

as well, although we subsequently realized that grading on a curve discourages many students from working collaboratively. More students in the interactive course achieved a grade of "B" (51%) than in the traditional course (37%) (Table 3). Because of the curve, there were actually fewer students who achieved an "A" grade in the interactive course, although the overall distribution of point totals was higher (Figure 1).

As shown in Table 4, the average performance on exams and problem sets differed only slightly between the two semesters. Although the questions on the F'03 and S'04 exams were similar, they were not identical. Reflecting the differences in emphasis of the classes, the exams in S'04 were designed to test more conceptual and less factual knowledge than those in F'03. Because the exams were substantially different, we cannot make meaningful comparisons of exam performances between the two semesters. Several students in both semesters (nine in F'03, 11 in S'04) would have received below "C-" if their grades had been based solely on exam performance. In F'03, grades of these students were improved by problem set performance, and in S'04 by problem set performance and class participation (class participation grades in S'04 averaged 86%, which accounted for about 30% of the grade). We suspect that the in-class exercises also helped prepare students to perform better on the posttest, as described further below.

Normalized Learning Gains Were Significantly Higher for the Interactive Class

The average performance and standard deviation on the pretest were not significantly different in the two semesters:

Table 3. Percent of students receiving final grades of A, B, and C

	A	B	C
Traditional (fall '03)	23%	37%	38%
Interactive (spring '04)	20%	51%	28%

traditional, 34% ($\pm 12\%$); interactive, 31% ($\pm 12\%$), indicating that the incoming students were equally well prepared. However, the average performance on the posttest was significantly higher in the interactive course, by 9 percentage points ($p = .001$, two tailed t -test).

The most compelling support for superiority of the interactive approach came from comparisons of normalized learning gains calculated from pretest and posttest scores in the traditional and interactive classes (Table 4). Normalized learning gain is defined as the actual gain divided by the possible gain, expressed as a percentage [$100 \times (\text{posttest} - \text{pretest}) / (100 - \text{pretest})$; (Fagan *et al.*, 2002)]. Normalization allows valid comparison and averaging of learning gains for students with different pretest scores. A comparison of the F'03 and S'04 courses showed a significant 16% difference ($p = .001$) in average learning gains, corresponding to a 33% improvement in performance by students in the more interactive S'04 course. Learning gains of greater than 60% were achieved by substantially more students in the interactive class (43/70) than in the traditional class (19/72) (Figures 2, 3). Both "A" and "B" students made higher gains in the interactive course, while "C" students achieved about the same learning gain range in both semesters (Figure 2). A more detailed breakdown of gains on individual posttest questions is presented in Appendix C (Figure C1).

In both the F'03 and S'04 courses, we saw a slight negative correlation between pretest performance and learning gain (Pearson correlation test coefficient: -0.4 in F'03, -0.1 in S'04). However, students exhibiting low pretest with high gain, low pretest with low gain, high pretest with high gain, and high pretest with low gain were all represented in both courses (data not shown). In the S'04 interactive class, we saw a higher positive correlation between learning gain and grade achieved (.54) than in the traditional F'03 class (.12). However, even in S'04, individuals with low grades sometimes achieved high learning gains, and vice versa. When we compared normalized learning gains by gender and ethnicity, we found that average learning gains for minority and male students were somewhat lower than for female

Table 4. Comparison of average performance on different assessments for all three courses

Assessment	Performance (average percent of maximum score)		
	F'03	S'04	S'05
Pretest (12 questions) ^a	34	31	37
posttest (12 questions) ^a	65	74	72
Raw learning gain	31	43	38
Normalized learning gain ^b	46	62	61
Hourly exams	71	71	73
Final exam	77	71	76
Problem sets	82	85	90
Participation	N/A	86	86
Final total points	76	81	81

^aData based only on the 12 questions that were common to all three pretests and posttests (see Appendix A).

^bAverage for each class is shown. Normalized learning gains were computed as described in the text and the legend to Figure 2.

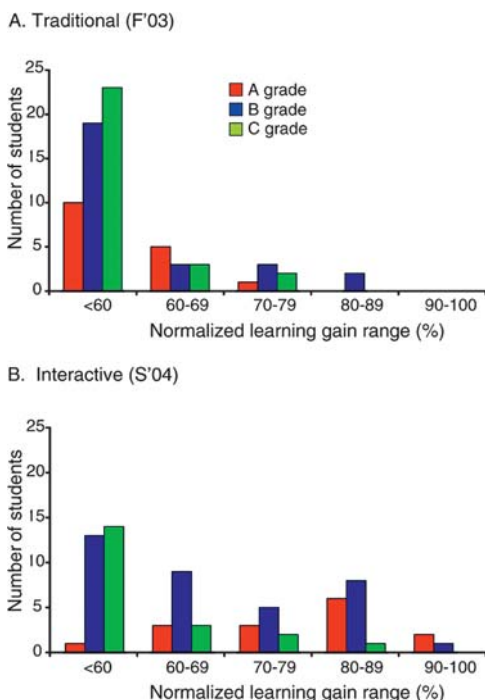


Figure 2. Comparison of normalized learning gain ranges (% of possible maximum) achieved by students in each passing grade range (“A,” “B,” and “C”) in the F’03 and S’04 courses. Normalized learning gains were computed as $100 \times (\text{posttest score} - \text{pretest score}) / (100 - \text{pretest score})$ (see text). A. F’03 (traditional class). B. S’04 (interactive class).

students in F’03, and that the minority average was slightly lower than the total average in S’04 (Table 5); however, these differences were not significant. The ranges of gains were similar across all three groups.

Reproducibility of Learning Gain Comparisons

In the spring of 2005, we again presented the interactive course using the same syllabus and course content, in a very

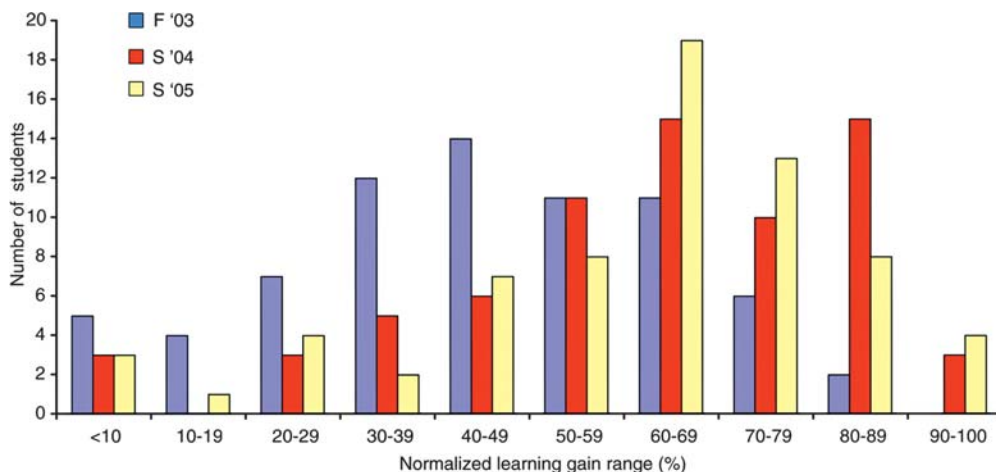


Figure 3. Comparison of normalized learning gains (% of possible maximum) in 10% increments on 12 common pretest and posttest questions for students in one traditional (F’03) and two interactive (S’04, S’05) classes. Normalized learning gains were computed as in Figure 2.

Table 5. Average percent normalized learning gains by ethnicity and gender

Group	F’03 ^a	S’04 ^a
Minority (male and female)	41 ± 21 (17)	56 ± 21 (13)
Males	41 ± 32 (22)	62 ± 19 (31)
Females	51 ± 18 (51)	62 ± 24 (42)
Total	46 ± 23 (73)	62 ± 22 (73)

^aEntries are indicated as ± standard deviations, with number of students in parentheses.

similar format. There were a few changes from the S’04 course: we eliminated the curved grading; we removed the 5% incentive on exams; we allowed students to choose their own group members; and we incorporated additional group activities, such as concept mapping (Novak and Gowin, 1984), into the classes at the expense of lecture time. We also omitted the previous structured class discussions of current journal articles and instead gave students data from current developmental biology articles with questions to answer as a group in class. To maintain approximately the same amount of course content in spite of reduced lecture time (more reduced than in S’04), we asked students to take more responsibility for their own learning. We assigned reading for each week in advance and gave a graded multiple-choice quiz on the reading material (using clickers) at the start of each Tuesday class; the points earned (five maximum), were counted toward students’ participation grade. Prior to each Thursday class, we required that students work with their groups to complete a set of homework problems and post their answers on the course Web site. This allowed us to review their performance in advance and identify the concepts that were causing frequent difficulties, which we could then focus on in class.

Normalized learning gains from the S’05 class are compared with those from F’03 and S’04 in Table 4 and Figure 3, based on results from the 12 validated pretest and posttest questions that were the same in all three semesters.

As in the S'04 course, a substantial fraction of the S'05 class (34/69) achieved greater than 60% normalized learning gains, and 4/69 achieved greater than 90% gains, compared with 3/70 in S'04 and 0/72 in F'03. The increases in average normalized learning gains over those in the F'03 course were almost identical in the S'04 and S'05 courses. The variation of gains on individual questions was also generally reproducible (Appendix C). We conclude that the effects of interactive engagement seen between the F'03 and S'04 courses are reproducible.

Performance on Repeated Homework Problems

We also compared student performance on 19 homework problems that were assigned in the F'03 and S'04 courses. All were short-answer questions graded by different TAs, but following the same guidelines in both semesters. Although performance data from the two semesters are not strictly comparable, because of the increased collaboration on homework in S'04 (see Experimental Design), the comparison nevertheless suggests an additional benefit from the interactive format. Performance on all but two of the questions was significantly better in S'04 than in F'03. The 19 questions could be grouped into three categories. In the first were six questions that involved primarily recall. The remaining 13 required conceptual understanding, and we had designated six of these as "puzzlers" or difficult questions with no obvious answer, for which students had to apply facts and concepts they had learned in an unfamiliar context. Whereas improvement over F'03 performance was <10% on all but one of the six recall questions, it ranged from 10% to 40% on the six puzzlers. These results suggest that the interactive, collaborative format followed in the S'04 course positively affected students' ability to learn and apply conceptual understanding.

Use and Success of In-Class Clicker Questions and Group Work

Personal response systems using clickers have been used for several years in our Physics Department. As a result, about one-half of the students in our S'04 and S'05 classes had used them previously. However, none had experienced clickers in a biology class, since ours was the first to use them. Although we did not manage to make every student like using the clickers (see Student Attitudes below), we were impressed by the effectiveness of this technology in changing the dynamic of the classroom and providing us as well as students with valuable feedback (Wood, 2004). The ICQs revealed student misconceptions and helped identify concepts that were difficult for students to grasp. They provided a way for individual students, who otherwise might have assumed they alone had misunderstood a concept, to realize that in many cases, one-half the class was similarly confused. They allowed us to determine, on the spot, when a concept was not getting across and needed more explanation or discussion. Most importantly, the clicker questions actively engaged the students, allowing them to talk and think about the topic of the class rather than simply record it, and to interact with teaching staff during the class period. Students quickly became comfortable asking questions of us and each other because interaction was part of the classroom dynamic from the first day.

The clickers also synergized effectively with group work in promoting active engagement and conceptual learning during class. We found that ICQs for which about half the class chose an incorrect alternative to the correct answer provided valuable teaching and learning opportunities (Wood, 2004). There was often a palpable tension in the classroom until the disagreement was resolved. To exploit this tension, it was important not to reveal the correct choice immediately, but rather to let the students work it out through discussion with the members of their group. Almost inevitably, when a second vote was taken after 3–4 min of discussion, more than 75% of the class chose the correct answer (Appendix B).

After the second vote, we always held a brief discussion of why students might have first chosen incorrect answers and then changed their choices. For some questions it became clear that conceptual misunderstanding was not the problem; rather, students had simply misunderstood the question, which was somehow poorly phrased or otherwise ambiguous. For questions that did reveal conceptual misunderstanding, one could imagine two possible extreme explanations for the greatly improved performance following group discussion: 1) students were simply identifying the group member most likely to know the correct answer based on previous experience and following that person's choice, or 2) students were meaningfully debating the alternatives and explaining them to each other, ultimately leading their group to the best choice. During these discussions, the instructors, TAs, and LAs circulated among the students (limited by the constraints of classroom design; see Discussion), listening in on discussions and attempting to facilitate them if necessary (without giving away the answer). From our own observations and subsequent debriefing of the TAs and LAs, we concluded that most groups were engaging in constructive dialog that led them to the correct answer.

Group Dynamics

Group work on homework problems during the lab periods was also moderately successful, but worked better for some groups than for others. Whereas some groups discussed and worked through the problems together as we had intended, with all members participating, other groups split up the work, each member doing one problem, and pooling their answers. A few groups were dominated by one or two members who did most of the talking, while the other members listened and copied their work.

Despite our view that group work was clearly successful in class and moderately successful outside, four of the 20 groups in the S'04 class reported that they did not work well together because of personality conflicts, differing work ethics, or differing levels of interest. Because the same groups were also working together in the accompanying lab course, in which there were ongoing projects, we did not reconstitute them during the course of the semester as recommended by Beichner and Saul (2003).

We also found that the incentive for group members to work together, described under Experimental Design, was largely unsuccessful. For the first exam, only one of the 20 groups achieved an average of greater than 80%, and for the second exam, only three groups did so. Students reported that the incentive made them uncomfortable because they did not like sharing their scores, nor did they feel it was fair. Finally, students reported, quite vehemently, that they did

not like being assigned to groups. Overall, we did not see any benefit to the assigned groups; whether a group worked well seemed to depend more on personalities and similar motivation to succeed than on a balance of previous student performance records.

In the S'05 course, in which students chose their own group members, the groups worked together much better than in S'04, both in class and in solving homework problems during scheduled problem-solving sessions. This conclusion was based on both self-reporting and our observations. In S'05, only two out of 20 groups shared complaints with us about poor functioning; in both cases, the problem was a single member of the group not participating, while the rest of the group worked well together.

Student Attitudes

General Reactions to the Interactive Format. Although many students in the S'04 and S'05 courses at first disliked and distrusted the interactive classes and group activities, most became comfortable with the unfamiliar format and ultimately reported that it helped their learning. Some students, however, never fully accepted the new approaches. One common complaint was that the interactive techniques promoted grade leveling (since students received points for participation, and could discuss answers with group members). Other complaints included annoyance at being "forced" to attend class and work in groups (see Discussion).

In the traditional class, complaints tended to focus not on the style of the class, but rather on the content or sheer volume of information. For example: "I was so concerned with just getting everything done that I'm not sure I really learned anything." "A ton of information was crammed into our heads."

When we asked students in the traditional class to anonymously write down what they had learned at the end of a particular lecture, one telling comment was: "I can't answer this. In class I just take notes. Then I go home and try to figure out what we talked about."

In the interactive class, students in general tended to feel strongly about the format. We received negative comments, such as: "Clicker questions took too long and took away from the learning experience." "Clickers are promoting mass mediocrity by evening everyone out."

But we also received positive ones; for example: "I've never had a class where I felt like the teachers cared as much for the students as this one." "It was good to see what other people thought about problems because sometimes even if we were both wrong, information from the two wrong solutions could lead to a correct solution."

Reactions to Specific Aspects of the Interactive Format. In the interactive S'04 and S'05 courses, we asked students to rate various components of the course for helpfulness in learning the material on a scale of one (not) to five (very) useful (Table 6). We do not have comparable data from the F'03 course, except on the usefulness of posted lecture notes and PowerPoint slides, which were equally highly appreciated in all three semesters. Students in the S'04 class were on average fairly neutral regarding usefulness of the interactive techniques overall and somewhat more positive about in-class group work. In the S'05 course, average ratings were significantly higher for both these categories,

Table 6. Student reactions to specific components of the interactive course format

Course component	Semester		
	F'03	S'04	S'05
Posted lecture notes and PowerPoint slides ^a	4	4	4
In-class interactive activities, in general ^a	—	2.8	4.2
In-class group work ^a	—	3.7	4.8
Significant learning took place during class (as well as outside of class)	—	65%	70%

^aRated by students on a Leichert scale of one (not useful for learning) to five (very useful).

possibly for reasons suggested in the Discussion. A substantial majority of students in both S'04 and S'05 reported that significant learning took place during class as well as outside of class; unfortunately we do not have comparable data for the F'03 course.

On other topics, about 90% of the students in both S'04 and S'05 reported that they still studied for the exams on their own rather than in groups, but only 40%–50% of these felt this technique was successful. In comparing the results of F'03 and S'04 Faculty Course Questionnaires, we were interested to note no difference in ratings of the course workloads, despite substantially more assignments and requirements that demanded student time in the interactive S'04 class: about two-thirds of students rated the workload as too high in both semesters. In F'03, 69% of the students gave the course an overall rating of "B–" or higher. They rated the S'04 course slightly less positively (61%), while in S'05, the course rating was much more positive (86% \geq "B–").

We were particularly interested in student attitudes toward the clickers, which we viewed as a powerful aid to both teaching and in-class learning. At the beginning of the S'04 course, students were on average neutral in their overall reaction to using clickers in class, although there were strong negative and positive opinions on the ends of the scale (Figure 4). By the end of the semester, the distribution of attitudes had shifted toward a more positive opinion, although some students were still strongly negative, possibly for reasons discussed below.

DISCUSSION

Positive Effects of Interactive Classes and In-Class Cooperative Activities on Student Learning

The experience we report here, in an advanced biology class of students previously taught almost exclusively in lecture courses, suggests that even a partial shift toward a more interactive and collaborative course format can lead to significant increases in student learning gains. Moreover, students in the interactive environment apparently developed better skills for solving conceptual problems than the similar group of students taught only through lectures. In particular, our results indicate that, as observed in elementary physics courses by Beichner and Saul (2003), both "A" and "B" students had higher learning gains in an interactive environment than in a traditional environment, whereas "C"

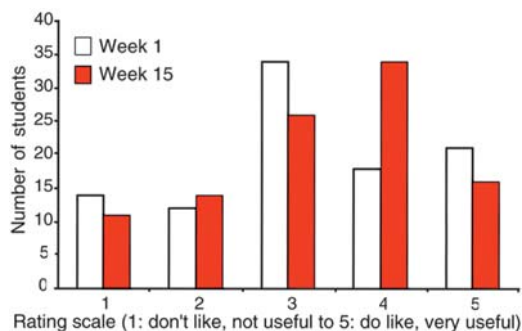


Figure 4. Changes in student attitudes about clickers during the semester in the interactive S'04 course. Near the beginning (open bars) and near the end (solid bars) of the course, students were given an ICQ asking them to rate the usefulness of the personal response system for their learning in class, on a Leichert scale of 1 (not at all useful; don't like), to 5 (very useful; like very much). Reasons for why some students persisted in their dislike of the clickers are discussed in the text.

students made relatively low learning gains in both environments. Recognizing and addressing the needs of these students is an important future goal.

Even though our assessment instrument for learning gains was rudimentary, and our experiment was not ideally controlled, we believe our results are meaningful for several reasons. First, we have been interacting closely with this student population (junior and senior MCDB majors) for years and have identified the topics that challenge them; we designed our assessment instrument around these topics. A second reason is the reproducibility of our results; a repeat of the interactive course with a different set of students showed a very similar increase in learning gains over the control course. A third reason is an uncontrolled variable that almost certainly worked against learning in the interactive format. Both our interactive "experimental" classes, in S'04 and S'05, included greater than 60% spring-term seniors (Table 1), a generally undermotivated group! Our finding that these students nevertheless made greater learning gains and apparently achieved better conceptual understanding than their peers in the fall control course strengthens our confidence in the results. A final reason is that our findings parallel those of similar studies conducted in other disciplines, notably physics. During the past decade, physicists have used a standardized, validated assessment of conceptual understanding of force and motion known as the Force-Concept Inventory (FCI; Hestenes, 1992) to measure learning gains, which they could then compare between traditionally taught and interactive classes. Their results show compellingly that in comparison to traditional lecture courses, interactive courses including features we have described here result in significantly higher normalized student learning gains at all achievement levels (e.g., Beichner and Saul, 2003; Hake, 1998).

We believe that a substantial impact could be achieved if interactive and collaborative teaching were introduced in introductory courses and continued throughout the curriculum. Such reforms may also be necessary simply to maintain enough interest among incoming students to make them want to continue as biology majors. More and more of these students are likely to experience interactive, inquiry-based teaching and collaborative learning in secondary-

school biology courses, where there is also increasing evidence for its efficacy (e.g., Lord, 1998, 2001). The Advanced Placement Biology syllabus is currently undergoing revision to a more inquiry-based format, with less emphasis on factual knowledge and more on in-depth understanding of biological concepts (Educational Testing Service, 2002). Incoming students with this sort of background are likely to require more to maintain their interest in biology than large impersonal lecture courses emphasizing memorization of factual knowledge.

Student Reactions to Interactive Classes and Group Learning

The dynamic in our interactive classes was almost certainly affected by the stage of the students in their undergraduate careers. By their junior and senior years, most had developed study habits they believed were effective, and they were reluctant to question or change these habits, even when faced with evidence that they do not work well for long-term retention or conceptual understanding. They also had become comfortable with traditional course structures, and some resented having to adapt to a new format (see Allen and Tanner, this issue). Many students have unfortunately adopted the misconception (promulgated by some professors) that meaningful teaching and learning in the classroom take place only while the instructor is actually lecturing (i.e., transmitting information; see Kliensky, 2004). Particularly at the start of the course, these students saw other classroom activities as distractions and a waste of time. Even at the end of the semester, students who disliked the interactive format complained that we were not teaching them very much, but rather making them learn the material on their own. We tried to explain to them why we found this perception gratifying.

We can suggest two reasons for why some students persisted in their dislike of the clickers. First, in the traditional course, some of the better students were able to learn the course content adequately from reading assignments, posted lecture notes, and posted PowerPoint slides. These students often did not come to class, feeling (perhaps justifiably) that the lectures were not providing any additional value. In the interactive courses, because participation via the clickers was rewarded by course points, these students felt compelled to come to class in order to achieve a high grade. In fact, class attendance increased from 60% to 70% in the F'03 course to greater than 95% in the interactive courses. A second reason for disliking clickers was the time they took away from lecturing, which most students feel is essential for their learning, as mentioned above. Finally, rewarding correct clicker answers with points led to an atmosphere in which students were constantly thinking about how many points they had in the class. This created a higher level of concern about their grade on a day-to-day basis, something we did not intend to emphasize.

Regarding group work, students who had performed well in previous courses were particularly prejudiced against it, since they felt they would be helping their group members without getting anything in return. Their competitive instincts, fostered by years of being graded on a curve, tended to strengthen this prejudice. Students in the S'05 course had fewer complaints about group work than those in S'04, probably for two main reasons. First, the students in S'05 chose their own groups, and thus seemed to feel more

loyalty to other group members than students in S'04, who were assigned to their groups. A second factor was our own increased comfort level with the interactive format after a semester of experience. Students in S'04 knew they were part of an "experiment," while students in S'05 accepted the unfamiliar format as simply the way this course was taught.

Solving the Content Problem

A common concern among faculty who contemplate introducing more interactive classroom activities in place of lecturing is that course content will have to be limited. Because clicker responses and discussion take time, less material can be "covered" during each class. As pointed out in the Introduction, content knowledge is not the only benefit, and perhaps not the most important benefit, that students gain from a good course. Ability to solve problems and in-depth understanding of underlying concepts will probably be of more use to them in the long run than any particular piece of factual information (Kitchen *et al.*, 2003).

Nevertheless, there is a certain amount of content knowledge that will seem essential for any required course in the curriculum, particularly for students who are continuing in the discipline. In our experience, the best way to make sure that students assimilate all the essential content is to place on them more of the responsibility for learning it, outside of class time. We employed this approach successfully in our S'05 course as described under Results. Because students are required to demonstrate their understanding of each week's material by posting answers to homework problems on the course Web site before class, the instructors receive the added benefit of knowing in advance which concepts the students are finding difficult, so that they can tailor classroom activities to address these problems. The purpose of classes then changes, from the traditional one of transmitting information to one of helping students understand and apply concepts from their reading and homework assignments. This approach, first described several years ago for use in physics courses as "Just-In-Time Teaching" (Mazur, 1996; Novak *et al.*, 1999), has also been used successfully in teaching biological sciences (e.g., Kliensky, 2004).

Concerns for Instructors

The active-engagement teaching we have described here presents several practical problems for instructors. First, the extra time and effort required to design and teach such a course is substantial if an existing lecture course is being transformed, as in our case. We note, however, that for a beginning instructor preparing to teach a subject for the first time, creating an interactive course may not require substantially more effort than preparing a semester's worth of lectures.

Second, the physical facilities provided by most universities do not favor this kind of teaching. Interaction between teaching staff and students during discussion of our ICQs and other in-class group activities was greatly hampered by the theatre-style design that is unfortunately standard for large university classrooms. Café-style classrooms, such as those designed by Beichner and his colleagues (2003), would be far more conducive to active-engagement teaching, and should be introduced in universities as new classrooms are built.

Third, because students at present are used to having most large courses taught in the lecture format, the unfamiliar demands of an active-engagement course may take them out of their comfort zone, resulting in lower student ratings for the instructor. In our experience, some students did react in this manner, but others were enthusiastic about the high level of instructor involvement and its effect on their learning. These students rated the course as superior to their other lecture-style courses. We believe an antidote to this concern is more emphasis on student learning gains as part of the faculty teaching evaluation. Although biology so far generally lacks assessments comparable to the FCI in physics (but see Anderson *et al.*, 2002; Khodor *et al.*, 2004; Klymkowsky *et al.*, 2003), we found it fairly straightforward to develop an assessment tool for measuring learning gains (see also Dancy and Beichner, 2002). Once baseline learning gains have been established for a course, the effects of new teaching strategies can be determined by comparing measured learning gains to baseline data (the principle of scientific teaching, Handelsman *et al.*, 2004). If preassessments and postassessments such as those we have used here were to become standard practice, then evaluation of teaching could be based on actual student learning gains as well as student course ratings.

An additional, more subtle concern for instructors is that we, like the students, have generally become comfortable with the lecture format. We found that adjusting to a decreased emphasis on information transmission during class was difficult not only for our students, but also for us. Part of what inspired us to become teachers was a delight in explaining fascinating aspects of our discipline to students. Good lecturing is a skill in which many of us take pride, and there is no doubt that an outstanding presentation can be a formative experience for students, remembered for many years. However, such rare lectures generally have enduring effects not because of the factual information they convey, but more likely because they inspire by example, or reveal unexpected new insights, or open up new worlds to their listeners. Few instructors can hope to deliver at best more than one or two such lectures during a semester. Thus, while lecturing in small doses remains a valuable teaching technique, lecturing for an entire period of 50 or 75 min is unlikely to be the best use of class time.

Like the students who bring prior knowledge to our courses in the form of misconceptions about biology (Duit and Treagust, 2003; Wandersee *et al.*, 1994), we as instructors must face up to the common pedagogical misconception that students will learn effectively only what we tell them in class. Like our students, we must allow ourselves to undergo conceptual change (Tanner and Allen, 2005), based on evidence for the greater learning effectiveness of active engagement over passive listening in lectures. This evidence is now overwhelming (see National Research Council, 1999, for a review of earlier work). Even in large courses, clicker technology (Wood, 2004) and Web-based course management software (Novak *et al.*, 1999) make active-engagement classroom activities feasible without sacrificing content. We have shown here that even a partial change in this direction can lead to significantly increased learning gains. We urge our colleagues to adopt these approaches throughout the biology curriculum.

ACKNOWLEDGMENTS

We are grateful to Valerie Otero and Carl Wieman for helpful discussions and comments on the manuscript, and to Carl Wieman for partial support of J.K. We also acknowledge the dedicated participation of several undergraduate learning assistants; those participating in the spring 2004 course were supported by a STEM-TP grant from the National Science Foundation (Richard McCray, Principal Investigator; DUE-0302134).

REFERENCES

- Allen, D., and Tanner, K. (2005). Infusing active learning into the large enrollment biology class: seven strategies, from simple to complex. *Cell Biol. Educ.* 4, 262–268.
- Anderson, D.L., Fisher, K.M., and Norman, G.J. (2002). Development and evaluation of the Conceptual Inventory of Natural Selection. *J. Res. Sci. Teach.* 39, 952–978.
- Beichner, R.J., and Saul, J.M. (2003). Introduction to the SCALE-UP (Student-Centered Activities for Large Enrollment Undergraduate Programs) Project. Proceedings of the International School of Physics “Enrico Fermi,” Varenna, Italy. <http://www.ncsu.edu/per/scaleup.html> (accessed 7 June 2005).
- Dancy, M.H., and Beichner, R.J. (2002). But are they learning? Getting started in classroom evaluation. *Cell Biol. Educ.* 1, 87–94.
- Duit, R., and Treagust, D.F. (2003). Conceptual change: a powerful framework for improving science teaching and learning. *Intl. J. Sci. Educ.* 25, 671–688.
- Educational Testing Service (2002). Access to Excellence: Report of the Commission on the Future of the Advanced Placement Program. Princeton, NJ: The Educational Testing Service.
- Fagan, A., Crouch, C.H., and Mazur, E. (2002). Peer instruction: results from a range of classrooms. *Phys. Teach.* 40, 206–209.
- Gilbert, S. (2003). *Developmental Biology*, 7th ed., Sunderland, MA: Sinauer Associates.
- Hake, R.R. (1998). Interactive-engagement vs. traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66, 64–74.
- Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S.M., and Wood, W.B. (2004). Policy forum: scientific teaching. *Science* 304, 521–522.
- Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher* 30, 141–158.
- Khodor, J., Halme, D.G., and Walker, G.C. (2004). A hierarchical biology concept framework: a tool for course design. *Cell Biol. Educ.* 3, 11–121.
- Kitchen, E., Bell, J.D., Reeve, S., Sudweeks, R., and Bradshaw, W. (2003). Teaching cell biology in the large-enrollment classroom: methods to promote analytic thinking and assessment of their effectiveness. *Cell Biol. Educ.* 2, 180–194.
- Klionsky, D.J. (2004). Talking biology: learning outside the book—and the lecture. *Cell Biol. Educ.* 3, 204–211.
- Klymkowsky, M.W., Garvin-Doxas, K., and Zeilik, M. (2003). Bioliteracy and teaching efficacy: what biologists can learn from physicists. *Cell Biol. Educ.* 2, 155–161.
- Lord, T. (1998). Cooperative learning that really works in biology teaching. *Am. Biol. Teach.* 60, 580–587.
- Lord, T. (2001). 101 reasons for using cooperative learning in biology teaching. *Am. Biol. Teach.* 63, 30–37.
- Mazur, E. (1996). *Peer Instruction: A Users Manual*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- National Research Council (1999). *How People Learn: Brain, Mind, Experience and School*. Washington, DC: National Academies Press.
- Novak, G.M., Gavrin, A.D., Christain, W., and Patterson, E.T. (1999). *Just-In-Time Teaching: Blending Active Learning with Web Technology*. Upper Saddle River, NJ: Prentice-Hall, Inc..
- Novak, J., and Gowin, D.B. (1984). *Learning How To Learn*. Cambridge, UK: Cambridge University Press.
- Tanner, K., and Allen, D. (2005). Understanding the wrong answers—teaching toward conceptual change. *Cell Biol. Educ.* 4, 112–117.
- Wandersee, J.H., Mintzes, J.J., and Novak, J.D. (1994). Research on alternative conceptions in science. In: *Handbook of Research on Science Teaching and Learning*, ed. D. Gabel. New York: Simon & Schuster Macmillan, 177–210.
- Wood, W.B. (2004). Clickers: a teaching gimmick that works. *Dev. Cell* 7, 796–798.

APPENDIX A

Pretest and posttest questions used to measure learning gains in the F'03 and S'04 courses. The 12 questions that were also used in the S'05 course are marked with an asterisk. The answers scored as correct are in bold-face type.

1. The five most commonly studied metazoan model organisms: *C. elegans*, *Drosophila*, *Xenopus*, chick, and mouse, are important for biomedical research because all of them
 - a) **are simpler and/or experimentally more convenient than humans.**
 - b) **are descended from a common metazoan ancestor.**
 - c) are representative of five different phyla.
 - d) **use many of the same developmental and physiological mechanisms as humans, though they appear superficially very different.**
 - e) **have one or more characteristics that facilitate study of certain aspects of development.**

 - 2* We still don't understand very well how genes control the construction of complex structures, like the antennae of the fruit fly, *Drosophila melanogaster*. If you wished to identify genes that control antennal development and find out what proteins they encode, the best way to *begin* would be to:
 - a) isolate a gram of *D. melanogaster* antennae and extract mRNAs to make cDNA clones.
 - b) find another *Drosophila* species with different antennal morphology and genetically map the genes responsible for the difference.
 - c) obtain a large population of embryos at the stage when antennae are beginning to form, label them with ³²P, extract the labeled mRNAs, make the corresponding cDNAs, and sequence them.
 - d) **mutagenize a population of *D. melanogaster* wild type and screen their progeny for mutants with no antennae or altered antennal morphology; then genetically map the mutations responsible.**
 - e) search the database of sequenced *D. melanogaster* genes for homologs of antennal genes in other organisms.

 - 3* *Drosophila* strains carrying a mutation in a gene you have named *ant* have no antennae. You suspect that *ant* could encode a known transcription factor called PT3. The PT3 gene has been cloned and sequenced. A good test of whether *ant* is the PT3 gene would be to:
 - a) determine whether a PT3 gene probe will hybridize to any mRNAs from an *ant* mutant.
 - b) **determine whether the cloned PT3 gene injected into an *ant* mutant embryo could rescue (correct) the antennal defect.**
 - c) **determine whether double-stranded RNA made from the cloned PT3 gene and injected into the embryo causes lack of antennae.**
 - d) **isolate the PT3 gene from *ant* mutant fly DNA and determine whether its sequence is different from that of the normal PT3 gene.**
 - e) use PT3 DNA to probe Southern blots of digested genomic DNA from wild-type and *ant* mutant embryos, and ask whether the hybridization patterns are different.

 - 4* Which of the following statements about ligands and receptors is/are true?
 - a) Components of the extracellular matrix never serve as signaling ligands.
 - b) The receptors for steroid ligands are membrane bound.
 - c) **Many ligands interact with receptors in target cell membranes to activate signaling pathways.**
 - d) Juxtacrine signaling involves diffusible ligands and membrane bound receptors.
 - e) Two cells with the same receptor will always respond identically to a given ligand.

 - 5* The number of different signaling pathways involved in embryonic development is
 - a) less than 5.
 - b) **between 5 and 10.**
 - c) between 10 and 15.
 - d) between 15 and 20.
 - e) more than 20.

 6. The position and orientation of the cleavage furrow that separates two mitotically dividing cells during cytokinesis are
 - a) usually determined by external cues.
 - b) **equidistant from the two poles of the mitotic spindle and orthogonal to it.**
 - c) equidistant from the two poles of the dividing cell and orthogonal to a line connecting them.
 - d) often orthogonal to the cleavage plane of the preceding division.
 - e) controlled by interaction of microtubules with the cortex of the dividing cell.

 - 7* Epithelial cells are different from mesenchymal cells in that epithelial cells:
 - a) **have polarity, defined by an apical and basal side.**
 - b) **are tightly adherent to each other.**
 - c) are loosely adherent to each other.
 - d) are usually defined as migratory.
 - e) **typically comprise the linings of organs.**

 8. Which of the following is/are likely to bind to specific DNA response elements and activate or repress the transcription of specific genes?
 - a) A growth factor receptor.
 - b) A protein kinase.
 - c) A G protein.
 - d) **A steroid hormone receptor.**
 - e) None of the above.
- From the numerical choices below, pick the one experimental technique that could best be used to answer each of the questions 9–11.
- 1) Gel mobility shift assay with labeled DNA.
 - 2) RNase protection assay.
 - 3) In situ hybridization.
 - 4) RNA interference.

- 9* Which of two tissues contains more of a particular mRNA? **2**
- 10* Where is a particular transcript (for which you have an RNA probe) present in an embryo? **3**
- 11* What is the phenotype of an embryo in the absence of a particular transcript? **4**
- 12* Consider a recessive, maternal-effect *C. elegans* mutation, *m*, that causes embryos to die. Which of the following statements is/are true?
- Whether an embryo dies will depend on the genotype of the embryo.
 - Whether an embryo dies will depend on the genotype of the hermaphrodite parent.**
 - If a heterozygous (*m/+*) hermaphrodite is mated to a heterozygous male 1/4 of the progeny embryos will die.
 - If a homozygous (*m/m*) hermaphrodite is mated to a wild-type male, all of the progeny embryos will die.**
 - The experiment in (d) cannot be done, because *m/m* hermaphrodites will always die as embryos.
- 13* In all animal embryos, the process of gastrulation accomplishes the following important function(s):
- Patterning the anterior-posterior axis.**
 - Establishing which side of the embryo will be dorsal.
 - Bringing endodermal cells into the interior of the embryo.**
 - Bringing ectodermal cells into the interior of the embryo.
 - Bringing endodermal and ectodermal cells into contact for inductive interactions.**
- 14* Targeted alteration (knockout or mutation) of a specific gene in the germ line of an animal requires
- that the animal's genome has been completely sequenced.
 - that the gene in question has been cloned.**
 - a method for introducing DNA into cells that are in or will give rise to the germ line.**
 - homologous recombination of introduced DNA with the resident gene on a chromosome.**
 - nonhomologous recombination of introduced DNA with the resident gene on a chromosome.
- 15* Programmed cell death (apoptosis)
- occurs only in invertebrates.
 - occurs in response to injury.
 - is necessary to prevent cancer in mammals.**
 - occurs only in certain degenerative disease conditions.
 - is important in limb morphogenesis.**

APPENDIX B

A selection of clicker questions supporting the effectiveness of peer instruction (Mazur, 1996). Questions were used in the order shown during the course. For each question, the table shows the percentage of correct answers initially, based on individual student responses prior to discussion, and the percentage of correct answers following a 4–5 min discussion of the initial responses in student groups. The peer instruction format worked best when the initial responses were ~50% correct. Correct answers are shown in bold-face type.

In-class question	Responses (% correct)	
	Before discussion	After discussion
1. This question was asked in the second class period, to determine how well students understood the roles of regulatory elements and transcription factors in cell fate determination. ICQ: Choose the best answer. Different kinds of cells in an animal synthesize different proteins. The reason(s) for why this happens include the fact that different cells have different a) assortments of chromosomes. b) coding sequences in their DNA. c) combinations of transcription factors present. d) combinations of DNA regulatory elements for some coding genes. e) more than one of the above four reasons.	28%	79%
2. Maternal-effect mutants. This question followed a 20-minute lecture on why embryonic lethal mutations in <i>C. elegans</i> show a maternal effect if they inactivate genes encoding proteins essential for embryogenesis that the hermaphrodite parent puts into the oocyte during oogenesis, with physiological and genetic explanations. A pedigree was displayed to the class, in which a recessive, strict maternal-effect embryonic lethal mutation <i>m</i> arises in the germ line of a mutagenized <i>Po</i> animal, so that an <i>m/+</i> F1 animal is produced (Wood, 2004). It will produce F2 embryos of three genotypes: <i>+/+</i> , <i>m/+</i> , and <i>m/m</i> . ICQ: will the <i>m/m</i> F2 progeny embryos a) live or b) die?	49%	95%
3. Determination of cell fates in the embryonic cellular blastoderm of <i>Drosophila</i> embryos. This question followed a lecture on how the embryo is sequentially patterned, starting with anterior Bicoid and posterior Nanos gradients, by formation of overlapping expression domains of gap gene-encoded transcription factors in the syncytial blastoderm, which then act in different combinations along the A-P axis to regulate genes with complex promoters in a region-specific manner after cellularization.		

ICQ: How is the positional identity of each cell established in the early *Drosophila* embryo?

- a) By signaling between cells.
- b) By segregation of different mRNAs into each cell. 29% 76%
- c) **By access to different levels of transcription factors.**

4. This question followed discussion of sex determination and dosage compensation in *C. elegans*, near the end of the course.

ICQ: A homozygous *lf* mutation in the *C. elegans* gene *sdc-2* will kill otherwise normal hermaphrodite but not male embryos. Consider double-mutant embryos of the following four genotypes. Which two of them will die?

- 1) *sdc-2*(0); *tra-1*(0) XX (male).
 - 2) *sdc-2*(0); *tra-1*(0) XO (male).
 - 3) *sdc-2*(0); *her-1*(0) XX (hermaphrodite).
 - 4) *sdc-2*(0); *her-1*(0) XO (hermaphrodite). 60% 87%
- a) 1 and 2.
 - b) 3 and 4.
 - c) **1 and 3.**
 - d) 2 and 4.

APPENDIX C

Increased average learning gains were seen for some but not all pretest/posttest questions in the interactive class. Average student learning gains differed on individual questions in the pretest/posttest. To determine whether certain questions produced a higher average learning gain in one style of class, we subtracted the average learning gain for the traditional class from the average learning gain for each of the interactive classes for each question. Figure C1 shows the range of differences per question between the two interactive classes (S'04 and S'05) and the traditional class (F'03). There were five questions from the pretest/posttest in which the difference between average learning gain increased by greater than 10% in both S'04 and S'05 (questions 2, 5, 7, 10, 13) and one in which the gain decreased by greater than 10% (15; this question was on a topic, apoptosis, which we covered in F'03 but only mentioned in S'04 and S'05 and did not hold students responsible for). There was only one question on which the performance in the two interactive semesters showed a significant mismatch (9). Differences on the remaining questions are probably not significant.

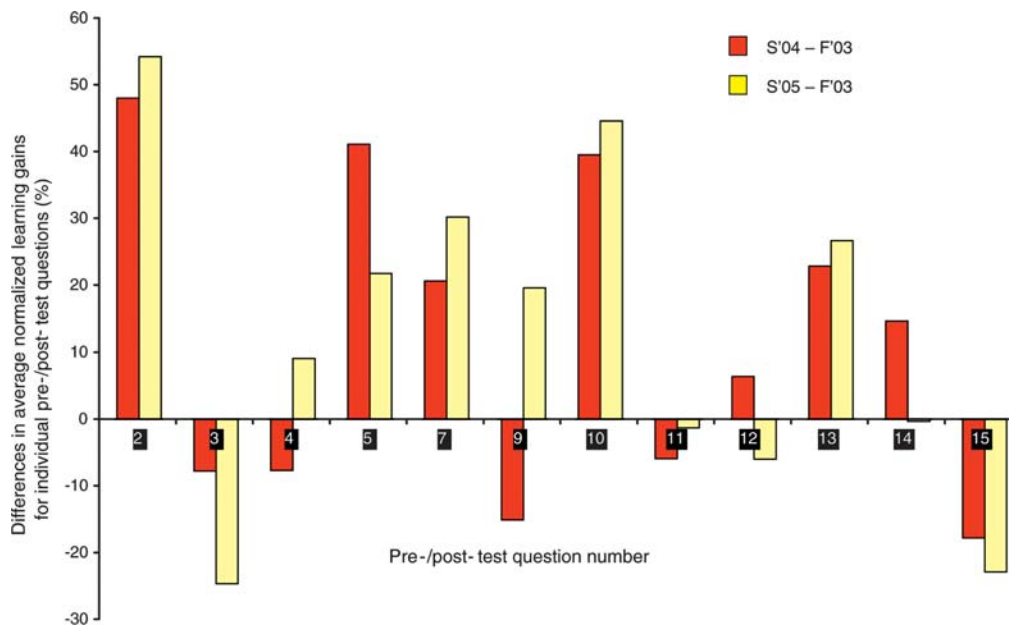


Figure C1. Differences in average percent normalized learning gain for each of the 12 repeated questions on the pretest/posttest in S'04 and S'05. Question numbers are indicated on the horizontal axis; we obtained the differences for each question by subtracting the average gain in F'03 from the average gains in S'04 (red bars) and S'05 (yellow bars). On questions for which more than one answer was correct, we gave partial credit for each correct answer and deducted partial credit for incorrect answers (60% of the value of each correct answer). See text for further discussion.