

Article

Cooperative Learning in Industrial-sized Biology Classes

Norris Armstrong,* Shu-Mei Chang,[†] and Marguerite Brickman[†]

Departments of *Genetics and [†]Plant Biology, University of Georgia, Athens, GA 30602

Submitted November 13, 2006; Revised February 22, 2007; Accepted February 28, 2007
Monitoring Editor: Diane Ebert-May

This study examined the impact of cooperative learning activities on student achievement and attitudes in large-enrollment (>250) introductory biology classes. We found that students taught using a cooperative learning approach showed greater improvement in their knowledge of course material compared with students taught using a traditional lecture format. In addition, students viewed cooperative learning activities highly favorably. These findings suggest that encouraging students to work in small groups and improving feedback between the instructor and the students can help to improve student outcomes even in very large classes. These results should be viewed cautiously, however, until this experiment can be replicated with additional faculty. Strategies for potentially improving the impact of cooperative learning on student achievement in large courses are discussed.

INTRODUCTION

Most people would agree that one of the best formats for teaching is one-to-one interaction between an instructor and student. In this setting, continuous feedback is possible, enabling the student to work at his/her own pace and level and for the instructor to tailor the lesson to each student's individual needs. Close interaction with the instructor also helps to engage students and encourages them to become an active partner in the learning process (Bloom, 1984).

Unfortunately, large class sizes and the many demands placed on instructors make it difficult to devote even small amounts of time to single students. The reality is that many college-level science classes are taught almost exclusively using a lecture format even though lectures by themselves are relatively ineffective at engaging students and promoting learning (Bransford *et al.*, 1999; Powell, 2003; Handelsman *et al.*, 2004). Therefore, a significant challenge facing many instructors today is how to help students learn in situations where it is not possible to interact with them on an individual level. One partial solution to this dilemma that has grown in prominence over the past few decades is the use of cooperative learning strategies.

Cooperative learning encompasses a variety of approaches that encourage students to work together in small groups to achieve success rather than compete for a grade. Collaborating in this manner encourages students to take

responsibility for their own learning and provides a non-threatening environment in which they can receive help from their peers. Students who provide help to others actually clarify their own knowledge of the concepts they are trying to explain and construct a more elaborate and sophisticated understanding of the material themselves. Conversely, the students who receive individualized attention fill in gaps in their knowledge and correct their misconceptions. When explaining a concept, students may also use terms or examples that differ from the instructor's and that other students may find more familiar. Lastly, cooperative learning can make large classes less impersonal and can increase participation, especially by quieter students. Increased participation can, in turn, help the instructor identify those topics that need further exploration (Lord, 1994; Cooper, 1995; Webb *et al.*, 1995; Johnson *et al.*, 1998).

Using cooperative learning as a teaching tool is not new, and research comparing the effectiveness of this method with individual/competitive learning approaches has been carried out since the 1800s. A review of 168 studies involving college-aged students from 1924 to 1997 reported that cooperative learning promoted greater achievement than competitive or individual approaches over a wide range of measures (acquisition, retention, and accuracy of knowledge), skills (reading, writing, math, etc.), and tasks (knowledge vs. activities). In addition, cooperative learning also helped students become more aware of their own knowledge (metacognition) and improved relationships among students, self-esteem, social skills, and attitudes toward the course (Johnson *et al.*, 1998). In a meta-analysis of more

DOI: 10.1187/cbe.06-11-0200

Address correspondence to: Norris Armstrong (narmstro@uga.edu).

recent studies that examined cooperative learning in college science, technology, engineering, and mathematics (STEM) courses, Springer *et al.* (1999) found that this approach had a significant positive effect on achievement, persistence, and attitude. In addition, the impact of cooperative learning on achievement was consistent regardless of the STEM fields involved, the method used to assign groups, group composition (male/female), or the type of group learning that was used.

Although cooperative learning is generally considered an effective teaching strategy, not all reviews of this approach are completely positive. For example, reviews of problem-based learning (PBL) curricula at medical schools, a version of cooperative learning that asks students to work in small groups to solve problems, have found that these programs have had relatively little impact on student achievement or problem-solving skills. In fact, although PBL appeared to promote independent and deeper study skills, in some cases it appeared to have a negative impact on content learning (Albanese and Mitchell, 1993; Berkson, 1993; Vernon and Blake, 1993; Colliver, 2000; Colliver *et al.*, 2003). Despite these potential drawbacks, however, students generally viewed the use of PBL in medical school programs favorably (Albanese and Mitchell, 1993).

Most research examining cooperative learning and achievement in science courses has examined the effectiveness of this strategy in relatively small classes (<100 students). Few studies have looked at large university courses that can enroll hundreds of students in a single section. Furthermore, schools that have used cooperative learning in large classes have often relied on additional classroom support (Posner and Markstein, 1994; Klionsky, 1998; Udovic *et al.*, 2002; Anderson *et al.*, 2005). Unfortunately, many schools do not have the resources to hire additional personnel. Also, although cooperative learning had a positive impact on student learning in these studies, it cannot be ruled out that students may have performed better in these classes, in part, because of the additional resources that were available.

Results from studies in which cooperative learning was incorporated into large classes without additional personnel is mixed. Either 1) no effect on student achievement was seen (Pukkila, 2004), 2) the impact of these activities was lower than detected in smaller classes (Hake, 1998), or 3) it was difficult to distinguish between individual and group performance (Hanshaw, 1982; Rao and DiCarlo, 2000; Zimbardo *et al.*, 2003). One exception was a study by Ebert-May *et al.* (1997) in which formal cooperative groups were incorporated into moderately large class sections (140 students). At the end of the course, students in the cooperative learning sections had higher self-efficacy about doing, analyzing, and explaining science than students in traditional classes of similar size. The cooperative learning students also performed just as well as the traditional class on questions dealing with content knowledge. A subsequent experiment examining cooperative learning in a very large class (450 students) using informal groups found that student participation, attendance, and satisfaction with the course increased but did not examine student performance (Ebert-May *et al.*, 1997).

Research Question

In this study we wanted to examine the impact of cooperative learning activities, designed so they could be administered by just one individual even in a very large course (>250 students per section), on student learning outcomes. We hypothesized that cooperative learning would improve retention and understanding of course material. To determine if this was the case, we compared the performance of students taught using a cooperative learning or a traditional lecture format on traditional exams. In addition, we examined students' perceptions of the activities used in the cooperative learning sections using surveys and attendance records.

MATERIALS AND METHODS

Participants and Classroom Setting

All of the participants in the study were enrolled in a first semester Introductory Biology course taught in the spring of 2004 and spring of 2005. The goal for this course is for students to develop a conceptual understanding of basic biology and to be able to apply their understanding to solve biological questions. Topics covered during the semester included biochemistry, cell biology, reproduction, genetics, and evolution. This Introductory Biology course is designed for nonscience majors and, from previous surveys, we know that the vast majority of students who take the class do so primarily to meet a university graduation requirement. Approximately two-thirds of the students (68%) were in their first 2 yr of study and more than 71% report that previously they had attended at least one other college-level science course.

Each section of Introductory Biology had an initial enrollment of more than 300 students, with the exception of one section in the spring of 2004 that had a beginning enrollment of 289. During the semester, the sections met for three 50-min sessions or two 75-min sections per week (two sections of each). All lecture sections met in the same classroom equipped with stadium-style seating. Approximately two-thirds of the students in the lecture portion of the course also enrolled in an optional 2-h weekly laboratory section.

The students were free to enroll in any section of the course offered each semester. No information regarding the format of instruction was provided before the start of classes. On the first day of class, the researchers explained the goals of the study to the students and requested their participation. Students were informed about the requirements to take part in the study and were given the option to sign consent forms. No incentives were offered for participation.

Classroom Instruction and Activities

The Introductory Biology classes were taught by two instructors who each have 10+ yr experience teaching introductory biology and 5–10 yr teaching the specific course used for this study. Both instructors were working to incorporate cooperative learning strategies into their classes for the previous 2 yr. In the spring of 2004, each instructor taught two sections of the course, one using a traditional lecture format and one using a format that incorporated group activities and cooperative learning strategies. In the spring of 2005, the instructors again taught two sections of the course each but used a cooperative learning format exclusively.

In the cooperative learning sections, the instructors used the random number generator function in Microsoft Excel (Microsoft, Redmond, WA) to assign students to 50 groups of 6–8 students during the second week of class. Special needs students and other individuals who had a valid reason for sitting in a particular location in the classroom were asked to contact the instructor in the first week of class. Students were given a handout with their group assignment

and a seating chart the first week that group activities were used. Students were asked to sit with their groups in designated areas for the remainder of the course. No formal instruction was given on how to carry out group activities. However, the instructor pointed out at the beginning of the semester that each person evaluates his or her peers at the end of the term for how well they contributed to the functioning of the group. Contributing to group function was defined to include the following: preparation for class, willingness to speak up during group activities as well as to ask and answer each others' questions, and to show respect and consideration for other group members and their ideas. In the traditional lecture sections no groups were assigned, and students were allowed to choose their own seats.

Instruction in the cooperative learning sections was carried out using short PowerPoint (Microsoft) lectures interspersed with multiple-choice questions posed to the students. No points were associated with these questions, and at least two or more questions were posed each class period. Most questions were designed to require the students to apply their understanding of a concept to a problem or situation. (For examples of the questions used during class, see Appendix A in the Supplemental Materials.) Students were given time to discuss possible answers in their groups, after which a single group answer was submitted using the Interwrite Personal Response System "clickers" that were given to each group at the beginning of the class. Pickup and return was facilitated using a box in which the clickers could be stored in numbered slots corresponding to the different groups. After the groups submitted their responses, the instructor displayed a summary of their answers to the class. The instructor then solicited the class to explain why they chose particular answers in order to reveal students' thought processes and to uncover misconceptions.

Several topics were covered during the semester using a pseudocase format that organized the concepts taught around a central story line. Example cases include the following:

- A murder in which the students need to understand DNA structure and replication, PCR, gel electrophoresis, and DNA fingerprinting to convict or release a suspect (Brickman, 2005).
- A malpractice trial involving a case of Down syndrome in which the students must understand chromosome structure and function, mitosis, and meiosis in order to determine the guilt or innocence of a doctor.
- A genetic testing case in which the students must understand DNA structure and function, transcription and translation, and the nature of mutations in order to determine if a patient carries a mutation for the disease phenylketonuria.

In addition to their textbook, students were provided with a study guide through WebCT courseware (Blackboard, Phoenix, AZ) that contained an outline of major concepts including images and animations.

Instruction in the traditional lecture classes covered the same topics using the same PowerPoint lectures and course materials as in the cooperative learning sections. The primary difference between the lecture and cooperative learning classes was that instead of initially encouraging the students to work on questions in groups, the instructor worked through each problem in front of the class. In addition, students in the traditional section were given individual quizzes only and did not participate in group quizzes (see below).

Assessment of Content Learning

We examined the relative impacts of traditional lecture and cooperative learning approaches on student learning using the sections of the course taught in the spring of 2004. To examine whether the cooperative learning (treatment groups) and traditional lecture (control groups) sections were equivalent before the study, the students were given a survey at the beginning of the semester asking them to provide demographic information regarding their gender, self-reported grade point average (GPA), and previous

college science courses taken. The survey also asked 15 multiple-choice questions corresponding to topics to be covered during the upcoming semester and similar in format to content questions on typical exams (pretest).

Because of the large size of the classes and limited instructional personnel, student performance in the lecture portion of the course was determined using machine-graded, multiple-choice exams including six quizzes (15 questions each), a cumulative midterm exam (50 questions), and a cumulative final exam (80 questions). The quizzes and midterm exams in the separate sections used different questions, but the instructors strove to ensure that the tests dealt with the same concepts and were of the same level of difficulty. We administered the same final exam to all sections.

In the lecture-only sections, students completed each quiz just one time. In the cooperative learning sections, the students completed each quiz individually, turned in the exam, and then retook the same quiz with their groups (modified from Michaelsen *et al.*, 1985). The students needed to discuss each question and come to a consensus regarding the appropriate answer in order to fill out a single answer sheet that they would turn in as a group. The group quiz scores could contribute up to one-third of each individual's total quiz score.

Test periods were the only times during which additional personnel were used to support the instructor. This was true for both the cooperative learning and traditional lecture sections. In both cases, graduate students served as proctors to help distribute and collect the exams as well as monitor the students while the exam was in progress.

To encourage participation and prevent "freeloading" during the group tests and classroom activities, each student was asked to evaluate the other members of their group on how well they contributed to group functioning (modified from Herreid, 2001; see Appendix B in the Supplemental Material). Peer evaluation results were used to determine how many group quiz points (see below) each student would receive. If a student received an average score of 80% from their peers, that student would receive 80% of their group's test points. The instructor reserved the right to overrule any peer evaluation score if it appeared to be inaccurate or inappropriate such as when evaluations have been biased because of personality conflicts.

We measured student achievement in five ways: 1) individual performance on in-class exams excluding the final exam, 2) individual performance on a cumulative final exam, 3) performance on questions requiring factual recall, 4) a conceptual understanding of course material, and 5) relative improvement from the beginning to the end of the semester. To determine performance on questions requiring different levels of cognition (Areas 3 and 4), we categorized the questions from the final exam according to the type of knowledge demonstrated in agreement with Bloom's taxonomy (Krathwohl, 2002). The questions on the final exam were ranked by the two instructors for the course. Disagreement on the categorization of a question was resolved by discussion and re-evaluation of the specific questions involved. Of the 80 questions on the final exam, 32 demonstrated simple recall of information and 48 demonstrated comprehension or the ability to apply knowledge to a novel situation. Relative improvement from the beginning to the end of the semester (Area 5) was determined by subtracting each student's percent score on the pretest administered on the first day of class from the percent score on the cumulative final exam.

Student Attendance and Perceptions

To determine if the instructional method used affected classroom attendance, we photographed the sections taught in the spring of 2004 multiple times during the semester (6 times by instructor 1 and 18 times by instructor 2). Students in each photograph were visually counted.

At the end of the spring 2004 semester, both instructors received numerous favorable comments regarding the instructional approach used in the cooperative learning sections. To more accu-

rately measure students' views regarding this approach, we administered a survey to four sections of the Introductory Biology course at the end of the spring 2005 semester. All four of these sections used cooperative learning in a format identical to that of the treatment sections in spring of 2004.

Data Analysis

Difference before the Study. To evaluate the difference among students of different sections before this study, we compared the students' self-reported GPA, the number of previous college science courses taken (ScienceCourse hereafter), gender composition, and biological knowledge at the beginning of the semester (referred to as Pretest hereafter) between the treatment and the control groups. Self-reported GPA was used as a measure of the students' overall academic achievement and accurately reflects students' actual GPA (Cassady, 2001). Because of limitations inherent to the machine-graded forms used to collect survey data, prior GPA was divided into five levels representing the numerical values of 0–1.6, 1.7–2.2, 2.3–2.8, 2.9–3.4, and 3.5–4.0, respectively. GPA and ScienceCourse were analyzed using a nonparametric Kruskal-Wallis test (using Proc NPAR1WAY in SAS, 1999) because the distributions of these two measurements do not conform to normality and none of the commonly used transformation methods was able to achieve the normality assumption required by more conventional analytical methods such as ANOVAs. Gender, recorded as 0 for females and 1 for males, was analyzed using a logistic regression approach (Proc CATMOD in SAS, 1999). The pretest scores were analyzed using an ANOVA approach (Proc GLM in SAS, 1999).

Even though the students were self-enrolled into these sections and we did not expect the sections to differ from one another before the study, results of the above analyses did reveal some significant differences in regards to GPA and previous science credits (see more details in *Results*). To evaluate if any differences observed between the treatment and control groups were due to these pre-existing differences, we included gender, GPA, and previous scientific courses as covariates in the analyses of student performance. We carried out an ANOVA to analyze whether instructor, experimental treatments, and their interaction terms affected the student attendance in the sections used in this study.

Because attendance was determined as the percentage (%) of the students present on each of the days in which photographs were taken, the data were transformed using the following equation: $\text{attendance} = \arcsin[\text{square root}(\text{attendance in } \%/100)]$. This transformation helped satisfy the normality assumption of ANOVA. In the tables, we report the average attendance in % to allow easy assessment for the readers.

Performance Measurements. To evaluate the effect of treatment on performance measurements, we carried out an analysis of covariance (ANCOVA hereafter) for each measurement. Treatment (cooperative learning vs. control), Instructor, and their interaction were included as the main effects in these analyses along with the covariates described in the previous paragraph. The residuals of these analyses were normally distributed.

In addition to the above measurements of student performance during the semester, we also wanted to examine whether "improvement" in students' biological knowledge during the semester was influenced by the type of instruction they experienced. To achieve this goal, we carried out a similar ANCOVA analysis as the one described in the last paragraph but used the difference between final exams scores and the pretest scores (both evaluated as % correct answers) as the response variable.

RESULTS

Differences before the Study

No significant difference was seen between the experimental and control sections in regards to their performance on the

pretest, indicating that students in the different sections had similar levels of biology knowledge at the beginning of the class. The individual sections did differ somewhat in regards to their gender ratios. All sections were female biased with treatment section 1 (taught by instructor 1) having the highest female ratio (71.8%), whereas the control section 1 had the lowest (53.7%). When examined together, the treatment sections had a slightly higher percentage of females (64.7%) than the controls (57.8%; Table 1).

The treatment and control sections differed significantly in regards to their self-reported GPA ($\chi^2 = 18.54$, d.f. = 1, $P < 0.001$). This difference was due mainly to a particularly low GPA for the control section 1 (3.80 vs. ≥ 4.2 in the other three sections). Students in the control group also took slightly more science courses than the treatment group before this study, but the two groups did not differ significantly ($\chi^2 = 3.69$, d.f. = 1, $P = 0.055$; Table 1).

Student Performance

Student performance on questions demonstrating comprehension of course material was significantly affected by the Instructor (Table 2). Sections taught by instructor 2 tended to have higher scores on conceptual questions than sections taught by instructor 1 (Figure 1D). In addition, there was a significant interactive effect between Instructor and Experimental Treatment for three of the five performance measures examined (Table 2). More specifically, the control section taught by instructor 1 performed significantly worse than the treatment section on the final exam as well as on questions requiring factual knowledge and comprehension. The reverse trend was seen for these measures in the treatment and control sections taught by instructor 2 though the differences between these two sections were not significant (Figure 1B, C and D).

When we compared the final exam and pretest scores, we found that students in the cooperative learning sections showed significantly more improvement than students in the control sections (Table 2 and Figure 1E). On average, the control sections showed a 44.7% improvement in score, whereas the treatment sections improved by 47.6%. This effect was consistent across instructors as indicated by a nonsignificant interaction between the instructor and the experimental treatment.

Lastly, student performance on all five measures examined was significantly influenced by prior GPA (Table 2).

Student Attendance and Perceptions

Average classroom attendance was significantly higher in the cooperative learning sections relative to the controls ($P < 0.001$, Table 1). Attendance was particularly low in control section 1. This section also showed a substantially higher dropout rate than the other three sections (11.9 vs. $< 5\%$).

When students in the spring 2005 semester were surveyed regarding the instructional methods used for cooperative learning, the response was highly favorable (Table 3). Students indicated that the group activities and tests helped them understand the course material better. More than 92% of the students indicated that the cooperative learning strat-

Table 1. Gender ratio, prior GPA, number of other science courses taken, attendance, and withdrawals

	Four individual sections				Combined treatment and control sections		
	Instructor 1		Instructor 2		Combined		Difference ^a
	Treatment	Control	Treatment	Control	Treatment	Control	
Gender (% female)	71.8	53.7	58	69	64.7	57.8	
GPA ^b	4.26 (0.05)	3.80 (0.06)	4.27 (0.05)	4.20 (0.05)	4.26 (0.04)	4.02 (0.04)	$P < 0.001$
Other science	0.88 (0.06)	1.11 (0.07)	1.01 (0.06)	1.03 (0.06)	0.94 (0.04)	1.07 (0.04)	$P = 0.055$
Pretest (out of 15 pts)	4.83 (0.13)	4.85 (0.17)	5.02 (0.13)	4.55 (0.19)	4.94 (0.13)	4.66 (0.10)	$P = 0.149$
Attendance (%)	70.2 (3.04)	46.0 (3.04)	69.6 (1.81)	59.1 (1.75)	69.9 (1.74)	52.6 (1.74)	$P < 0.001$
Withdrawals (%)	3.4	11.9	5.0	4.0	4.2	7.7	

Values are means with the SEMs in parentheses.

^a P values are for the difference between combined Treatment and Control sections.

^b GPA was divided into 5 categories: 1 = 0–1.6; 2 = 1.7–2.2; 3 = 2.3–2.8; 4 = 2.9–3.4; and 5 = 3.5–4.0.

gies should be used in other classes. When asked to respond to the prompt, “Feel free to comment on the course or offer suggestions on how the course could be improved” students most frequently remarked that the cooperative learning activities helped them to understand the material better and that group work made the class more personal and enjoyable. They also indicated that using the clickers encouraged them to interact more effectively with their groups and that the problems they were asked to work on forced them to think about and apply the concepts they were trying to learn. Several students indicated that the group tests encouraged them to prepare more rigorously than they might have otherwise. Students who viewed cooperative learning less favorably appeared to do so primarily because they did not like being randomly assigned to groups or indicated that their groups were dysfunctional in some manner. Others did not like evaluating their peers or being evaluated by them. Very few students indicated that they did not like group work or found it unhelpful (Table 4).

DISCUSSION

The purpose of this study was to examine the impact of cooperative learning strategies on student achievement and attitudes in a very large introductory biology class. In particular, we wanted to determine if activities designed such that a single instructor could administer them would help students learn the course material better than a traditional lecture format.

We found that, although overall test scores did not differ significantly between the treatment and control groups, students in the cooperative learning sections showed greater improvement from the beginning to the end of the semester than students taught using a traditional lecture approach. This finding suggests that cooperative learning activities can improve student outcomes even in very large classes. However, the difference in improvement observed was fairly small (a 3.2% difference in the improvement), indicating that the impact of cooperative learning was relatively weak though statistically significant. In addition, instructor and treatment interacted strongly for three of the five performance measures examined, suggesting that these results

Table 2. Student performance on questions demonstrating comprehension of course material

Sources of variation	d.f.	ANCOVA of performance measurements ^a									
		Test average		Final exams		Factual questions		Conceptual questions		Improvement ^b	
		F	P	F	P	F	P	F	P	F	P
Main effects											
Instructor	1	0.30	0.58	0.09	0.76	0.24	0.63	5.85	0.02	0.64	0.42
Experiment	1	2.98	0.08	1.22	0.27	0.89	0.35	1.03	0.31	5.92	0.01
Instructor × Experiment	1	2.32	0.13	11.03	<0.001	11.20	<0.001	8.15	0.004	0.42	0.52
Covariates											
GPA	1	4.05	0.04	189.17	<0.001	156.89	<0.001	148.00	<0.001	26.61	<0.001
Gender	1	0.55	0.46	0.18	1.79	6.24	0.01	0.16	0.69	1.30	0.25
Other sciences	1	0.49	0.48	2.70	0.10	2.61	0.11	2.04	0.15	0.11	0.74

Values reported are the F statistics and corresponding P values for significance tests in the analyses of covariance (ANCOVA).

^a ANCOVA revealed a strong interaction between Instructor and Experimental treatments on three of the five performance measurements taken. The Control and Treatment sections taught by Instructor 1 differed significantly but not the sections taught by Instructor 2. Prior GPA showed a strong effect on overall performance of the students in this study.

^b Experimental treatment had a significant effect on students' improvement in their performance.

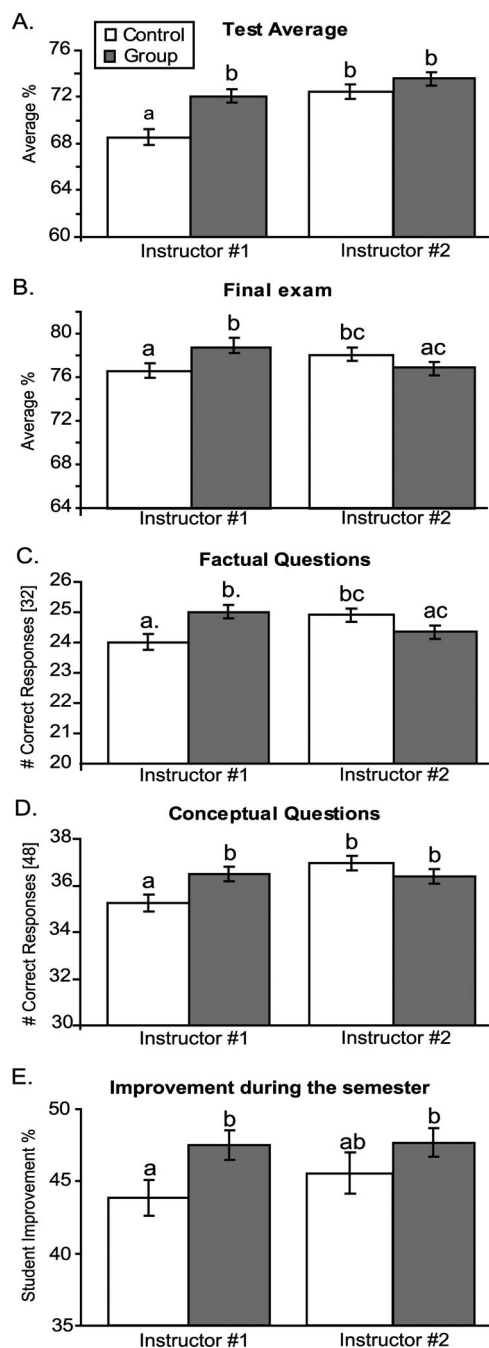


Figure 1. No significant difference was observed between the treatment and control sections in regards to their (A) test average, (B) final exam, (C) factual questions, or (D) conceptual questions scores. The treatment sections demonstrated significantly more improvement from the beginning to the end of the semester (E) than the control sections ($P = 0.01$, Table 2). The letters above each column indicate results of pairwise comparisons between sections. Sections that share the same letter do not differ significantly from one another. Bars, ± 1 SE for the least-square means for each section derived from the ANOVA.

provide preliminary evidence only until the study is replicated with a larger cohort of instructors.

A factor that may have influenced the outcome of this study is the manner in which content learning was measured. Cooperative learning activities are proposed to encourage a deeper understanding of course material and promote higher-level thinking skills (Cooper, 1995; Anderson *et al.*, 2005; Cortright *et al.*, 2005). As is common in many large science classes, personnel limitations demanded that the classes examined in this study use multiple-choice exams as a primary means of evaluating students' understanding of course content. However, Darling-Hammond *et al.* (1995) have argued that multiple-choice exams are not an effective tool for evaluating higher-order thinking skills, and it is possible that greater differences existed among students in our study than could be detected by our assessment method. Indeed, the effect of PBL, as well as other instructional approaches thought to promote higher levels of cognition, has been found to depend on the type of assessment used (Boyles *et al.*, 1994; Hand *et al.*, 2002; Gijbels *et al.*, 2005). Future experiments that examine cooperative learning should address this issue by incorporating questions that more effectively measure higher-order thinking skills into the assessments used.

When asked to comment on instruction of the cooperative learning sections, students viewed the methods used very positively. Most felt that this approach helped them to understand the material better as well as made the course more interesting, personal, and enjoyable. The students also appeared to sense a greater level of support for their learning through interactions with their peers. Very few students felt that group work was unhelpful. This finding is noteworthy because student attitude is considered to be an important component of overall engagement (National Research Council, 2004). Students who find a course more interesting and enjoyable and who feel they are supported are more likely to invest more effort into the course, potentially leading to improved learning outcomes (Marks, 2000; Handelsman *et al.*, 2005; Robbins *et al.*, 2006). This study suggests that, even in very large classes and with minimal personnel, it is possible to implement cooperative learning activities and that students find groups enjoyable and useful. However, additional studies are necessary in order to determine if these activities improve student engagement relative to a traditional lecture format.

Table 3. Response to questions by students in the spring 2005 semester who were taught using a cooperative learning approach

Question ^a	Response
1. The group tests helped me correct my confusion and misunderstanding about the material.	1.3
2. My understanding of the material increased due to my group's discussion of the quiz questions.	1.2
3. Would you recommend group testing be used in your other classes?	Yes (92%) No (8%)

^a Questions 1 and 2 were scored using a Likert Scale ranging from strongly agree (2) to strongly disagree (-2).

Table 4. Comments submitted by students taught using a cooperative learning approach

Comment category	No. of comments
Helped understand the course material better	61
Made the class more personal	19
Make the class more enjoyable or engaging	15
Encouraged them to study the material better	10
Improved their grade	5
Were in a bad/dysfunctional group	28
Didn't like random assignment to groups	14
Groups activities didn't help	13
Don't like peer evaluations	10
Disapprove of group work	5
Groups don't work in large classes	1
Examples of favorable comments	
<ul style="list-style-type: none"> ... with a class this size, the groups are a MUST. At first I was a little apprehensive about being in a group since I'm usually a very independent worker, but it turned out to be great. For such a big class, it ended up being one of the most personal I've had of that size. I think that having the group tests is a good idea. It encouraged me to study more for the class out of a desire to not let my group down. Group testing is definitely a plus. I believe that during group tests or discussion that if you don't understand a concept too well that hearing the explanation comes from another group member might help you to understand it better than if it came from the teacher or the book. Plus fellow students tend to be less intimidating when it comes to asking questions. I really enjoyed the group set up. I have never been in a class like this before, but I found that it really helped me understand what I was confused about because, in a group setting, your classmates know you better and know what you struggle with and what your strengths are. The groups also made the class more fun because I was put in a situation to make new friends. This was one of my favorite classes because we were able to be so involved in lecture. Normally it is just the instructor at the front of the room talking the entire time, but with activities and the clickers students were able to participate in class and relate the material to real life situations. 	
Examples of unfavorable comments	
<ul style="list-style-type: none"> I don't like the group evaluation idea—several people in my group participated as little as humanly possible, even skipping tests, and I fear that next week they will give me a low evaluation score out of mere spite; they weren't even in class to see my contributions. I didn't like the group thing because since they were randomly assigned, I happened to have a group in which none of us really understood any of the material, whereas a friend of mine was in a group with a couple of people who were brilliant at biology. It's random, but it's only beneficial for some people. I think the group tests were not all that beneficial to my overall comprehension of the course material. I feel those people who knew the answers consistently knew them for each test and just told the rest of the group what those were instead of more of a collaborative group effort. As for the student evaluations, I think they could be done without. I personally don't like the idea of having another person determining my grade. 	

Students in the spring 2005 semester were asked to respond to the following prompt: "Feel free to comment on the course or offer suggestions on how the course could be improved." A total of 202 students submitted comments of which 181 addressed the cooperative learning aspect of the course.

Cooperative learning activities potentially have a reliable and robust impact on achievement in large classes. How might this be accomplished? Slavin (1996) argued that cooperative activities are most likely to promote learning gains when students are rewarded for their participation and when group performance is measured as the sum of each member's individual performance rather than through a single group product. When a single product is submitted by the entire group, there is little incentive for the group to ensure that each member fully understands the concepts involved, and it is possible for a single individual to contribute most of the work/knowledge necessary for everyone to obtain a good grade. This type of behavior is especially likely with high-stakes assessments (James, 2006), and evidence that it may have occurred in this study can be seen in the student surveys. As one student commented, "I feel those people that knew the answers consistently knew them for each test and just

told the rest of the group what those were instead of more of a collaborative group effort."

The group activities we used may have had a more significant impact if rewards for cooperative work were based on the average of each member's individual performance rather than on a consensus answer. This would have provided a greater incentive for group members to explain the concepts to one another and help each member perform to his or her maximum ability. One potential problem with this approach, however, is that it may also lead to greater dissatisfaction among some students because their personal grades would become partially dependent on the performance of others.

Another way in which the cooperative learning approach used in this study might be improved is to provide students with training in effective group interactions and learning strategies. The debate of ideas within groups, referred to as argument, ideally promotes learning by helping students

become more aware of their own level of understanding (metacognition) and by forcing them to reconsider their ideas in response to alternative and potentially conflicting views (social constructivism). College students frequently display relatively weak metacognitive and argument skills (Kuhn, 1993; Pintrich, 2002), thus these strategies need to be explicitly taught in order for students to use them effectively (Penrose, 1992; Moore, 1993; Mason, 1996; Zohar and Nemet, 2002; Kuhn and Udell, 2003; Osborne *et al.*, 2004). It is possible that instructing students on the use of argument and on how and when to use metacognitive strategies would make cooperative learning more effective at promoting learning. In support of this possibility, studies have found that students receiving metacognitive training outperformed students who did not (Meloth and Deering, 1992; Mevarech, 1999; Kramarski *et al.*, 2001; Zohar and Nemet, 2002). Indeed, cooperative groups that did not receive such training performed similarly to students who studied individually (Kramarski and Mevarech, 2003).

CONCLUSIONS

The results of this study indicate that a cooperative learning approach helped students improve their knowledge of course material slightly better than a traditional lecture format. In addition, students viewed cooperative learning activities favorably, raising the possibility that this approach may improve student engagement. These findings suggest that encouraging students to work in small groups and improving feedback between the instructor and the students can help to improve student outcomes even in very large-enrollment classes.

Confirmation of these preliminary results will occur when this experiment is replicated with additional faculty. In addition, students have the potential to increase the impact of cooperative learning in large classes by learning to work together but they need explicit instruction and training about group interactions.

ACKNOWLEDGMENTS

We are grateful to Kathrine Stanger-Hall for helpful comments and discussions regarding the manuscript.

REFERENCES

Albanese, M. A., and Mitchell, S. (1993). Problem-based learning—a review of literature on its outcomes and implementation issues. *Acad. Med.* 68(8), 51–82.

Anderson, W. L., Mitchell, S. M., and Osgood, M. P. (2005). Comparison of student performance in cooperative learning and traditional lecture-based biochemistry classes. *Biochem. Mol. Biol. Educ.* 33(6), 387–393.

Berkson, L. (1993). Problem-based learning—have the expectations been met? *Acad. Med.* 68(10), 79–88.

Bloom, B. S. (1984). The search for methods of group instruction as effective as one-to-one tutoring. *Educ. Leadership* 41(8), 4–17.

Boyles, M. P., Killian, P. W., and Rileigh, K. K. (1994). Learning by writing in introductory psychology. 75(1), 563–568.

Bransford, J., Brown, A., and Cocking, R. (1999). *How People Learn: Brain, Mind, Experience, and School*, Washington, DC: National Academy Press.

Brickman, P. (2005). The case of the Druid Dracula. *J. Coll. Sci. Teach.* 36(2), 48–53.

Cassady, J. C. (2001). Self-reported GPA and SAT: a methodological note. *Pract. Assess. Res. Eval.* 7(12), 1–6.

Colliver, J. A. (2000). Effectiveness of problem-based learning curricula: research and theory. *Acad. Med.* 75(3), 259–266.

Colliver, J. A., Feltovich, P. J., and Verhulst, S. J. (2003). Small group learning in medical education: a second look at the Springer, Stanne, and Donovan meta-analysis. *Teach. Learn. Med.* 15(1), 2–5.

Cooper, M. M. (1995). Cooperative learning—an approach for large enrollment courses. *J. Chem. Educ.* 72(2), 162–164.

Cortright, R. N., Collins, H. L., and DiCarlo, S. E. (2005). Peer instruction enhanced meaningful learning: ability to solve novel problems. *Adv. Physiol. Educ.* 29(2), 107–111.

Darling-Hammond, L., Aness, J., and Falk, B. (1995). *Authentic assessment in action: studies of schools and students at work*, National Center for Restructuring Education, Schools, and Teaching, Teachers College, New York: Columbia University.

Ebert-May, D., Brewer, C., and Allred, S. (1997). Innovation in large lectures—teaching for active learning. *Bioscience* 47(9), 601–607.

Gijbels, D., Dochy, F., Van den Bossche, P., and Segers, M. (2005). Effects of problem-based learning: a meta-analysis from the angle of assessment. *Rev. Educ. Res.* 75(1), 27–61.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66(1), 64–74.

Hand, B., Prain, V., and Wallace, C. (2002). Influences of writing tasks on students' answers to recall and higher-level test questions. *Res. Sci. Educ.* 32(1), 19–34.

Handelsman, J. *et al.* (2004). Scientific teaching. *Science* 304(5670), 521–522.

Handelsman, M. M., Briggs, W. L., Sullivan, N., and Towler, A. (2005). A measure of college student course engagement. *J. Educ. Res.* 98(3), 184–191.

Hanshaw, L. G. (1982). Test anxiety, self-concept, and the test performance of students paired for testing and the same students working alone. *Sci. Educ.* 66 (1), 15–24.

Herreid, C. F. (2001). When justice peeks. *J. Coll. Sci. Teach* 30(7), 430–433.

James, M. C. (2006). The effect of grading incentive on student discourse in peer instruction. *Am. J. Phys.* 74(8), 689–691.

Johnson, D. W., Johnson, R. T., and Smith, K. A. (1998). Cooperative learning returns to college. *Change* 30(4), 26–35.

Klionsky, D. J. (1998). A cooperative learning approach to teaching introductory biology. *J. Coll. Sci. Teach.* 27(5), 334–338.

Kramarski, B., and Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: the effects of cooperative learning and metacognitive training. *Am. Educ. Res. J.* 40(1), 281–310.

Kramarski, B., Mevarech, Z. R., and Lieberman, A. (2001). Effects of multilevel versus unilevel metacognitive training on mathematical reasoning. *J. Educ. Res.* 94(5), 292–300.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: an overview. *Theor. Pract.* 41(4), 212–218.

Kuhn, D. (1993). Science as argument—implications for teaching and learning scientific thinking. *Sci. Educ.* 77(3), 319–337.

Kuhn, D., and Udell, W. (2003). The development of argument skills. *Child. Dev.* 74(5), 1245–1260.

- Lord, T. R. (1994). Using cooperative learning in the teaching of high-school biology. *Am. Biol. Teach.* 56(5), 280–284.
- Marks, H. M. (2000). Student engagement in instructional activity: patterns in the elementary, middle, and high school years. *Am. Educ. Res. J.* 37(1), 153–184.
- Mason, L. (1996). An analysis of children's construction of new knowledge through their use of reasoning and arguing in classroom discussions. *Int. J. Qual. Studies Educ.* 9(4), 411–433.
- Meloth, M. S., and Deering, P. D. (1992). Effects of two cooperative conditions on peer-group discussions, reading-comprehension, and metacognition. *Contemp. Educ. Psychol.* 17(2), 175–193.
- Mevarech, Z. R. (1999). Effects of metacognitive training embedded in cooperative settings on mathematical problem solving. *J. Educ. Res.* 92(4), 195–205.
- Michaelsen, L. K., Watson, W. E., and Shrader, C. B. (1985). Informative testing: a practical approach for tutoring with groups. *Org. Behav. Teach. Rev.* 9(4), 19–33.
- Moore, R. (1993). Does writing about science improve learning about science? *J. Coll. Sci. Teach.* 22(4), 212–217.
- National Research Council. Institute of Medicine (2004). *Engaging Schools: Fostering High School Students' Motivation to Learn*, Washington, DC: National Academies Press.
- Osborne, J., Erduran, S., and Simon, S. (2004). Enhancing the quality of argumentation in school science. *J. Res. Sci. Teach.* 41(10), 994–1020.
- Penrose, A. M. (1992). To write or not to write—effects of task and task interpretation on learning through writing. *Writ. Commun.* 9(4), 465–500.
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theor. Pract.* 41(4), 219–225.
- Posner, H. B., and Markstein, J. A. (1994). Cooperative learning in introductory cell and molecular biology: Binghamton's trial in cooperative learning—a startling success. *J. Coll. Sci. Teach.* 23(4), 231–233.
- Powell, K. (2003). Spare me the lecture. *Nature* 425(6955), 234–236.
- Pukkila, P. J. (2004). Introducing student inquiry in large introductory genetics classes. *Genetics* 166(1), 11–18.
- Rao, S. P., and DiCarlo, S. E. (2000). Peer instruction improves performance on quizzes. *Adv. Physiol. Educ.* 24(1), 51–55.
- Robbins, S. B., Allen, J., Casillas, A., Peterson, C. H., and Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *J. Educ. Psychol.* 98(3), 598–616.
- SAS (1999). *SAS/STAT User's guide*, Cary, NC: SAS Institute.
- Slavin, R. E. (1996). Research on cooperative learning and achievement: what we know, what we need to know. *Contemp. Educ. Psychol.* 21(1), 43–69.
- Springer, L., Stanne, M. E., and Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Rev. Educ. Res.* 69(1), 21–51.
- Udovic, D., Morris, D., Dickman, A., Postlethwait, J., and Wetherwax, P. (2002). Workshop biology: demonstrating the effectiveness of active learning in an introductory biology course. *Bioscience* 52(3), 272–281.
- Vernon, D.T.A., and Blake, R. L. (1993). Does problem-based learning work—a meta-analysis of evaluative research. *Acad. Med.* 68(7), 550–563.
- Webb, N. M., Troper, J. D., and Fall, R. (1995). Constructive activity and learning in collaborative small-groups. *J. Educ. Psychol.* 87(3), 406–423.
- Zimbardo, P. G., Butler, L. D., and Wolfe, V. A. (2003). Cooperative college examinations: more gain, less pain when students share information and grades. *J. Exp. Educ.* 71(2), 101–125.
- Zohar, A., and Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *J. Res. Sci. Teach.* 39(1), 35–62.