CBE—Life Sciences Education
Vol. 8, 182–192, Fall 2009

Article

Replacing Lecture with Peer-led Workshops Improves Student Learning

Ralph W. Preszler

Department of Biology, New Mexico State University, Las Cruces, NM 88003-8001

Submitted January 9, 2009; Revised April 20, 2009; Accepted April 22, 2009
Monitoring Editor: Laura L. Mays Hoopes

Peer-facilitated workshops enhanced interactivity in our introductory biology course, which led to increased student engagement and learning. A majority of students preferred attending two lectures and a workshop each week over attending three weekly lectures. In the workshops, students worked in small cooperative groups as they solved challenging problems, evaluated case studies, and participated in activities designed to improve their general learning skills. Students in the workshop version of the course scored higher on exam questions recycled from preworkshop semesters. Grades were higher over three workshop semesters in comparison with the seven preworkshop semesters. Although males and females benefited from workshops, there was a larger improvement of grades and increased retention by female students; although underrepresented minority (URM) and non-URM students benefited from workshops, there was a larger improvement of grades by URM students. As well as improving student performance and retention, the addition of interactive workshops also improved the quality of student learning: Student scores on exam questions that required higher-level thinking increased from preworkshop to workshop semesters.

INTRODUCTION

Learning activities designed to include peer and instructor discussion increase interactive student engagement and learning (Hake, 1998; Knight and Wood, 2005). This interactive engagement can occur within the constraints of large lecture sessions through the use of creative teaching strategies (Ebert-May *et al.*, 1997; Klionsky, 2001; Udovic *et al.*, 2002; Allen and Tanner, 2005; Armstrong *et al.*, 2007; Walker *et al.*, 2008) as well as technologies (Jensen *et al.*, 2002; Novak, 2002; Wood, 2004; Preszler *et al.*, 2007; Crossgrove and Curran, 2008). Alternatively, breaking large lectures into smaller workshop sections in which peer instructors facilitate cooperative group work dramatically increases the interactive engagement of students; however, this change in course structure is associated with an equally dramatic loss of lecture. In this study, we assess the net change in student learning associated with sacrificing one-third of our lectures to add weekly peer-facilitated workshops. Although faculty lecturers are best positioned to describe the course information in the context of a larger knowledge base, they may not provide the most accessible entry points to the discipline for many students. The large content-specific cognitive gaps between instructors and students limit students' abilities to assimilate information presented by their instructors and as a result, limit the students' abilities to generate knowledge. Undergraduates leading workshop sessions are closer to the students' zone of proximal development, the region of students' potential knowledge gain, and can therefore communicate and facilitate learning through channels not available to faculty (Tien *et al.*, 2002). In addition to the exchange of ideas between peer leaders and their students about course-specific knowledge and general learning strategies, peer leaders also enrich the overall flow of information within the course as they interpret the instructor's expectations and presentations to students and as they explain student's responses and potential misconceptions to instructors. Peer leaders are typically upper-division undergraduates who have successfully completed the course (Arendale, 1997; Tien *et al.*, 2002; Hockings *et al.*, 2008). Some programs are effectively facilitated by postbaccalaureate students who may not have been students in the course in its current form but who have been trained to facilitate cooperative learning (Fullilove and Treisman, 1990; Rath *et al.*, 2007). This suggests that perhaps the most important benefit of peer-facilitated workshops is that they provide a learning environment that promotes communication and coopera-

DOI: 10.1187/cbe.09–01–0002
Address correspondence to: Ralph W. Preszler (rpreszle@nmsu.edu).

tive learning among students (Johnson *et al.*, 1998; Springer *et al.*, 1999).

The Department of Biology at New Mexico State University (NMSU), with support from the Howard Hughes Medical Institute's (HHMI) Undergraduate Science Education Program, transformed our introductory biology lecture course into a course with required peer-facilitated workshops beginning in the spring 2007 semester. In the workshops, peer facilitators help cooperative groups of students work through challenging problems, case studies, and activities that promote the development of general learning skills. Our goals in developing this new course structure were to increase students' general learning and critical-thinking skills, increase course-specific content acquisition, and facilitate more meaningful student learning. We also hoped that peer-facilitated cooperative workshops, while benefitting all students, would be especially effective at improving the performance of underrepresented minority (URM) students. The Department of Education considers NMSU a minority institution that significantly serves Hispanic and Native-American student groups. Across all campuses, 50% of our students are members of ethnic minorities that are underrepresented in the sciences; across all semesters used in this study, 56.6% of the population in our Natural History of Life course are URM students. The development of our goals to improve students' content and general learning skill development, and the use of peer facilitators to do so, was strongly influenced by the Supplemental Instruction (SI) program developed at the University of Missouri–Kansas City. However, our desire to develop a program that impacts all the students, rather than the subset of students electing to participate in supplemental workshops, led to the development of a workshop structure that has more in common with the Peer-Led Team Learning (PLTL) program developed at City College of the City University of New York (Tien *et al.*, 2002; Gafney and Varma-Nelson, 2008). In addition to requiring the participation of all students, our program and most PLTL programs differ from typical SI programs in that faculty are more involved in coordinating our workshops and our workshop sessions are built around challenging problems and case studies written by the faculty course instructors. Although our program shares the critical components of PLTL programs described in Gafney and Varma-Nelson (2008), we did not set out to replicate the PLTL program and our structure does significantly differ from the PLTL model in three respects: 1) our workshops contain an average of 19 students, whereas the PLTL program recommends six to eight students in each session; 2) our weekly workshop sessions last 65 min, whereas PLTL recommends 90- to 120-min sessions; and 3) we chose to have our peer leaders grade students' workshop reports in order to encourage students to take the workshops and their peer leaders seriously, whereas the PLTL program recommends not having peer leaders grade student work to reinforce their role as role models rather than instructors.

Arendale (2004) identified six peer-facilitated cooperative learning programs that have been successfully replicated and evaluated across universities. All of these programs focus on course-specific knowledge acquisition, as well as more general learning skill development. He found that the programs fit into two categories: 1) those that offer activities adjunct to the primary course (accelerated learning groups, structured learning assistance, and supplemental instruction); and 2) programs that are fully integrated into the course structure (Emerging Scholars, PLTL, and video-based Supplemental Instruction). The elective or voluntary nature of adjunct programs reduces the number of students who participate and limits the ability of instructors to refer to and reinforce programmatic activities in lecture. A pilot study at our institution (Preszler, 2006) demonstrated that although elective workshops promoted student learning, a majority of students did not enroll in and so did not benefit from the workshops. Arendale (2004) reports that only one-third of students typically participate in Supplemental Instruction programs. He concluded that programs that are fully integrated into the course structure have the highest likelihood of improving student performance and retention. Although these programs may have the highest potential, their success depends on a high level of sustained institutional support. The institution must be willing to modify course policies and expectations. Also, faculty need to develop the curriculum and keep the course components well integrated. Training and supervision of peer facilitators also is essential in both adjunct and fully integrated programs (Supplemental Instruction training is widely available from the International Center for Supplemental Instruction at The University of Missouri–Kansas City; PLTL training workshops are supported by the PLTL Workshop program at City College of the City University of New York). Although national and institutional programs can contribute to coordinator, faculty, and peer facilitator training, our experience and descriptions of particularly successful programs (e.g., Rath *et al.*, 2007) have convinced us that faculty involvement in the training and supervision of peer facilitators maximizes communication across all levels and maximizes integration of course content and pedagogy between workshops and lectures. External support from HHMI through the Undergraduate Science Education Program, as well as our College of Arts and Sciences at NMSU, was critical for the development and implementation of our workshops that are fully integrated into our course structure. The detailed assessments of the project included in this study will play a critical role in supporting the transition from external to internal support of this course structure, which admittedly requires more resources than a standard lecture format.

Many successful adjunct as well as required peer-led workshops represent an addition to the original course structure (Fullilove and Treisman, 1990; Arendale, 1997; Ramirez, 1997; Rath *et al.*, 2007; Hockings *et al.*, 2008). It is not surprising that when students are required to, or elect to, participate in additional learning activities their performance increases. There are fewer studies that demonstrate that peer-led workshops are more beneficial than a reasonable alternative approach. Tien *et al.* (2002) compared required recitation sessions with required PLTL sessions and found that PLTL resulted in improved grades, retention, and attitude. Although we would have liked to add required workshops to our existing course structure, many students from a variety of majors rely on our course. We could not expect all of these programs to accept additional credit hours for the biology course required by their majors. These interdepartmental considerations made it necessary to reduce the number of weekly lectures so that we could add workshops without increasing the total course credits. This reduction of lectures with the addition of workshops has allowed us to evaluate student learning in two versions of the course that require a similar level of commitment from students. We suspect that many departments with a desire to reform their courses with required workshops face this same dilemma of choosing between course structures rather than adding to a course. Tien *et al*. (2002) demonstrated that replacing tutoring sessions with PLTL workshops increased student exam scores by 17%. They also reported significant improvements in retention and in student attitudes toward the course. Our analyses rely primarily on longitudinal comparisons of student performance in preworkshop semesters to student performance after the replacement of one of our weekly lectures with a peer-led workshop.

METHODS

Course Design

The Natural History of Life (Biology 111) is the first biology course taken by biology majors. All students at NMSU are required to successfully complete at least two General Education Core Curriculum Laboratory Science courses, such as The Natural History of Life. Thus, in addition to being the introductory course for biology majors, this course serves other science majors, majors from the School of Agriculture, Secondary Education majors, and a variety of students who choose to take a nonterminal biology course to satisfy part of their general education requirements. We do offer an alternative, terminal course for students who are confident they do not need a course that will serve as a prerequisite for subsequent biology courses. Natural History of Life introduces students to scientific process, classical and population genetics (molecular genetics is primary covered in a subsequent course), evolution, and ecology (Supplemental Material lists concepts covered in the preworkshop and workshop versions of the course). Throughout the seven preworkshop semesters of the course, as well as the three workshop semesters, the course was composed of a three-credit lecture course (three lectures or two lectures and a workshop) and a one-credit laboratory course that met for 2½ h each week. Most students, 90–95%, were concurrently enrolled in the lecture and laboratory courses. Most of the laboratory activities are student-centered inquiry-based experiments (e.g., Preszler, 2004b).

Preworkshop Course. In this study, we have used student performance in the fall 2003 through fall 2006 semesters, excluding summer sessions, of the Natural History of Life as a baseline preworkshop control; student performance in these control semesters are compared with student performance in three workshop semesters (Table 1 describes the characteristics of each preworkshop and workshop semester of the course used in this study). During these preworkshop semesters, the lecture course met for 50 min three times each week. In the spring 2004 semester, I conducted a pilot workshop project in which I taught four sections of a one-credit Learning Strategies for Biology course that ran concurrently with the lecture course. This Learning Strategies for Biology course included teacher-centered and student-centered review activities (Preszler, 2004a). Only 36 of the 248 students enrolled in the lecture course also enrolled in Learning Strategies for Biology. Student response system "clickers" were introduced to the course in fall 2004 semester and were used in every semester from fall 2005 to the present (as described in Preszler *et al*., 2007), and they have been used in every subsequent semester. In the spring 2006 semester, additional weekly discussion sections were led by the course instructors. Students were given a choice of completing short essay assignments or participating in discussion sessions. Again, student participation was low with approximately 30 students/wk (of the 263 students enrolled in the lecture course) attending the discussion sections.

Workshop Course. In the spring 2007 semester, with support from a grant to NMSU from HHMI and from the NMSU College of Arts and Sciences, required peer-facilitated workshops were added to the Natural History of Life lecture course in place of one weekly 50-min lecture. This study reports on student evaluations and performance in the first three semesters (spring 2007, fall 2008, and spring 2008) of the workshop version of the lecture component of the Natural History of Life. In this new version of the lecture course, students are required to attend two 50-min lectures each week and one 65-min workshop; as in the preworkshop semesters, most also attend one 2½-h weekly laboratory course. The workshops are not supplemental to the course nor are they an associated elective course. They are a required component of the lecture course. In the spring 2007 semester, assignments completed in the workshop contributed to 25% of students' grades; in the fall 2007 semester, quizzes on preworkshop readings and workshop assignments contributed to 28% of the course grades; and in the spring 2008 semester, quizzes on preworkshop readings and workshop assignments contributed to 19% of the course grades. The number of points derived from preworkshop readings and workshop assignments was reduced in the spring 2008 semester because the instructor felt that the

Table 1. Semester course characteristics; the characteristics of the preworkshop and the workshop lecture courses are shown for each semester

	Preworkshop							Workshop		
Semester	F03	SP04	F04	SP05	F05	SP06	F06	SP07	F07	SP08
Students in lecture	297	248	355	285	318	263	364	246	310	229
Lecturer	A	B	C/D	E	B	B/F	G	B	G	B
Clickers in lecture	No	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
WebCT quizzes	No	No	Yes	Yes	Yes	No	No	Yes	No	Yes
No. of 50-min lectures/wk	3	3	3	3	3	3	3	2	2	2
Required 65-min workshop/wk	No	No[a]	No	No	No	No[b]	No	Yes	Yes	Yes

[a] In spring 2004, instructor B offered four elective workshops that met each week. In total, 36 students enrolled in these elective supplemental workshops.

[b] In spring 2006, course instructors led four workshops each week. Students could attend the workshops or complete a homework assignment. Approximately 30 students/wk participated in these voluntary workshops.

scores derived from workshops, WebCT quizzes, and clickers were overweighted, and due to the lack of variation among students in these scores, were dampening the relationship between total course scores and the students' understanding of biology. In all three workshop semesters, lecture exams included questions associated with workshop activities.

The workshop peer facilitators, biology learning catalysts (BioCats), were chosen based on their transcripts (students who had earned high grades in the Natural History of Life and subsequent science courses were preferred), written statement of their teaching and learning philosophy, and two interviews. Eight BioCats were hired for the larger fall semester course; six BioCats were hired for each of the spring semester courses. BioCats are hired as juniors or seniors, so they are often rehired for two or three semesters. BioCats were paid $1500/semester for 10 h of work per week to attend preparation meetings and lectures, prepare individually for their workshops, lead their two workshop sessions, and grade workshop reports. The lecture sessions met in a large lecture hall; the workshop sessions were held in a much smaller room with six tables sitting up to four students at each table. Before each semester, BioCats were required to attend 1- or 2-d training sessions where they were introduced to procedures and responsibilities associated with their position, overall course structure, details of the first workshops, and the purpose and philosophy of the workshops. BioCats were required to attend lectures and to attend one meeting each week to prepare for the next workshop. Each BioCat taught the same two sections throughout the semester. The weekly meetings were led by the lecture instructor with the help of a graduate teaching assistant. One teaching assistant, working up to 20 h/wk, was assigned to the course each semester. They helped with administering the course, tutoring students, and mentoring the BioCats. Lecture instructors were responsible for designing the workshop activities along with keys for the workshop activities, although the activities and keys were often refined in response to input from BioCats. The discussion of the grading keys ensured that all the BioCats graded the workshops consistently. The students were not provided with the workshop keys, although they were provided with written comments from their BioCats on each of their reports. During both the presemester training sessions and the weekly meetings it was emphasized to BioCats that they were responsible for keeping cooperative learning groups on task as they worked through case studies, problem sets, and review activities. BioCats were encouraged to use prompts rather than direct answers or mini-lectures in response to student questions. It was repeatedly emphasized that their primary task was to facilitate cooperative-learning activities, rather than lecture or tutor their workshop students.

Workshop activities (see boxed material below) ranged from learning-skill sessions to case studies. Biological concepts were first introduced in the lecture and the textbook. A subset of these concepts was explored in more depth in the workshops; the workshops were not designed to introduce students to new biological concepts. In case-study activities, students analyzed data and applied the results to topical issues, such as bacterial resistance to antibiotics, pandemics, endangered species, human population growth, climate change, and pollution. In other cases, students used data to evaluate hypotheses that were not directly tied to topical issues, such as species interactions and speciation. If there was time left over at the end of a case-study session, it was used to conduct short learning-skill activities. The majority of the sessions, and a majority of the time within each session, focused on working through case studies. In the most recent workshop semester, a majority of workshop sessions (eight of 13) were devoted to investigative cases, two sessions were focused entirely on learning-skill development, and three sessions were a mix of problem sets and review activities. All sessions emphasized cooperative learning within teams of three or four students.

Learning-Skill Activity. An example of a learning-skill activity is the first workshop session in spring 2008, which focused on note-taking skills. Note-taking strategies had been identified as a strong predictor of student success in the course (Preszler, unpublished). The lecture instructor or the graduate student lecture assistant visited each workshop and asked students to take notes as they presented a short, fast-paced lecture. The lecturer then left and the BioCat took over the workshop session. Students processed their notes; compared and discussed the information within their cooperative teams; and to simulate the relationship between lectures and lecture exams, answered a series of multiple-choice questions about the mini-lecture material. The BioCats led workshop discussions of problems students had taking notes and studying from their notes and they helped the class arrive at solutions to these problems.

Case-Study Activity. An example of a case study in which students explore a current research topic is our use of evolutionary and behavioral studies of the side-blotched lizard, *Uta stansburiana* (Sinervo and Lively, 1996; Bleay *et al*., 2007). This case study allows students to discover a sequence of concepts: 1) the utility of a biological application of a mathematical null model (Hardy–Weinberg equilibrium [HWE]); 2) frequency-dependent selection is a form of balancing selection that maintains the frequencies of two alternative alleles in a population; and 3) behavioral observations can illuminate the selection pressures associated with balancing selection. In the first step, students are given estimates of the numbers of orange-, blue-, and yellow-throated *U. stansburiana* (morphs) based on surveys of Arizona and California populations. The genetic model is simplified from a three-allele to a two-allele system. Students convert the counted phenotypes to counts of genotypes and alleles. They use HWE, starting with the observed allele frequencies, to predict genotype frequencies. When they compare the genotype frequencies predicted by HWE to the observed genotype frequencies, they find that the genotype frequencies are more evenly distributed than predicted by HWE. This suggests that an evolutionary process is tending to equalize the frequencies of the three genotypes. In the second step of the workshop, students consider behavioral descriptions of interactions between the three lizard morphs. Students use conclusions they construct from the behavioral observations to predict and explain the results of an experiment that measured the mating success of males of the three lizard morphs in populations that were manipulated to have a high proportion of orange-, yellow-, or blue-throated lizards. Each cooperative group of students then draws a figure on their group's whiteboard that illustrates the interactions between the three *U. stansburiana* morphs and explains how these behavioral interactions result in frequency-dependent balancing selection.

Assessment. The success of our workshop course structure was assessed with student evaluations, longitudinal demographic analyses of course grades, a comparison of student performance on paired exam questions given in preworkshop semesters and workshop semesters, and an analysis of the quality of exam questions in preworkshop and workshop semesters. The protocol for using students as human subjects in our assessment of our workshop course structure was approved by the NMSU Institutional Review Board (NMSU Institutional Review Board 6463).

Student Evaluations. In each of the three semesters of workshops, students used a Likert scale (strongly agree, agree, neither agree nor disagree, disagree, and strongly disagree) to respond to statements about the workshops. These questions were included in anonymous lecture course evaluations that were filled out by students in class during the last week of the semester. Student responses to the following statements were averaged across the three workshop semesters: 1) In general, the workshops made me more interested in the material/content of the course; 2) In general, the workshops helped me better understand the material/content in this course; and 3) Attending the two lectures and one workshop each week is more valuable than attending three lectures and no workshops.

Demographic Analyses of Course Grades. Two-way contingency table analyses were used to evaluate the relationships between course grades and course structure, gender, ethnicity, and academic discipline. For each two-way contingency table analysis, as well as the three-way analyses described below, if the probability associated with the Pearson χ^2 was <0.01, the null hypothesis of no effect of the variable (two-way) or variables (three-way) on student grades was rejected. The two-way contingency table analyses were used to address the following four questions.

1. Course structure. Does the distribution of "A," "B," "C," "D," "F," and "W" (withdrawal from course by midsemester) grades differ between students in the preworkshop version of the course (fall 2003–fall 2006) and students in workshop semesters (spring 2007–spring 2008)?
2. Gender. Do grade distributions differ between males and females?
3. Ethnicity. Do grade distributions differ between students who have self-identified as being a member of an URM (African American, Latino, or Native American) and students who have self-identified as being a member of an ethnic group that is not underrepresented (non-URM: Asian American or Caucasian)? Students who selected "other" or did not select a category in response to the ethnicity question during the university application process were excluded from the ethnicity analyses. Only students who indicated they were U.S. citizens were included in the ethnicity analyses.
4. Academic discipline. Do grade distributions differ among students with majors in the School of Agriculture, majors in biological sciences (biology, microbiology, and biochemistry), majors in the School of Education, other declared majors, and students who have not declared a major?

Three-way contingency table analyses were used to evaluate the effects of our change in course structure on the relationships between gender and grades, ethnicity and grades, and academic discipline and grades. They were used to evaluate the following questions:

1. Was the shift in grade distributions from preworkshop to workshop semesters the same for female students as it was for male students?
2. Was the change in grade distribution from preworkshop to workshop semesters the same for URM students and non-URM students?
3. Did the grade distributions of various academic disciplines (described above) shift in the same way from preworkshop to workshop semesters?

Pre- and Postworkshop Exam Questions

To determine whether student learning, as reflected by performance on lecture exam questions, improved with the onset of workshops, a subset of exam questions in spring 2007 and in spring 2008 semesters were derived from preworkshop exams used in the spring 2006 semester. These paired exam questions were used to determine whether student performance in workshop biology in the spring 2007 semester was higher than in the preworkshop in the spring 2006 semester, and also to determine whether student performance was higher in the spring 2008 semester than in the spring 2006 semester. The significance of the difference in scores was determined using one-tailed paired t tests with an α level of 0.05. Both analyses included questions from the genetics, evolution, and ecology exams. Most of the questions addressed single concepts; 11% of the questions challenged students to draw conclusions from the synthesis of multiple concepts; the questions used in these analyses did not address novel problems or examples. The recycled questions used in 2007 were a different set of questions from those recycled in 2008, so that if any students in 2008 had access to the previous year's exams they would not have an unfair advantage and they would not be encouraged to study at a simple recognition level of understanding.

Quality of Exam Questions

By the end of the third workshop semester the current course instructor (R. Preszler) had the impression that improved student performance had allowed him to ask more higher-level questions (questions that asked students to apply concepts learned in the course to novel problems or examples) on exams. He also had the impression that students were better prepared to answer such questions as a result of improved critical-thinking skills. These post hoc hypotheses were evaluated by counting the proportion of exam questions that asked students to address novel examples or problems in the most recent version of the workshop course (spring 2008) in comparison with when the same instructor taught the course in the preworkshop format (fall 2005 and spring 2006); the hypothesis that workshop students perform better on higher-level questions was tested by comparing student scores on higher-level exam question in the most recent workshop semester (spring 2008) with their performance on this type of question in preworkshop semesters (fall 2005 and spring 2006). The last comparison was evaluated with a one-tailed two-sample t test ($\alpha = 0.05$).

Quantity of Concepts Covered in Lecture

Before the implementation of workshops (fall 2006), the author constructed a hierarchical concept list describing the concepts he felt students should learn in the Natural History of Life lecture course. This list of concepts was approved by the NMSU Department of Biology's Undergraduate Curriculum and Teaching Committee and was presented to and approved by the biology faculty. In this study, to determine whether the loss of one weekly lecture has reduced the quantity of concepts covered in the course, the number of concepts at each level in the hierarchy (primary through quaternary) covered in the preworkshop semesters has been counted and compared with the number at each level covered in the most recent (spring 2008) workshop semester. The complete concept hierarchy indicating which versions of the course covered each concept is included in Supplemental Material.

RESULTS

Student Evaluations

Clear majorities of students either strongly agreed or agreed that workshops improved their understanding of biology (Table 2, first row) and that workshops improved their interest in biology (Table 2, middle row). A very strong majority of students strongly agreed or agreed that attending two lectures and one workshop is more valuable than attending three lectures each week (Table 2, bottom row).

Grades

Main Effect of Course Structure. There was a significant change in the distributions of student grades from the seven preworkshop semesters to the three workshop semesters (Table 3). This change was due to large percent increases in

Table 2. Combined student responses over three semesters of workshops to three questions about the workshop

Question	Student response category				
	Strongly disagree	Disagree	Neither	Agree	Strongly agree
Structure	5.11	8.25	12.38	18.27	55.99
Interest	6.23	12.65	18.68	32.10	30.35
Learn	4.30	8.98	13.48	30.66	42.58

The first row lists the percentage of students choosing each response to the statement "The workshops made me more interested in the material/content of the course" (n = 514 students). The middle row lists the percentage of students choosing each response to the statement "The workshops helped me better understand the material/content of the course" (n = 512 students). The bottom row lists the percentage of students choosing each response to the statement "Attending the two lectures and one workshop each week is more valuable than attending three lectures and no workshops" (n = 509 students).

students earning "A's" and "B's" and large percent decreases in students earning "F's" and "W's." There was a 45% increase in the proportion of students earning an "A" or "B" with the advent of workshops. There was little overall percent change in "D's" and "C's" with the advent of the workshops due to the shift of students from lower grades into "D's" and "C's" and the concordant shift of students out of "C's" and "D's" into higher grades.

Gender. Averaged across all 10 semesters, the differences in grade distributions between females and males are so small as to be statistically inconclusive (Table 4, significance of gender over all semesters). The grade distributions of both females and males improved in response to the introduction of workshops (Table 4, significance of course structure, females only; significance of course structure, males only). However, the responses of females to the change in course structure are clearly different from the responses of males (Table 4, significance of different effects of course structure on males and females). Male students showed a sharper increase in "A's," nearly bringing them up to the same percent "A's" as for females. Females and males had similar percent "B's" before the introduction of workshops, but during the workshop semesters 29.89% of females earned "B's," whereas males only increased up to 22.48%. The increase in the proportion of females earning "A's" or "B's" with the advent of workshops was 49%; the increase in the proportion of males earning "A's" or "B's" was 37%. Although course withdrawals of both females and males went down with the implementation of workshops, this increased retention of students in the workshop semesters was a particularly strong effect on female students (52.8% reduction in "W's") in comparison with males (20.8% reduction in "W's").

Ethnicity. Overall, there was a significant performance gap (Table 5, significance of URM in comparison with non-URM grades across all semesters) between URM (combined African Americans, Latinos, and Native Americans) and students who have self-identified as being a member of an ethnic group that is not underrepresented (non-URM combined Asian Americans and Caucasians). The difference in the percentage of URM in comparison with non-URM students earning "A's" was −7.5%; the difference in the percentage of URM in comparison with non-URM students earning "B's" also was −7.5% (Table 5). Both groups of students benefited from the introduction of workshops (Table 5, significance of course structure URM students only; significance of course structure non-URM students only). This positive shift in the distribution of grades differed between the two groups (Table 5, significance of different effects of course structure on URM and non-URM students). Underrepresented minority students experienced a 47% increase in the proportion of students earning "A's" or "B's," whereas non-URM students showed a 36% increase. However, non-URM students experienced a larger reduction in "F's" and in course withdrawals ("W's") than URM students.

Academic Discipline. There was a difference in the distributions of student grades across academic disciplines (Table 6). Much of the difference was due to the higher grades of students with majors in the biological sciences. In general, students with the lowest grades were those who had not declared a major. These students had the lowest percent "A's" and "B's" and a higher percentage of these undeclared students received "F's" than did students who had declared a major—regardless of the field of their declared major.

There was a significant difference among academic disciplines in the response of students, as indicated by their course grades, to the introduction of workshops (Table 6, significance of different effects of course structure among

Table 3. Course structure

Course structure	Grade						n
	"A"	"B"	"C"	"D"	"F"	"W"	
Preworkshop (% each grade)	9.97	17.96	23.93	13.12	22.94	12.08	2127
Workshop (% each grade)	13.55	26.98	24.42	13.30	14.58	7.16	782
Difference	3.59	9.02	0.49	0.18	−8.37	−4.92	2909
% Change	36.00	50.24	2.06	1.39	−36.46	−40.73	
Significance	Two-way Pearson $\chi^2 = 61.85$, df =5, $p < 0.001$						

Grade distributions of students in preworkshop and workshop semesters are shown followed by the difference in percentages, and the percentage of change. The significance of a two-way contingency table analysis of grades × course structure is reported in the last row.

Table 4. Gender

	Grade						n
	"A"	"B"	"C"	"D"	"F"	"W"	
Gender							
Female (% each grade)	11.95	21.45	23.72	12.06	20.16	11.95	1716
Male (% each grade)	9.47	18.86	24.56	14.75	21.46	9.47	1193
Difference	2.47	2.59	−0.84	−2.69	−1.30	−0.23	2909
Significance of gender over all semesters	Two-way Pearson $\chi^2 = 10.93$, df = 5, $p = 0.053$						
Females							
Preworkshop	11.20	18.21	23.85	11.76	22.48	12.49	1241
Workshop	13.89	29.89	23.37	12.84	14.11	5.89	475
Differences	2.69	11.68	−0.48	1.08	−8.38	−6.60	
% Change	24.05	64.16	−2.03	9.16	−37.26	−52.80	
Significance of course structure, females only	Two-way Pearson $\chi^2 = 50.28$, df= 5, $p < 0.001$						
Males							
Preworkshop	8.24	17.61	24.04	15.01	23.59	11.51	886
Workshop	13.03	22.48	26.06	14.01	15.31	9.12	307
Differences	4.79	4.87	2.02	−1.00	−8.28	−2.39	
% Change	58.14	27.65	8.39	−6.69	−35.10	−20.78	
Significance of course structure, males only	Two-way Pearson $\chi^2 = 17.04$, df = 5, $p = 0.004$						
Course structure by gender							
Significance of different effects of course structure on males and females	Three-way Pearson $\chi^2 = 81.36$, df = 16, $p < 0.001$						

Two-way contingency table analyses of gender × grades across all semesters is shown in the first section followed by analyses of the effects of course structure on females and on males. The last row describes the results of a three-way contingency table analysis of the differences in the effects of the change in course structure on the grades of females in contrast to males.

disciplines). The grade distributions of students with declared majors in the School of Agriculture, the biological sciences within the School of Arts and Sciences, the School of Education, and any other declared major not in the abovementioned categories all improved with the onset of workshops. In contrast, students without a declared major showed no significant change in their grades with the onset of workshops (Table 6).

Pre- and Postexam Question Scores

There was no significant change in student scores on recycled exam questions from the spring 2006 preworkshop semester and spring 2007, the first workshop semester (paired t test = −1.35, df = 65, one-tailed $p = 0.090$). However, there was a significant 6.19% increase in scores on exams questions used in both spring 2006 (preworkshops) and the most recent workshop semester included in this study, spring 2008 (paired t test = −2.44, df = 44, one-tailed $p = 0.009$).

Quality of Exam Questions

There was a threefold increase in the proportion of exam questions that asked students to solve problems or explain examples that had not been presented in lecture, workshop, or the textbook. In the spring 2008 semester, 15.8% of exam questions fell into this category, whereas only 5.1% of the exam questions presented novel problems or examples in spring 2006 and fall 2005 preworkshop semesters. Students in the spring 2008 semester also performed better on these questions, which required a higher level of understanding (65.9% correct), than students in the spring 2006 and fall 2005 preworkshop semesters (56.6% correct). This was a significant improvement in student performance on these higher-level learning questions (t test = −1.67, df = 32, one-tailed $p = 0.052$).

Quantity of Concepts Covered in Lecture

As indicated in Table 7, there was a slight reduction in the number of targeted concepts covered in the workshop version of the course in comparison with the preworkshop course.

DISCUSSION

Our comparison of student performance in pre- and postworkshop semesters indicates that the replacement of one weekly lecture with a weekly peer-led workshop has dramatically improved student learning. This effect was strongest for females and for URM students. Ideally, longitudinal analyses of student grades across semesters would be based on iterations of the course with identical grading schemes. Because we are continually refining our course, our grading schemes varied between preworkshop semesters and also varied between workshop semesters. Also, the addition of workshops added points associated with workshop activities. Males and females experienced the same sequence of changes in grading schemes from one semester to the next. Therefore, these changes do not bias comparisons of grades by gender or the analysis of the effects of the interaction between course structure and gender on student grades. The same reasoning suggests that changes in grading schemes do not bias comparisons of grades of URM and non-URM students and of students by academic discipline. The addition of workshop points to the course with the onset of workshops does bias our comparison of the overall effect of course structure on student grades. However, our analyses of student scores on matched exam questions indicate that there has been a significant improvement in students' learning of basic course content. This analysis did not show significant improvement in comparison with preworkshop exams in the spring 2007 semester, the first semester of the workshop program; it did show significant improvement over preworkshop performance in spring 2008, the third semester of the workshop program. This indicates that refinements of the workshop activities that occurred from the first to third workshop semesters increased the impact of the workshops on student learning, bringing the learning up beyond the levels seen in the preworkshop semesters. It is unlikely that this is due to a reduction in the total course content because a comparison of concept inventories from pre- and postworkshop semesters shows a very modest decline in total concepts covered. One of the most gratifying results of the transition to a workshop approach is the improvement in the quality of learning; students are performing significantly better on exam questions that require higher-level thinking and instructors are now able to ask more of these questions on lecture exams. Improved student performance on exam questions that require higher-level thinking is consistent with the development and exercise of these skills as case studies are explored in the workshop.

A variety of programs offer students adjunct workshops that are either completely voluntary or are offered as elective courses. These adjunct workshop programs have demonstrated positive impacts on student's grades (Lundeberg, 1990; Van Lanen and Lockie, 1997; Gattis, 2000; Wright *et al*., 2002; Grise and Kenney, 2003; Ogden *et al*., 2003). In a 15-yr study of 39,439 students, in 375 courses at the University of Missouri–Kansas City, Arendale (1997) reported that 37.2% of students in the targeted courses attended voluntary SI sessions, and 14.6% more of the SI participants earned a grade of "C" or better in the at-risk courses in comparison with students who chose not to attend SI (derived from data in Arendale 1997). In a meta-analysis of students from 270 institutions, Arendale (1997) documented a similar improvement in performance associated with participation in SI. Students enrolled in SI also had higher retention and grad-

Table 5. Ethnicity

	Grade						
	"A"	"B"	"C"	"D"	"F"	"W"	% of n[a]
URM							
African American	10.81	21.62	24.32	12.16	21.62	9.46	2.8
Latino	7.48	15.86	25.06	15.11	25.43	11.07	49.9
Native American	3.77	19.81	15.09	20.75	34.91	5.66	4.0
Combined URM	7.38	16.41	24.32	15.36	25.91	10.61	56.6
Non-URM							
Asian American	9.68	22.58	29.03	12.90	19.35	6.45	1.2
Caucasian	15.04	23.98	24.78	10.62	15.22	10.35	42.2
Combined non-URM	14.90	23.94	24.89	10.68	15.33	10.25	43.4
Ethnicity × grades							
Differences URM − non-URM	−7.52	−7.53	−0.57	4.68	10.57	0.36	
Significance of URM in comparison to non-URM grades across all semesters	Two-way Pearson $\chi^2 = 99.34$, df = 5, $p < 0.001$						
URM × course structure							
Preworkshop	6.95	14.34	23.92	15.22	27.97	11.61	1137
Workshop	8.68	22.63	25.53	15.79	19.74	7.63	380
Differences	1.74	8.30	1.60	0.57	−8.23	−3.98	1517
% Change	24.99	57.87	6.70	3.77	−29.43	−34.26	
Significance of course structure, URM students only	Two-way Pearson $\chi^2 = 25.16$, df = 5, $p < 0.001$						
Non-URM × course structure							
Preworkshop	13.38	21.78	24.70	10.83	17.27	12.04	822
Workshop	18.58	29.20	25.37	10.32	10.62	5.90	339
Differences	5.20	7.43	0.67	−0.50	−6.66	−6.14	1161
% Change	38.87	34.11	2.72	−4.64	−38.53	−51.01	
Significance of course structure, non-URM students only	Two-way Pearson $\chi^2 = 25.76$, df = 5, $p < 0.001$						
Ethnicity × course structure							
Significance of different effects of course structure on URM and non-URM students	Three-way Pearson $\chi^2 = 161.97$, df = 16, $p < 0.001$						

Descriptive statistics for each ethnic group are followed by two-way contingency table analyses of ethnicity (URM vs. non-URM) × grades across all semesters, effects of course structure on grades of URM students, and then of non-URM students. The last row shows results of a three-way contingency table analysis of the effects of course structure on the grades of URM in contrast to non-URM students.

[a] The far-right column indicates overall percent of students in each ethnic group in the URM and non-URM rows. In the URM × course structure and non-URM × course structure rows, the far-right column indicates the number of students in each category.

uation rates. An alternative explanation for the higher performance of students who have chosen to participate in SI is that these students were more motivated to learn and so would have outperformed other students even if SI were not available. Students who claim to have wanted to attend SI, but were prevented from doing so by scheduling conflicts, have been used as a motivational control by Arendale (1997) and by Ramirez (1997). In both studies, SI student performance exceeded that of students in the motivational control group. The results of Arendale (1997) indicate that approximately one-third of the improvement in performance of SI students is associated with initial motivational differences, rather than the effects of SI. In a discussion of improved success of students who have elected to enroll in a Treisman Model Math Excel program, Duncan and Dick (2000) suggest increased motivation and increased study time does not necessarily translate into increased performance if students do not know how to study effectively. Our results, and those of Tien *et al*. (2002), indicate that improved student performance as a result of participation in peer-led cooperative learning workshops is not limited to students (approximately one-third of a typical course) with enough motivation to participate in voluntary, or enroll in elective, workshops. Required workshops improved overall performance

Table 6. Academic discipline

	Grade						
	"A"	"B"	"C"	"D"	"F"	"W"	% of n[a]
Academic discipline							
Agriculture	9.83	20.22	25.21	14.68	19.25	10.80	25.1
Biological sciences	18.98	23.02	22.32	9.14	19.51	7.03	19.8
Education	8.47	22.48	27.69	14.01	17.26	10.10	10.7
Other declared majors	9.14	19.90	24.14	13.87	19.90	13.05	21.3
Other undeclared	7.84	18.55	23.08	13.73	24.89	11.92	23.1
Significance of academic discipline across all semesters	Two-way Pearson $\chi^2 = 79.59$, df = 20, $p < 0.001$						
Agriculture							
Preworkshop	9.09	16.83	25.34	15.67	21.47	11.61	517
Workshop	11.71	28.78	24.88	12.20	13.66	8.78	205
Differences	2.62	11.95	−0.46	−3.47	−7.81	−2.82	722
% Change	28.78	71.03	−1.82	−22.16	−36.38	−24.34	
Significance of course structure, agriculture students only	Two-way Pearson $\chi^2 = 18.35$, df = 5, $p = 0.003$						
Biological sciences							
Preworkshop	16.98	20.99	23.35	8.02	22.41	8.25	424
Workshop	24.83	28.97	19.31	12.41	11.03	3.45	145
Differences	7.85	7.97	−4.04	4.39	−11.37	−4.81	569
% Change	46.21	37.99	−17.30	54.81	−50.75	−58.23	
Significance of course structure, biological sciences students only	Two-way Pearson $\chi^2 = 20.28$, df = 5, $p = 0.001$						
Education							
Preworkshop	8.65	17.84	23.78	15.14	20.54	14.05	185
Workshop	8.20	29.51	33.61	12.30	12.30	4.10	122
Differences	−0.45	11.67	9.82	−2.84	−8.25	−9.96	307
% Change	−5.23	65.42	41.30	−18.76	−40.14	−70.84	
Significance of course structure, education students only	Two-way Pearson $\chi^2 = 17.57$, df = 5, $p = 0.004$						
Other declared majors							
Preworkshop	7.92	17.19	24.21	13.12	22.85	14.71	442
Workshop	12.28	26.90	23.98	15.79	12.28	8.77	171
Differences	4.36	9.71	−0.23	2.67	−10.57	−5.93	613
% Change	55.09	56.45	−0.96	20.33	−46.26	−40.35	
Significance of course structure, other declared majors only	Two-way Pearson $\chi^2 = 19.29$, df = 5, $p = 0.002$						
Other undeclared							
Significance of course structure, undeclared students only	Two-way Pearson $\chi^2 = 5.60$, df = 5, $p = 0.347$						
Course structure by discipline							
Significance of different effects of course structure among majors	Three-way Pearson $\chi^2 = 231.43$, df = 49, $p < 0.001$						

The first section shows the percentage of students earning each grade by academic discipline and results of a two-way contingency table analysis of academic discipline × grades. The next four sections show the analyses of the effects of course structure on each academic discipline individually. The last row reports the three-way interaction of course structure × grades × academic discipline.

[a] The far-right column shows the overall percent of students in each group in the Academic discipline rows. It indicates the number of students in each category in subsequent rows.

of the entire lecture course. However, we did find that a minimal level of motivation may be necessary to benefit from the workshop experience. Although there were absolute differences in grade distributions between academic majors, all majors benefited from the workshops. However, the grades of students who had not declared a major did not improve from preworkshop to workshop semesters. This suggests that students' motivation and performance may benefit if they are encouraged to develop career goals and academic majors, even at a freshman level, and even though we know that in many cases these goals and majors will change in the coming years.

Although the grades of both males and females benefited from the workshops, females' grades and retention improved significantly more than that of males. Qualitative analyses of student interviews led Seymour and Hewitt (1997) to conclude that retention of female undergraduate science, mathematics, and engineering (SME) majors is enhanced if they are able to develop an individualized dialogue with their instructors that provides them with personal encouragement; they concluded that this was less important in the retention of male SME majors. Peer instructors in a small workshop setting are more likely to individually encourage students than are instructors in a large lecture hall. This may help explain the dramatic increase in retention of female students with the introduction of workshops (12.5% withdrawal rate of female students in preworkshop semesters, 5.9% withdrawal rate during workshop semesters) in comparison with the much more modest improvement in retention of males (12.0% preworkshop withdrawal rate, 9.5% workshop withdrawal rate). This more personalized dialogue between peer facilitators and students also may have contributed to the improved performance of female students.

Although the grades of both URM and non-URM students benefitted from the workshops, the grades of URM students improved more dramatically than those of non-URM students. Fullilove and Treisman (1990) found that African-American students who chose to enroll in a mathematics workshop program supplementing a first-year calculus course had dramatically improved grades and graduation at the university compared with a historical control group and with students in the calculus course who did not enroll in the workshop; these effects were stronger for African-American males than they were for African-American females. The design of this approach was based on observational studies of the study strategies of historically successful Chinese-American students in comparison with African-American students who historically had low success rates in the course. The underlying philosophy of the workshops was to create a more positive social context for learning mathematics that encouraged groups of students to cooperatively study challenging problems. The PLTL program was developed in part in response to the success of Fullilove and Treisman workshops. URM students' grades have improved when PLTL workshops are added to courses (Gafney and Varma-Nelson, 2008). Born *et al*. (2002) compared lecture exam performance of URM students participating in a workshop associated with a biology course with the performance of URM students in a previous version of the course that did not have associated workshops. Interestingly, the workshop URM students sharply improved from exam 2 to exam 3, whereas the performance of the historic control URM students sharply declined between exams 2 and 3. Non-URM student volunteers assigned to the workshops benefitted more rapidly than the workshop URM students. The non-URM workshop students showed a sharp improvement between exams 1 and 2 and maintained this success through exam 3. Although the reasons for this more delayed benefit for URM students are unclear, it may be that the workshop learning environment is more similar to previous educational experiences of non-URM students than it is to URM students (Born *et al*., 2002). Subsequent research has documented positive effects of SI on the performance of African-American students in a large number of courses associated with SI at the University of Missouri–Kansas City (Arendale, 1997), of SI on URM students in Biology I at San Francisco University (Rath *et al*., 2007), and of an elective variant of PLTL in general chemistry at Washington University (Hockings *et al*., 2008). Formal analyses of the interaction between the effects of workshops and ethnicity or gender on student performance have produced mixed results. Our analysis and those of Rath *et al*. (2007) revealed a greater effect on URM students, and our results also showed that females benefited more from the workshops; in contrast, Hockings *et al*. (2008) found that the benefits of PLTL did not differ between males and females, URM and non-URM, low income and non–low-income, or first-year and upper-level students. As Rath *et al*. (2007) discuss, ethnicity-based performance gaps are a result of many factors, only some of which may be impacted by changes in teaching strategies. Our results suggest the workshops not only helped ameliorate problems associated with performance gaps between URM and non-URM students but also were more beneficial to female students. Contrasting conclusions from studies of differential impacts of peer-led workshops could be due to variation in the structure of the workshops, the student populations, or the nature of the course content. We eagerly await development of further understanding of the relative importance of peer instruction in comparison to cooperative learning; and of understanding of the relative importance of case studies in comparison to challenging problems or learning skill development activities. Although the relative importance of each of these components is unknown, it is clear that our introductory biology students benefited from peer-led cooperative learning in a workshop setting.

Table 7. Number of concepts from the hierarchical list of targeted concepts covered, at each of the four levels of the hierarchy in preworkshop and workshop semesters

Hierarchical level of concepts	Version of course: Preworkshop	Version of course: Workshop
Primary	9	8
Secondary	31	27
Tertiary	48	48
Quaternary	33	31

ACKNOWLEDGMENTS

The success of the workshops was a direct result of the dedication of three semesters of BioCats. Dr. Avis James and Bethany Cook helped write workshop activities and implement the program. Drs. Michele Shuster and Amy Marion helped develop and implement the project. Tonia Lane was instrumental in hiring BioCats and coordinating the NMSU–Howard HHMI program. Dr. M. Shuster kindly reviewed an early draft of this manuscript. The NMSU biology workshop program was supported in part by a grant to NMSU from the HHMI through the Undergraduate Science Education Program; it also was supported by the College of Arts and Sciences at NMSU.

REFERENCES

Allen, D., and Tanner, K. (2005). Infusing active learning into the large-enrollment biology class: seven strategies, from the simple to complex. Cell Biol. Educ. *4*, 262–268.

Arendale, D. R. (1997). Supplemental instruction (SI): review of research concerning the effectiveness of SI from the University of Missouri-Kansas City and other institutions from across the United States. In: Proceedings of the 17th and 18th Annual Institutes for Learning Assistance Professionals: 1996 and 1997, ed. S. Mioduski and G. Enright. University of Arizona, Tucson, AZ: University Learning Center, 1–25.

Arendale, D. R. (2004). Pathways of persistence: a review of post-secondary peer cooperative learning programs. In: Best Practices for Access and Retention in Higher Education, ed. I. M. Duranczyk, J. L. Higbee, and D. B. Lundell, Minneapolis, MN: Center for Research on Developmental Education and Urban Literacy, General College, University of Minnesota, 27–40.

Armstrong, N., Chang, S., and Brickman, M. (2007). Cooperative learning in industrial-sized biology classes. CBE Life Sci. Educ. *6*, 163–171.

Bleay, C., Comendant, T., and Sinervo, B. (2007). An experimental test of frequency-dependent selection on male mating strategy in the field. Proc. R. Soc. B *274*, 2019–2025.

Born, W. K., Revelle, W., and Pinto, L. H. (2002). Improving biology performance with workshop groups. JOST *11*, 347–365.

Crossgrove, K., and Curran, K. L. (2008). Using clickers in nonmajors- and majors-level biology courses: student opinion, learning, and long-term retention of course material. CBE Life Sci. Educ. *7*, 146–154.

Duncan, H., and Dick, T. (2000). Collaborative workshops and student academic performance in introductory college mathematics courses: a study of a Treisman Model Math Excel program. School Sci. Math. *100*, 365–373.

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

Fullilove, R. E., and Treisman, E. M. (1990). Mathematics achievement among African American undergraduates at the University of California, Berkeley: an evaluation of the mathematics workshop. J. Negro Educ. *59*, 463–478.

Gafney, L., and Varma-Nelson, P. (2008) Peer-led team learning: evaluation, dissemination, and institutionalization of a college level initiative. In: Innovations in Science Education and Technology, Vol. 16, ed. K. C. Cohen. New York: Springer.

Gattis, K. W. (2000). Long-term knowledge gains due to supplemental instruction in college chemistry courses. J. Res. Dev. Educ. *33*, 118–126.

Grise, D. J., and Kenney, A. M. (2003). Nonmajors' performance in biology. J. Coll. Sci. Teach. *33*, 18–21.

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

Hockings, S. C., DeAngelis, K. J., and Frey, R. F. (2008). Peer-led team learning in general chemistry: implementation and evaluation. J. Chem. Educ. *85*, 990–996.

Jensen, M., Moore, R., and Hatch, J. (2002). Electronic cooperative quizzes. Am. Biol. Teach. *64*, 169–174.

Johnson, D. W., Johnson, R. T., and Smith, K. A. (1998). Cooperative learning returns to college: what evidence is there that it works? Change *30*, 26–35.

Klionsky, D. J. (2001). Constructing knowledge in the lecture hall. J. Coll. Sci. Teach. *31*, 246–251.

Knight, J. K., and Wood, W. G. (2005). Teaching more by lecturing less. Cell Biol. Educ. *4*, 298–310.

Lundeberg, M. A. (1990). Supplemental instruction in chemistry. J. Res. Sci. Teach. *27*, 145–155.

Novak, J. D. (2002). Meaningful learning: the essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. Sci. Educ. *86*, 548–571.

Ogden, P., Thompson, D., Russell, A., and Simons, C. (2003). Supplemental Instruction: short- and long-term impact. J. Dev. Educ. *26*, 2–8.

Preszler, R. W. (2004a). Cooperative concept mapping improves performance in biology. J. Coll. Sci. Teach. *33*, 30–35.

Preszler, R. W. (2004b). Expanding the nature of science in teaching laboratories: from ethology to investigating animal behavior. In: Tested Studies for Laboratory Teaching, Vol. 25, Proceedings of the 25th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), ed. M. A. O'Donnell, Hartford, CT: Association for Biology Laboratory Education, 255–268.

Preszler, R. W. (2006). Student- and teacher-centered learning in biology workshops. Bioscene. J. Coll. Biol. Teach. *32*, 21–25.

Preszler, R. W., Dawe, A., Shuster, C., and Shuster, M. (2007). Clickers: assessing their effects on student attitudes and student learning in biology lectures. CBE Life Sci. Educ. *6*, 29–41.

Ramirez, G. M. (1997). Supplemental instruction. In: Proceedings of the 13th and 14th Annual Institutes for Learning Assistance Professionals: 1992 and 1993, ed. S. Mioduski and G. Enright, University of Arizona, Tucson, AZ: University Learning Center, 78–91.

Rath, K. A., Peterfreund, A. R., Xenos, S. P., Bayliss, F., and Carnal, N. (2007). Supplemental instruction in introductory biology I: enhancing the performance and retention of underrepresented minority students CBE Life Sci. Educ. *6*, 2003–2216.

Seymour, E., and Hewitt, N. M. (1997). Talking about Leaving: Why Undergraduates Leave the Sciences, Boulder, CO: Westview Press.

Sinervo, B., and Lively, C. M. (1996). The rock-paper-scissors game and the evolution of alternative male strategies. Nature *380*, 240–243.

Springer, L., Stanne, M. E., and Donavan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. Rev. Educ. Res. *69*, 21–51.

Tien, L. T., Roth, V., and Kampmeier, J. A. (2002). Implementation of peer-led team learning instructional approach in an undergraduate organic chemistry course. J. Res. Sci. Teach. *39*, 606–632.

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*, 272–281.

Van Lanen, R. J., and Lockie, N. M. (1997). Using supplemental instruction to assist nursing students in chemistry. J. Coll. Sci. Teach. *26*, 419–423.

Walker, J. D., Sehoya, S. H., Baepler, P. M., and Decker, M. D. (2008). A delicate balance: integrating active learning into a large lecture course. CBE Life Sci. Educ. *7*, 361–367.

Wood, W. B. (2004). Clickers: a teaching gimmick that works. Dev. Cell *7*, 796–798.

Wright, G. L., Wright, R. R., and Lamb, C. E. (2002). Developmental mathematics education and supplemental instruction: pondering the potential. J. Dev. Educ. *26*, 30–35.