

Article

Facilitating Learning in Large Lecture Classes: Testing the “Teaching Team” Approach to Peer Learning

Kathrin F. Stanger-Hall,^{*†} Sarah Lang,^{‡§} and Martha Maas[‡]

^{*}School of Biological Sciences, Integrative Biology, and the [‡]UT Learning Center, the University of Texas at Austin, Austin, TX 78712

Submitted December 18, 2009; Revised May 12, 2010; Accepted May 12, 2010
Monitoring Editor: Gary Reiness

We tested the effect of voluntary peer-facilitated study groups on student learning in large introductory biology lecture classes. The peer facilitators (preceptors) were trained as part of a Teaching Team (faculty, graduate assistants, and preceptors) by faculty and Learning Center staff. Each preceptor offered one weekly study group to all students in the class. All individual study groups were similar in that they applied active-learning strategies to the class material, but they differed in the actual topics or questions discussed, which were chosen by the individual study groups. Study group participation was correlated with reduced failing grades and course dropout rates in both semesters, and participants scored better on the final exam and earned higher course grades than nonparticipants. In the spring semester the higher scores were clearly due to a significant study group effect beyond ability (grade point average). In contrast, the fall study groups had a small but nonsignificant effect after accounting for student ability. We discuss the differences between the two semesters and offer suggestions on how to implement teaching teams to optimize learning outcomes, including student feedback on study groups.

INTRODUCTION

Developments during the last two decades have made many college classrooms less supportive environments for student learning. Between 1990 and 2005 the number of full-time college students in the United States increased by 38% from 13.8 million students in 1990 to 17.5 million in fall 2005 (National Center for Education Statistics [NCES], 2008). Due to budget cuts faced by many colleges and universities, this trend contributed to an overall increase in class sizes in introductory college classes at public universities (NCES, 1995). At the same time students were shifting toward a

consumer attitude regarding their education, tending to become more passive in their learning and less engaged in the classroom (Weimer, 2002; Lord, 2008). This combination of effects presents both a teaching challenge for faculty and a learning challenge for many students.

Many different methods to motivate and engage students have been developed and tested for different classes, especially in the sciences (see Ebert-May and Brewer, 1997; Miller *et al.*, 2001; Crouch and Mazur, 2001; Prince and Felder, 2007; Eberlein *et al.*, 2008 for reviews). One such invention is the teaching team model for large lecture classes, a model that is closely aligned with the recommendations of the Boyer Commission on Educating Undergraduates in the Research University (1998) for improving undergraduate education (Larsen *et al.*, 2001). The teaching team model also directly applies the recommendations of the National Science Foundation (1996) to shift the focus in the science classroom from teaching to student learning and to provide students with opportunities for active learning outside the classroom.

During the 2003/2004 academic year, the University of Texas Learning Center piloted the University of Texas Teaching Team Program (TTP). This program, modeled after the University of Arizona Teaching Team Program (Libarkin

DOI: 10.1187/cbe.09-12-0093

Present addresses: [†]University of Georgia, Plant Biology Department, Athens, GA 30602; [§]Center for Teaching and Learning, Indiana University–Purdue University Indianapolis, Indianapolis, IN 46202.

Address correspondence to: Kathrin F. Stanger-Hall (ksh@uga.edu).

© 2010 K. F. Stanger-Hall *et al.* CBE—Life Sciences Education © 2010 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

and Mencke, 2001), aims to bridge the gap between faculty expectations and student readiness in large classes by creating course-specific learning communities or teaching teams. The TTP trains student volunteers, previously or concurrently enrolled in large lecture classes, to be preceptors (study group leaders), and creates an apprenticeship system that involves faculty, graduate teaching assistants (TAs), preceptors, and the students in class.

Similar to the peer-instruction model championed by Mazur (1997), the TTP uses peers to support student learning, but it differs in several aspects. Peer-instruction engages students during class mainly through peer discussions and in-class testing (e.g., Mazur, 1997; Crouch and Mazur, 2001; Peters, 2005) and has now become one of the more common approaches in science education (for additional approaches to student-assisted teaching see Miller *et al.*, 2001; e.g., Eberlein *et al.*, 2008). In many instances peer-instruction allows the traditional lecture-only approach to be supplemented with problem-based learning (PBL) or process-oriented guided inquiry learning (POGIL) inside the classroom (Eberlein *et al.*, 2008). Sometimes out-of class “workshops” or “tutorials” (e.g., peer-led team-learning [PLTL]) are used to support student learning (e.g., Gosser *et al.*, 2001; Born *et al.*, 2002; Sharma *et al.*, 2005; Eberlein *et al.*, 2008; Hockings *et al.*, 2008), but the common approach in peer instruction (in and outside the classroom) is to provide instructor-designed questions or problems to be discussed and solved in peer groups. Research suggests that students who work collaboratively in peer-led groups answering questions on instructor-provided worksheets (e.g., Goodlad, 1998; Born *et al.*, 2002; Peters, 2005; Sharma *et al.*, 2005) learn more than students who don’t have this opportunity. This seems especially true for the weakest students in the class (Born *et al.*, 2002; Peters, 2005) or for students who attended more than six sessions (Sharma *et al.*, 2005).

In the Teaching Team model implemented in this study, preceptors were recruited from current classes, participation was voluntary, and the peer-facilitated groups functioned as study groups, rather than peer-guided discussion sections. More importantly, the actual topics or questions discussed in the preceptor-facilitated study groups were not set by the instructor but determined by the individual study group members. This gave students more control and therefore more responsibility for their own learning. Preceptors merely facilitated this process and ensured that study groups used active-learning strategies (e.g., creating flow charts, comparing/contrasting structures and processes, step-by-step analysis of complex processes), while giving all group members the opportunity to contribute. As a result, the individual study groups were more heterogeneous and more tailored to individual needs than the more traditional groups (with instructor-controlled content) used in other peer-instruction models.

It was the goal of this study to test the effectiveness of the preceptor-facilitated study groups as implemented in the Teaching Team model on student learning in large introductory biology classes. We report the effects of the TTP on student learning, the demographics of the student population reached, and student feedback on study groups.

MATERIALS AND METHODS

We studied the effects of the TTP on student learning over two semesters (spring and fall 2005) in six introductory biology classes (80–120 students each), which were all taught by the same instructor (K.S.H.). The introductory biology course in this study was the second in a sequence of three required courses for science majors (or the second of two for nonmajors) and could not be taken without at least a grade of C in the prerequisite course. The main focus of the course was on plant and animal physiology, and most students ($\approx 67\%$) were working toward a career in the health professions (e.g., medical, veterinary, dental, or pharmacy school).

Every semester, the instructor (K.S.H.) used one of the first class periods in this course to introduce Bloom’s six levels of cognition (remembering, basic understanding, applying, analyzing, evaluating, synthesizing; Anderson *et al.*, 2001) to all students. This was done for several reasons: 1) to make students aware of the learning expectations for the course, 2) to give them tools to generate their own study questions on learning levels beyond remembering, and 3) to facilitate communication between students and the instructor about learning goals and/or learning difficulties in this course.

Students attended two lecture hours per week, and spent one hour in a discussion section (≈ 20 students each), led by a graduate TA. In addition, during spring 2005, 11 preceptors offered seven weekly study group times, and during fall 2005, nine preceptors offered six weekly preceptor-led study group times. To accommodate student schedules, students could opportunistically attend any of the offered study groups, thus group membership and size were not consistent. The students in the three parallel classes taught each semester sometimes attended the lectures of one of the other two classes, and each study group was open to students from all three classes. As a consequence, we treated the three classes each semester as one class.

To motivate students to participate in spring 2005 study groups we followed a suggestion of our fall 2004 students to award extra credit. We offered three points (2.8%) extra credit toward the first exam if students attended a study group before then. We did not offer the extra credit during fall 2005 due to a combination of larger class sizes and fewer preceptors/weekly study group times.

Recruitment of Preceptors and Training

Preceptors were recruited during the first two weeks of the semester. All applications were screened using grade point average (GPA; 3.0 or higher), grade in the prerequisite course (A or B), and ability to meet weekly at a scheduled time, as criteria. Qualified candidates were interviewed to assess their motivation, level of engagement, overall time commitment, and communication skills. Based on these interviews, candidates were accepted as independent preceptors, teamed with another preceptor to form a preceptor team, or rejected. Preceptors in this study were not paid and did not receive special course credit, but at the end of the semester preceptors received a certificate attesting to their leadership training and their services. Usually one or two preceptors returned the following semester.

Table 1. Centralized workshops offered for preceptors (in biology, chemistry, and mathematics) in the fall semester

- Leading your own study group
- Problem solving and active learning
- Group dynamics
- Keys to memory and retention
- Basic counseling skills for peer tutors
- Motivation and self-leadership
- Tutoring skills to foster independent learning
- Learning preferences
- Leadership skills

Each preceptor had to attend at least four workshops and share what was learned during the weekly Teaching Team meetings.

Each semester all members of the Teaching Team (preceptors, graduate TAs, and the instructor) met weekly to discuss course content, common misconceptions regarding the week's class material, teaching methods, active-learning strategies, and other relevant topics. Each preceptor was responsible for keeping current with the course material, attending the weekly Teaching Team meetings, and leading a weekly study group (one hour) on campus (library, class rooms, learning center), open to all students in the class. To prepare for their weekly study groups, the preceptors discussed, practiced, and received feedback on possible active-learning strategies during the Teaching Team meetings.

In addition, the preceptors were trained in leadership skills, mentoring, and study strategies based on current learning theory by University of Texas Learning Center staff. This training was part of the weekly teaching team meetings in spring 2005 but not in fall 2005 (Figure 1). Due to increasing participation in the TTP (e.g., by more biology, but also chemistry and mathematics classes), and limited Learning Center staff availability, this part of the training was centralized as a workshop series for all preceptors outside of the weekly Teaching Team meetings in fall 2005 (see Table 1 for a list of offered topics). During a typical peer-facilitated study group students would question and explain concepts to each other, break complex processes down into individual steps, draw diagrams and overviews, work out problems on

the board, create and answer potential test questions, and/or generate study questions on all Bloom levels. During spring, preceptors did not prepare any materials that they provided to their study groups. In contrast, the fall preceptors tended to prepare handouts for their groups.

Data Collection and Analysis

Student Demographics. We used Registrar data to access GPA and number of semesters at the university. In an end-of-semester survey we collected data on class rank, major, gender, and ethnicity of our students. We examined the associations between college experience (freshman, sophomore, junior/senior), major (biology major or other), gender (male, female), and ethnicity (African, Asian, European, or Hispanic heritage) of students and their study group participation using Pearson's χ^2 tests. Due to considerable demographic differences (Table 2) between spring and fall students, we did not collate the data from the two semesters but analyzed them separately instead.

Assessment of Student Learning. Final course grades (A, B, C, D, or F on a 10-point grading scale), final course percentage (based on in-class exercises, three exams, and one cumulative final exam), and the final exam scores were used to assess the effects of study groups on student learning. The final exam scores included multiple choice (MC) and short essay (E) scores. In addition, we categorized the short essay questions using Bloom's taxonomy (Bloom, 1956; Anderson *et al.*, 2001) as Level One (L1: remembering facts) or Level Two (L2: basic understanding: remembering explanations) questions, and as Level Three and higher questions (L3+: application, analysis, evaluation, and synthesis). Student answers for each category were analyzed and scored as a percentage of the total possible score in that category.

Study Group Participation. Study group attendance was monitored with sign-in sheets. Study group participation was measured using two variables: SG attendance ("Never attended" or "Attended") and SG attendance level ("Never attended," "Attended Once," and "Attended two or more times"). The decision to use SG attendance or SG attendance level in a particular analysis was determined by the nature of the response variable and sample size. For the purposes of

Table 2. Student demographics (consenting students) during spring and fall semesters

| Demographics | Spring semester | Fall semester | χ^2 | <i>p</i> value |
|------------------------------|-----------------|-----------------|----------|----------------|
| Freshmen ^a | 71.8% (n = 79) | 19.6% (n = 54) | 92.524 | <0.001* |
| Sophomore ^a | 15.45% (n = 17) | 54.2% (n = 149) | 21.205 | <0.001* |
| Junior & Senior ^a | 12.7% (n = 14) | 26.2% (n = 72) | 7.442 | 0.006* |
| % Biology majors | 40.9% (n = 45) | 22.9% (n = 63) | 11.737 | <0.001* |
| % Female | 56.4% (n = 62) | 61.5% (n = 169) | 0.650 | 0.420 |
| % Minority ^b | 38.7% (n = 41) | 50.2% (n = 127) | 2.186 | 0.139 |

^a Based on time on campus. Spring: a freshman was enrolled up to two previous semesters (incl. summer), a sophomore 3–4 semesters, and juniors/seniors five or more. Fall: a freshman was enrolled up to one previous semester (incl. summer), a sophomore 2–3 semesters, and juniors/seniors four or more.

^b Minority: African, Asian, and Hispanic heritage.

* Significant at $p < 0.05$. There were significantly more freshmen, significantly fewer sophomores, juniors and seniors, and significantly more biology majors in the spring class.

clarity, students who attended study groups will be called “participants” and students who did not attend study groups will be called “nonparticipants.”

Influences on Student Learning other than Study Groups. Other possible influences on student learning include student motivation to attend class, student engagement during class, and student ability. We monitored the possible influences of these variables on student performance.

Lecture Attendance. Lecture attendance was assessed with sign-in sheets on random class days throughout the semester and measured as the number of lectures attended. For comparison between semesters, attendance was compared as percent lectures attended of the total surveyed (spring 2005: 12 lectures (of 30); fall 2005: 21 lectures (of 30); exam days were not included). The association between lecture attendance and final course percentage (course%) was determined using a Spearman correlation. All consenting students were included in this analysis. To separate student motivation and study group effects we used lecture attendance as a measure of motivation: We did a separate analysis for students with a similar motivation level (i.e., students who attended between 33 and 66% of all lectures) to examine the relationship between study group attendance and final course percentage for those students.

Student Engagement during Class. Performance on in-class exercises (in the form of minute papers at the end of class; Angelo and Cross, 1993) served as a measure of student engagement during lecture. It was a low-stakes measure that tested student understanding of the day’s lecture content at the end of class (correct explanation in full = A; partial but correct explanation = B; attempted, but incorrect explanation = C; no attempt = F). The purpose of these exercises was to engage students in class to process the lecture information for understanding rather than passively consuming the day’s lecture material. These in-class exercises (lowest score dropped) counted 5% toward the final course grade.

Student Ability. We used start-of-semester GPA data to assess whether all students (across the GPA range) were equally likely to participate in study groups, and, if not, to take student GPA as a measure of student ability into account when evaluating study group effects. After obtaining the GPA data we found that many first-semester freshmen and transfers didn’t have a prior GPA, and when they did, it was usually a 3.0 or 4.0 calculated from AP credit. For example, in the fall semester, at least 30% of the students were missing a GPA or had a GPA possibly influenced by AP credit before enrollment. For statistical analysis involving the effects of student ability we excluded all students with missing GPA and GPA <1.0, but retained students with possible AP-influenced GPA.

Student Perceptions. At the end of the semester, we asked students whether they felt that they had a good grasp of the class material and whether they felt that they were a more active learner now than at the beginning of class. The relationships between SG attendance and a student agreeing that (s)he had a “good grasp of the course material” (disagree/neutral versus agree) and between SG attendance and

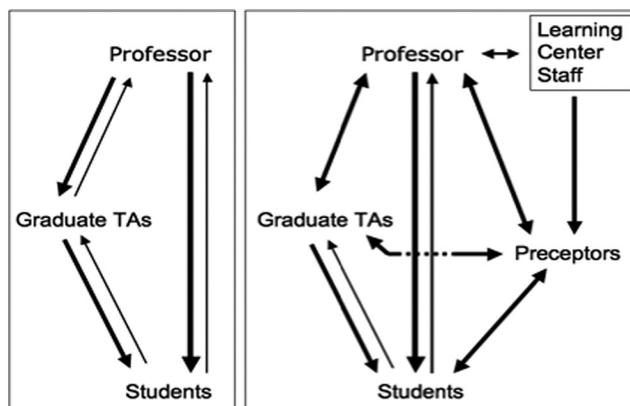


Figure 1. Interactions in a lecture class taught without (left) and with (right) the Teaching Team approach. Arrows indicate flow of information, and thickness of arrows represents intensity of information transfer and/or interaction. The interaction between Center for Teaching and Learning staff and faculty is only ensured if preceptor training is part of the weekly teaching team meeting; this has the added advantage that graduate TAs also benefit.

a student agreeing that (s)he was “a more active learner at the end of class than at the beginning of class” (disagree versus neutral versus agree) were analyzed using Pearson’s χ^2 tests. In addition, a brief (spring) and a more detailed (fall) survey were used to evaluate the TTP. We were especially interested in learning what would motivate the non-participating students to participate, what the main benefits were for the participating students, and what suggestions participants had for improving study groups. All consenting students were included in this analysis.

Between-Semester Comparisons. We used the final exams to compare student performance between semesters. The final exams included a total of 57 (spring) or 59 (fall) MC and true-false (TF) questions, with 50 of these questions being identical (the exams were not returned to students). None of the final exam questions had been asked in the three lecture exams before the final and were not part of the sample exams available to the students. The shared MC and TF questions (which were scored as part of the MC score) included 90% Level 1 or 2 questions (40% L1, 50% L2), as well as 10% Level 3 and higher (L3+) questions. The MC or TF that differed between the two semesters had equivalent difficulty (Bloom) levels. The 12 short answer/essay questions on the final exams included 58.3% Level 1 or 2 questions (8.3% L1, 50% L2), and 41.7% Level 3+ questions. The questions that differed between the semesters were equivalent in design and Bloom level but varied in the example used (e.g., blood sugar versus blood calcium regulation). For the final exam point distribution to different Bloom levels (by question type), please see Figure 2. Please note that the Level 1 or 2 final exam questions included applications and analysis questions that should have been familiar to students from class and thus were classified as remembering or basic understanding for the final exam. The original data from all consenting students were included in this analysis, and nonparametric tests were used.

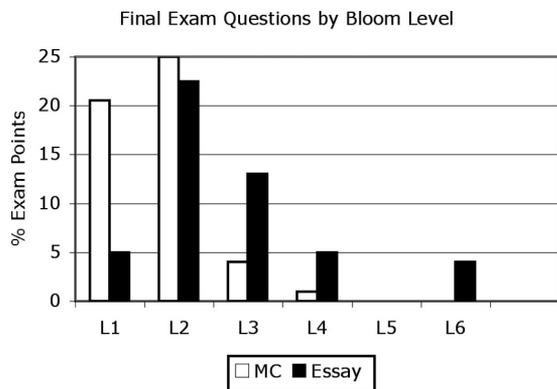


Figure 2. Final exam questions by Bloom level and question type. None of the final exam questions had been asked in the three lecture exams before the final and were not part of the sample exams available to the students. The final exam total was 200 points, with MC and essays contributing equally (50% each) to the point total. Please note that the Level 1 or 2 exam questions included applications (L3) and analysis (L4) questions that should have been familiar to students from class and thus were classified as remembering or basic understanding for the final exam. MC questions tested mostly lower-level thinking skills, essay questions also tested higher-level thinking. Students with factual knowledge and a basic understanding of the class material could earn 75% of the points (i.e., a grade of C) on the final exam. To earn a B or an A, students also needed to demonstrate higher-level thinking skills on the exam (25% of grade).

Statistical Analysis

We used SPSS 17.0 software (2008, SPSS, Chicago, IL), for quantitative statistical analysis. For the original data we used nonparametric Mann–Whitney *U* (two samples) or Kruskal–Wallis (more than two samples) tests to test for differences between means of ranked variables. A Kolmogorov–Smirnov test was used to test for differences in means and distribution. After excluding cases with missing data and outliers in the dependent variables (class%, Final%, MC%, E%, L1L2%, and L3+%) using JMP 8 software (SAS Institute, Cary, NC), we tested all variables for normality (Goodness of Fit: Shapiro Wilkes Test; JMP 8.0), and for homogeneity of variances (Levene’s Test: SPSS 17.0). If necessary, we used SQRT (square root) transformation to obtain normally distributed variables, or a weighted least squares (WLS) approach to adjust the variances before conducting parametric analyses (SPSS 17.0).

Class Data for Final Exam and Class Scores. The relationships between study group attendance and final course grade (A, B, C, D, F) and between attendance and passing the course (pass versus fail/drop) were analyzed using Pearson’s χ^2 test. Kruskal–Wallis tests were used to analyze the associations between level of study group attendance and final course percentage (FC%), as well as study group attendance level and lecture attendance level (LA). To compare SG effects between semesters, we used the Brown–Forsythe Robust Test for Equality of Means, which accounts for potential unequal variances between groups (SPSS 17.0).

Accounting for Student Ability. We tested whether study group participants had different (start-of semester) GPAs than nonparticipants using a Mann–Whitney *U* test (SPSS

17.0). If there was a significant difference, we tested whether study group participation had a significant effect on student performance beyond pre-existing student ability by including start-of-semester GPA as a covariate in our analyses. After testing whether our data met the assumptions of an analysis of covariance (ANCOVA; see Sokal and Rohlf, 1994), we conducted an ANCOVA for our final class score and all the individual scores for the final exam (Total%, MC%, Essay%, L1L2% and L3+%). For all final exam and class scores with significant study group effects, we tested whether the number of attended study groups (SG level: none, one, or two or more study groups) mattered, using the same approach as for study group participation (see above), and the Bonferroni adjustment for multiple comparisons (SPSS 17.0).

RESULTS

Spring 2005

Study groups (N = 7) were offered in 11 weeks, and 55% (N = 60) of the students who consented to this study (N = 110) attended at least one preceptor-led study group. Spring students attended between one and four study groups (mean = 1.4 \pm 0.85; mode = 1). Freshmen were significantly more likely to attend preceptor-led sessions than were upperclassmen ($\chi^2 = 4.366$, df = 1, $p = 0.037$). Female students were significantly more likely to attend than male students ($\chi^2 = 5.697$, df = 1, $p = 0.017$), but there was no significant association between major (biology versus other: $\chi^2 = 0.914$, df = 1, $p = 0.339$) or ethnicity ($\chi^2 = 6.087$, df = 3, $p = 0.107$) and study group participation.

Lecture Attendance, Engagement, and Motivation. Students who attended more classes earned higher final course percentages (calculated without in-class minute paper grades: $R^2 = 0.423$, $p < 0.001$). Similarly, students who attended more classes tended to attend more study groups; however, this trend was not significant. Students who attended between 1/3 and 2/3 of lectures (N = 33) were more or less equally split between participating (N = 16 or 48%) and not participating (N = 17 or 52%) in study groups. We analyzed this 1/3–2/3-attendance group in an effort to isolate the effects of study group participation from lecture attendance (as a measure of student motivation) and found that SG participants from this group (with similar motivation) had significantly higher final course percentages (82.52) than nonparticipants (74.88; Kolmogorov–Smirnov Z = 1.636, df = 1, $p = 0.009$).

Study Groups, Course Performance, and Retention. Participants were less likely to drop or fail the course (Figure 3) and were more likely to earn higher grades (Figure 4). In addition, participants performed better on the final exam. To determine whether this was a study group effect or due to better students attending study groups, we removed all students with missing GPA (or a GPA <1.0) from our data set. The reduced spring data set contained 90 students: 53 study group participants and 37 nonparticipants. The SG participants (N = 53) ranked between ranks 1 and 88 (of 90 students) on GPA scores (nonparticipants: ranks 1–90; Figure 5), but the mean GPA score of SG participants (3.24) was significantly higher than the mean GPA score of nonparticipants (2.98,

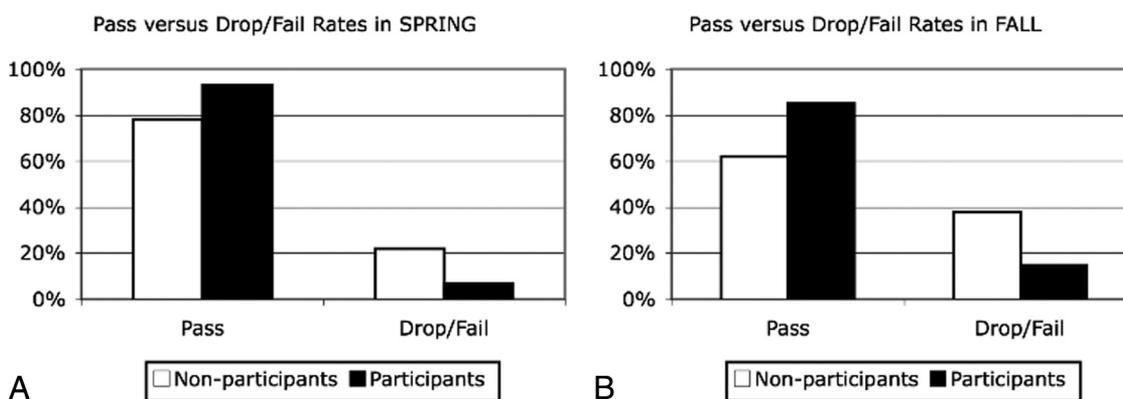


Figure 3. Study group participation and class retention. Pass (A, B, C) versus Drop/Fail (Drop, D, F) rates of participants in study groups versus nonparticipants for (A) spring ($N = 103$), and (B) fall ($N = 240$). Spring: 6.67% of participants dropped or failed the course (60 students attended: 4 dropped/failed) compared with 22% of nonparticipants (50 students never attended a study group: 11 dropped/failed). Fall: 14.7% of participants dropped or failed the course (95 students attended: 14 dropped/failed) compared with 37.8% of nonparticipants (180 students never attended a study group: 68 dropped or failed).

MWU = 717.0, $p = 0.031$), indicating that study group participants were better students overall (Table 3). As a consequence GPA was included as a covariate in subsequent analyses.

Study Groups and Final Exam Question Scores. GPA had a significant effect on all scores, except for lower level essays (Table 4). After controlling for GPA (covariate), study group participation still had a significant effect on all scores. The estimated marginal means (after controlling for GPA influence) are shown in Table 5. To assess the influence of GPA on study group attendance level, we analyzed all students with GPA data and compared the performance of those students that attended no study group ($N = 37$), one study group ($N = 36$), or two or more study groups ($N = 17$). For all performance measures, students who attended only one study group significantly outperformed nonparticipants (Table 6). Even though the estimated marginal means of participants increased (except for MC%) as they attended more study groups, this difference was not statistically significant (Table 6).

Study Groups and Student Perceptions. At the end of the spring semester significantly more study group participants (93.1%) felt they had a “good grasp of the course material,” compared with nonparticipants (76.1%; $\chi^2 = 7.413$, $df = 1$, $p = 0.025$). Students who attended at least one preceptor-led SG tended to agree more with the statement that they were more active learners at the end of class (81%) than nonparticipants (63%), but this difference was not significant; $\chi^2 = 4.269$, $df = 1$, $p = 0.118$).

Fall 2005

During the fall semester study groups ($N = 6$) were offered in nine weeks, and 34.5% ($N = 95$) of students who consented to participate in this study ($N = 275$) attended at least one preceptor-led study group. Fall students attended between one and six study groups (mean = 1.82 ± 1.12 ; mode = 1). There were several demographic differences between students enrolled in the fall and spring semesters, although their career goals were similar: There were signif-

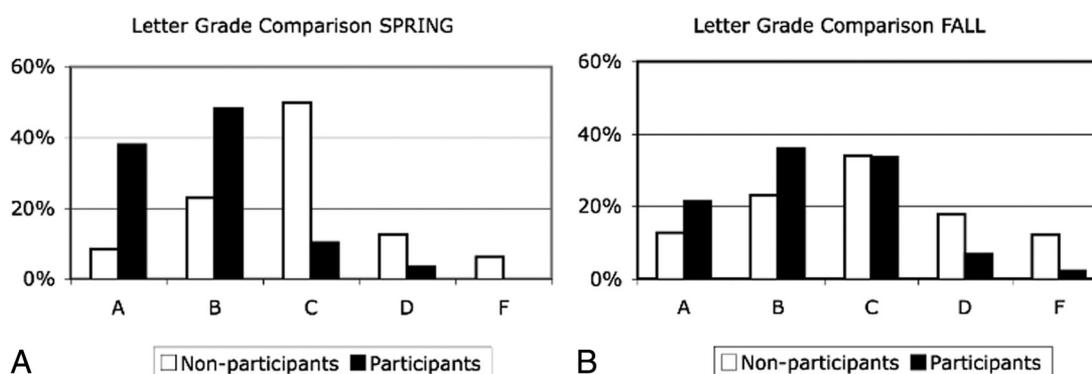


Figure 4. Study group participation and class grade. Class grade (A, B, C, D, F) distributions of participants in study groups versus nonparticipants for (A) spring ($N = 103$) and (B) fall ($N = 240$). Spring participants were more than 4 times more likely to earn an A, two times more likely to earn a B, four times less likely to earn a C, and five times less likely to fail the class (D/F; $\chi^2 = 35.04$, $df = 4$, $p < 0.001$). Fall participants were almost two times more likely to earn an A or B than nonparticipants and over three times less likely to fail the course (D/F; $\chi^2 = 17.627$, $df = 4$, $p = 0.001$).

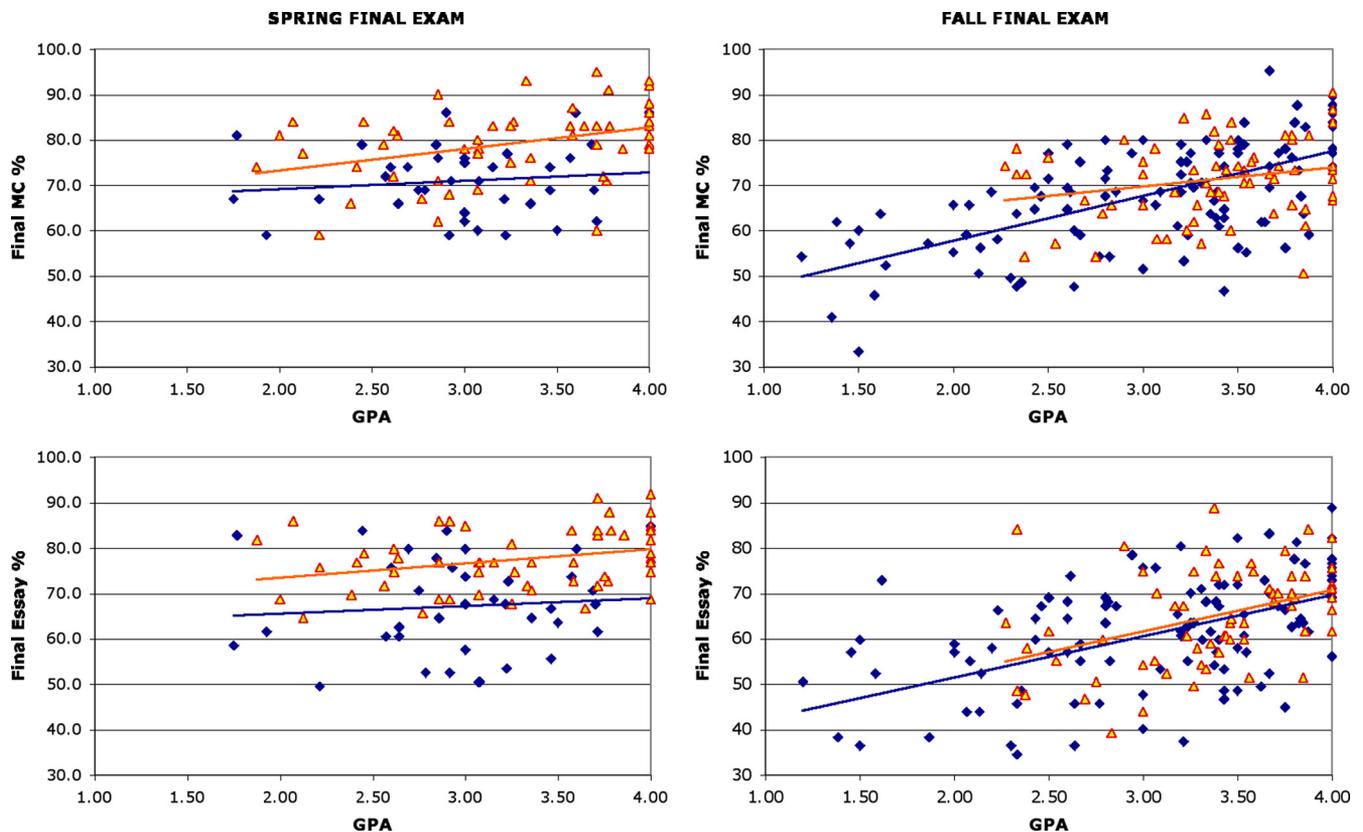


Figure 5. Final exam scores across the GPA range. Final exam scores (top row, MC, and bottom row, Essay) by GPA for study group participants \blacktriangle and nonparticipants \blacklozenge . (Left panels) Spring N: 53 participants and 37 nonparticipants, (right panels) fall N: 65 participants and 106 nonparticipants. Both study group participants and nonparticipants tended to score better on MC and essay exam questions (y axis) with increasing GPA (x axis). However, spring participants (left panel: top line) scored significantly better on the final exam across the GPA range than did nonparticipants (line below). In contrast, there was no difference between the scores of fall participants and nonparticipants (right panel: two lines).

ificantly fewer biology majors taking the class in fall (22.9%: $N = 63$) compared with spring (40.9%: $N = 45$; $\chi^2 = 11.737$, $df = 1$, $p < 0.001$, Table 2), and biology majors made up only 13.84% ($N = 9$ of 65) of study group participants in the fall, compared with 45.28% ($N = 24$ of 53) in spring ($\chi^2 = 12.804$, $df = 1$, $p < 0.001$).

Furthermore, the students in the fall semester had spent an average of 2.61 (median 2.00) semesters at the university, which was significantly more than students in the spring with an average of 2.07 (median 1.00) semesters ($U = 11234.0$, $p = 0.001$). Specifically, there were significantly fewer freshmen and significantly more sophomores and juniors taking the course in the fall than in the spring ($\chi^2 = 113.481$, $df = 3$, $p < 0.001$; Table 2). There were no demographic differences with respect to SG participation: there was no significant difference between freshmen and upper-class students ($\chi^2 = 0.184$, $df = 1$, $p = 0.668$), male or female students ($\chi^2 = 2.143$, $df = 1$, $p = 0.143$), major (biology versus nonbiology/undeclared: $\chi^2 = 0.053$, $df = 1$, $p = 0.818$), or ethnicity ($\chi^2 = 1.426$, $df = 3$, $p = 0.699$).

Lecture Attendance. As in the spring, students who attended more classes in the fall earned higher final course percentages (calculated without in-class minute paper grades: $R =$

0.574 , $R^2 = 0.330$, $p < 0.001$), and students who attended more lectures were more likely to attend preceptor-led study groups, although this trend was not significant.

Study Groups, Course Performance, and Retention. Participants ($N = 95$) earned significantly higher course percentages (75.52 versus 71.29; $R = 0.249$, $R^2 = 0.062$, $p = 0.002$) and scored significantly better on both the final exam MC questions (MWU = 5473.5, $p = 0.016$) and on the essay questions (MWU = 5534.5, $p = 0.023$) than nonparticipants ($N = 180$). The better essay performance was due to better lower-level (L1L2) essay scores (MWU = 5553, $p = 0.025$) but not to better higher-level (L3+) essay scores (MWU = 5798.5, $p = 0.076$). Participants were less likely to drop or fail the course (Figure 3) and were more likely to earn higher grades (Figure 4). After removing students with missing GPA (and GPA < 1.0), the fall data set was reduced to 171 students: 65 study group participants and 106 nonparticipants. Similar to the spring semester, the SG participants in fall ranged over a wide range of abilities: SG participants ($N = 65$) ranked between 1 and 153 (of 171 students) on GPA scores (nonparticipants ranked 1–171; Figure 5), but the mean GPA score of SG participants (3.36) was significantly

Table 3. Student scores of study group participants (SG) and nonparticipants (no SG) in spring and fall (for students with start-of-semester GPAs)

| Dependent variable | Spring semester | | | | Fall semester | | | |
|--------------------|-----------------|-------|-------|----------------|---------------|-------|--------|----------------|
| | No SG | SG | MWU | <i>p</i> value | No SG | SG | MWU | <i>p</i> value |
| N | 37 | 53 | | | 106 | 65 | | |
| GPA | | | | | | | | |
| Mean | 2.98 | 3.24 | 717.0 | 0.031* | 3.01 | 3.36 | 2529.5 | 0.004* |
| SD | 0.53 | 0.63 | | | 0.73 | 0.49 | | |
| Median | 3.00 | 3.27 | | | 3.20 | 3.42 | | |
| Course% | | | | | | | | |
| Mean | 76.98 | 86.08 | 294.0 | <0.001* | 76.28 | 81.32 | 2440.0 | 0.001* |
| SD | 6.87 | 5.79 | | | 10.14 | 7.85 | | |
| Median | 76.57 | 85.28 | | | 77.90 | 83.30 | | |
| Final exam% | | | | | | | | |
| Mean | 68.72 | 77.82 | 400.0 | <0.001* | 68.09 | 72.17 | 2788.5 | 0.004* |
| SD | 8.57 | 6.22 | | | 12.24 | 9.14 | | |
| Median | 67.50 | 78.50 | | | 69.25 | 73.00 | | |
| MC% | | | | | | | | |
| Mean | 70.97 | 79.11 | 453.0 | <0.001* | 67.80 | 71.31 | 2846.5 | 0.057 |
| SD | 7.74 | 8.21 | | | 11.76 | 8.67 | | |
| Median | 71.00 | 81.00 | | | 68.57 | 72.28 | | |
| Essay% | | | | | | | | |
| Mean | 67.16 | 77.30 | 434.5 | <0.001* | 60.73 | 64.93 | 2831.0 | 0.051 |
| SD | 10.51 | 6.78 | | | 12.70 | 10.81 | | |
| Median | 67.68 | 76.77 | | | 61.68 | 66.35 | | |
| Level 1 + 2% | | | | | | | | |
| Mean | 77.12 | 85.34 | 510.5 | <0.001* | 67.67 | 72.11 | 2791.5 | 0.037* |
| SD | 9.87 | 6.96 | | | 12.96 | 10.65 | | |
| Median | 78.57 | 83.93 | | | 69.23 | 72.30 | | |
| Level 3+% | | | | | | | | |
| Mean | 51.77 | 63.86 | 512.5 | <0.001* | 50.00 | 53.80 | 3053.0 | 0.212 |
| SD | 16.47 | 10.74 | | | 16.08 | 15.65 | | |
| Median | 51.11 | 64.44 | | | 50.00 | 52.38 | | |

MWU, Mann–Whitney *U* statistics for comparison of means; MC, multiple-choice questions. Level 1 + 2: lower-level essay questions (remembering facts and explanations), Level 3+: higher-level essay questions (critical thinking).

* Significant at $p < 0.05$.

Spring: participants had significantly higher GPAs and scored significantly better than nonparticipants in all scores measured; Fall: participants had significantly higher mean GPAs and scored significantly better than nonparticipants in only two scores: overall course score and the total final exam score. The latter was mainly due to a significantly higher lower-level (remembering facts and explanations) essay question score.

higher than the mean GPA score of nonparticipants (3.01; MWU = 2529.5, $p = 0.004$) indicating that study group participants were significantly better students overall (Table 3). As a consequence GPA was again included as a covariant in all subsequent analyses.

Study Groups and Final Exam Question Scores. GPA (covariate) had a significant effect on overall class performance and final exam scores (Table 4). In contrast, study group participation had a small but statistically not significant effect (beyond GPA; Table 4, Figure 6). As a result, the estimated marginal means (after controlling for GPA) were very similar between participants and nonparticipants (Table 5).

Study Groups and Student Perceptions. At the end of the fall semester 82.95% of participants felt they had a “good grasp of the course material,” compared with 75.17% of nonparticipants. This difference was not significant ($\chi^2 = 4.213$, $df = 1$, $p = 0.122$). Participants were also not significantly

more likely to agree with the statement that they were more active learners at the end of class (73.86%) than nonparticipants (67.79%; $\chi^2 = 1.223$, $df = 1$, $p = 0.543$).

Student Feedback

After the fall semester we asked students for feedback on the preceptor-led study groups regarding what was helpful, what could be improved, and, if not participating, why they did not attend. This feedback closely echoed the feedback we received during spring when we asked students after their first exam whether they attended a study group before the exam and what they thought. Below we report the more extensive data for fall.

What Participating Students Found Helpful about the Preceptor-Led Study Groups

Many of the students (43%) who participated in study groups felt that working together with their peers on the

Table 4. Study Group (SG) effects after accounting for GPA

| Dep. variable | ANCOVA | | Spring semester | | Fall semester | |
|---------------|------------------|----|-----------------|---------|---------------|---------|
| | Factor | df | F | p value | F | p value |
| Course% | GPA ^a | 1 | 26.685 | <0.001* | 85.945 | <0.001* |
| | SG | 1 | 35.406 | <0.001* | 3.115 | 0.079 |
| Final exam% | GPA ^a | 1 | 9.438 | 0.003* | 69.693 | <0.001* |
| | SG | 1 | 24.782 | <0.001* | 0.317 | 0.574 |
| MC% | GPA ^a | 1 | 7.425 | 0.008* | 61.227 | <0.001* |
| | SG | 1 | 17.797 | <0.001* | 0.137 | 0.712 |
| Essay% | GPA ^a | 1 | 4.762 | 0.032* | 54.310 | <0.001* |
| | SG | 1 | 22.525 | <0.001* | 0.513 | 0.562 |
| Level 1 + 2% | GPA ^a | 1 | 0.079 | 0.374 | 54.310 | <0.001* |
| | SG | 1 | 16.877 | <0.001* | 0.513 | 0.475 |
| Level 3+% | GPA ^a | 1 | 5.190 | 0.025* | 27.764 | <0.001* |
| | SG | 1 | 12.246 | 0.001* | 0.063 | 0.801 |

^a Covariate influence on student performance. MC, multiple-choice questions. Level 1 + 2: lower-level essay questions, Level 3+: higher-level essay questions. Spring (n = 90): study group participation had a significant effect on all scores (beyond GPA). Fall (n = 171): study group participation had no significant effect (beyond GPA) on student performance.

* Significant at $p < 0.05$.

course material led to a deeper understanding (Figure 7: A1). They reported that explaining what they knew to others was a great learning experience and helped solidify their understanding. They found that working in small groups was highly motivating and a great way to meet other students to study with. In addition, 26% reported that study groups helped them to clarify information and get an explanation in simpler terms (e.g., through the use of tables and diagrams to summarize difficult concepts; Figure 7: A2). Seventeen percent of students mentioned that handouts or worksheets prepared by the preceptors were very helpful in allowing them to work with the course material in a different context than during class (Figure 7: A3). Participation in study groups also encouraged students to keep up with the reading and come prepared to class and discussions (8%; Figure 7: A4). Finally, some students felt that asking questions during study group was less intimidating and allowed for more one-on-one feedback than in class (6%; Figure 7: A5).

What Participating Students Would Improve about the Study Groups

The most common suggestion on how to improve study groups was more organization (34%; Figure 7: B1). Students felt that an agenda should be set at the beginning of each group, and that groups should focus on questions that couldn't be answered by reading the book. In addition, students requested additional worksheets or test questions for each session (22%; Figure 7: B2). Students also wished that study group membership was more stable, i.e., that more students would attend more consistently, and that all students would interact more within the study group (18%; Figure 7: B3). Others felt that preceptors should be more assertive and know the answers to all questions (14%; Figure 7: B4). Some students explicitly mentioned that there was nothing to improve and that they were very happy with their study group (6%; Figure 7: B5). Finally, several students felt that they needed more time than the 60-min study sessions, some suggesting that the same study group meet twice weekly (3%;

Table 5. Study group effects: estimated marginal means of study group (SG) participants and nonparticipants (no SG) on course and final exam scores after controlling for GPA (covariate)

| Variable | Spring semester | | | | Fall semester | | | |
|--------------|-----------------|-------|--------|---------|---------------|-------|-------|-------|
| | No SG | SG | F | p | No SG | SG | F | p |
| Course% | 77.91 | 85.76 | 35.406 | <0.001* | 77.58 | 79.69 | 3.115 | 0.079 |
| Final exam% | 69.41 | 77.60 | 24.782 | <0.001* | 69.55 | 70.38 | 0.317 | 0.574 |
| MC% | 71.55 | 78.71 | 17.797 | <0.001* | 69.07 | 69.60 | 0.137 | 0.712 |
| Essay% | 67.74 | 77.14 | 22.525 | <0.001* | 61.96 | 62.95 | 0.337 | 0.562 |
| Level 1 + 2% | 77.36 | 85.26 | 16.877 | <0.001* | 68.89 | 70.13 | 0.513 | 0.475 |
| Level 3+% | 52.73 | 63.59 | 12.246 | 0.001* | 51.22 | 51.82 | 0.063 | 0.801 |

MC, multiple-choice questions. Level 1 + 2: lower-level essay questions, Level 3+: higher-level essay questions. Spring: study group participants scored significantly higher means in class and final exam scores than did nonparticipants (after removing GPA influence); Fall: marginal means for participants tended to be higher, but the effect was not significant (beyond GPA) on class and final exam scores.

* Significant at $p < 0.05$.

Table 6. Study group (SG) participation level and student performance in spring

| Spring variables | Mean scores by SG attendance level | | | Effects (<i>p</i> values) ^b | | |
|------------------|------------------------------------|----------------|-----------------|---|----------|----------|
| | No SG (n = 37) | SG: 1 (n = 36) | SG: 2+ (n = 17) | 0 vs. 1 | 0 vs. 2+ | 1 vs. 2+ |
| Course% | 77.581 | 84.902 | 87.273 | <0.001* | <0.001* | 0.572 |
| Final exam% | 69.384 | 77.176 | 78.712 | <0.001* | <0.001* | 1.000 |
| MC% ^a | 71.335 | 78.641 | 78.146 | <0.001* | <0.001* | 1.000 |
| Essay% | 67.430 | 75.869 | 79.858 | <0.001* | <0.001* | 0.230 |
| Level 1 + 2% | 77.225 | 84.400 | 87.322 | 0.002* | <0.001* | 0.478 |
| Level 3+% | 52.442 | 62.062 | 67.076 | 0.010* | 0.001* | 0.406 |

Estimated marginal means of study group participants (1: attended one; 2+: attended two or more) and nonparticipants (0) on course and final exam scores after controlling for GPA (covariate).

^a MC data were SQRT transformed before ANCOVA, and estimated marginal means were back-transformed.

^b Bonferroni adjustment for multiple comparisons.

* Significant at $p < 0.05$. All scores showed significant study group effects (after removing GPA influence), and all study group effects were significant after the first study group.

Figure 7: B6). A few students asked for more study group times (different time slots: 3%; Figure 7: B7).

What Would Motivate Nonparticipating Students to Attend Study Groups

The most common excuse for not taking advantage of the study groups was lack of time (35%). Students felt that their class or work schedule left little time for “extra” study sessions. Some would like to have study groups right after class when they are already on campus. They also expressed a need for more advertising of times and locations (e.g., during every class period; Figure 7: C1). Many students (23%) wanted to receive extra credit for their attendance, get information on exam questions, or have the study groups be mandatory and part of their grade (Figure 7: C2). Some students (15%) did not trust a classmate to be a preceptor and wanted an experienced study group leader (e.g., a faculty member who had all the knowledge, and specifically prepared them for the exam; Figure 7: C3). Another 11% of nonparticipants reported that they would attend study groups if they received low grades or were having trouble with a particular concept in class (Figure 7: C4), while another 8% requested proof that study groups would be beneficial (Figure 7: C5). Finally, 8% reported that studying alone or with a group of people they already knew was more helpful to them than studying with a group of unfamiliar people (Figure 7: C6).

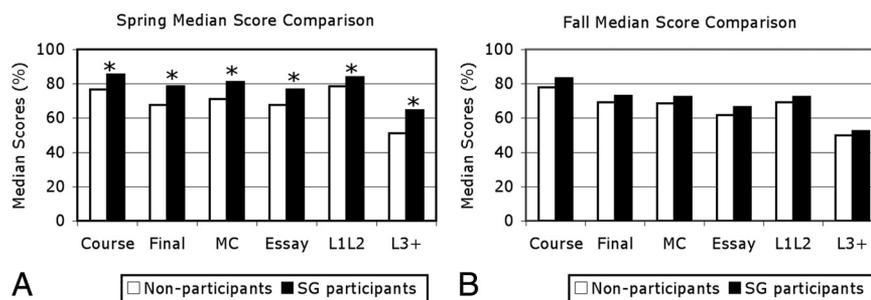
DISCUSSION

The Teaching Team model of peer instruction can indeed be an effective learning tool for large lecture classes. Both spring and fall participants reported a sense of community and valued the opportunity to exchange information with peers and to practice learning in their groups. They viewed study groups as a safe environment to expose what they did not know and as a motivation to stay current with the lecture material and be better prepared for class. In short, preceptor-facilitated study groups increased student engagement in class.

Higher Class Retention Rates for Study Group Members

Students across a wide range of abilities participated in study groups, and participation in study groups was associated with higher course completion in both semesters (Figure 3). Peer-led study groups could thus be an effective retention “tool” for science classes and college in general. For example, college dropout rates are highest during the first 2 years in 4-yr public institutions (NCES, 2003), and one of the reasons that are cited by college dropouts is academic challenge (see also Lang, 2008). By motivating students to stay in class while supporting their learning with study groups, we can address this issue while still challenging our

Figure 6. Median score comparison (% possible scores) of study group participants versus nonparticipants for the course, the final exam, and individual question groups within the final exam: MC questions, essay questions (all), and essay questions by Bloom levels (lower: L1L2 and higher: L3+); (A) spring (N = 103), (B) fall (N = 240). The direct comparison (not shown) between all spring (N = 53) and fall (N = 65) study group participants showed no difference in their GPAs (MWU = 1559, $p = 0.375$), but the spring participants outperformed the fall participants in MC scores (median scores: spring: 80.5, fall: 71.43; MWU = 1240.5, $p < 0.001$), and on the essay scores on all learning levels (median scores: L1L2 spring = 83.93 versus fall = 70.77; MWU = 824, $Z = -6.967$, $p < 0.001$; L3+ spring = 64.4 versus fall = 50.0; MWU = 1453, $Z = -4.473$, $p < 0.001$).



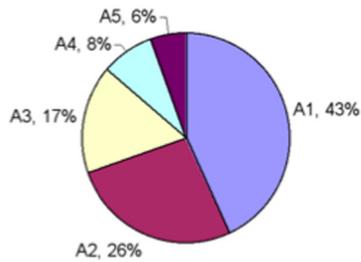
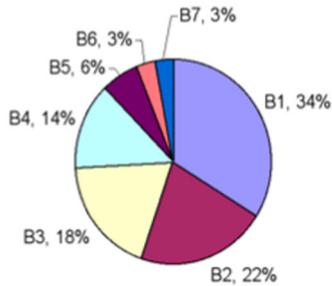
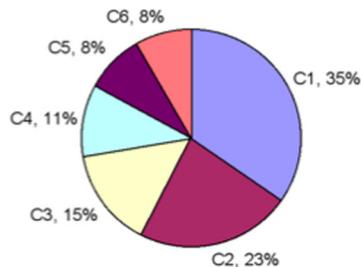
What Students Find Helpful about the Study Groups**What Students Would Improve about Study Groups****What Would Motivate Students to Attend Study Groups**

Figure 7. Student feedback on study groups (Fall semester). A: What students found helpful about study groups. A1: work with peers for deeper understanding, A2: clarify information, A3: work with worksheets to transfer knowledge, A4: motivation to keep up with reading and prepare for class, A5: less intimidating to ask questions and get one-on-one feedback. B: What students would improve about study groups. B1: better organization, more discussion and active learning, B2: provide more worksheets or test questions, B3: more consistent student attendance and more interaction, B4: preceptors should be more assertive and know the answers, B5: nothing to improve, work well, B6: meet longer or more often, B7: more available study group times. C: What would motivate nonattending students to attend study groups. C1: nothing: too busy, C2: extra credit, part of grade or information on exam content, C3: faculty should lead study group and prepare for exam, C4: low scores on exams, C5: proof that study groups will pay off, C6: nothing: prefer to study alone or with friends.

students to improve their learning and to develop critical thinking skills.

Course Performance

Spring study group participants tended to score significantly better on exams and earn higher course grades than

students who did not participate in peer-facilitated study groups. This effect was stable when controlling for overall motivation of students as measured by lecture attendance and engagement, both of which were positively associated with course performance in this and previous studies (see Moore, 2003; Knight and Wood, 2005; Freeman *et al.*, 2007). The study group effect on student learning was also stable when controlling for student ability. Even though better students were more likely to attend study groups, the better final exam scores of SG participants in the spring semester could not be explained by their ability (GPA) alone. In fact, SG participation consistently showed a larger impact (F-values) on student scores than did GPA (Table 4). After seeing the clear learning gains from study groups in the spring semester, we expected a similar outcome in the fall and were surprised to find that the better performance of the fall SG participants was almost entirely driven by student ability (GPA differences between participants and nonparticipants). After controlling for GPA, study group effects were not statistically significant (Table 4), and the marginal means of the performance scores of SG participants and nonparticipants were very similar (Table 5).

Possible Influences on Study Group Effects in Spring and Fall

Even though all students learned about Bloom's six cognitive levels (Bloom, 1956; Anderson and *et al.*, 2001) at the beginning of each semester and used this knowledge to communicate with the instructor about learning, this knowledge alone was not sufficient for most students to practice and master these learning skills on their own when studying for this course. However, spring semester students who took advantage of preceptor-led study groups to practice and get feedback on their learning skills did significantly better on the final exam questions than students who did not. In contrast it made no significant difference for the fall students. We discuss sampling issues and student experience as potential contributing factors to this difference below.

Sampling

In spring, fewer students consented to participate in this study ($N = 110$ or 42.9%) compared with fall ($N = 275$ or 90.5%). This difference in sample size could have an effect on the outcome, especially if better students are more likely to consent than weaker students. To address this issue, we took the anonymous final exam scores (MC% and Essay%) from the entire spring class and removed the matching scores of the consenting students. The remaining scores provided an anonymous summary sample of the nonconsenting students (for whom no GPA data or other demographic data including SG participation were accessed). When comparing the consenting ($N = 104$) and nonconsenting ($N = 118$) spring students we found that these two samples did not differ significantly from each other, neither in the final exam MC question scores (MWU = 5405, $p = 0.126$), nor in the essay question scores (MWU = 5333, $p = 0.093$). The same was the case for fall (15 nonconsenting students: MC: MWU = 1374.5, $p = 0.153$ and essay: MWU = 1427, $p = 0.216$). As a result, we can rule out sampling issues as a source for the differences between the two semesters.

Experience

Another possible explanation for the lack of significant study group effects in fall is that the more experienced fall students (80.4% sophomores, juniors, or seniors) already had better thinking and study abilities than spring students (71.8% freshmen), and as a result study group participation had no further effect. We investigated this possibility by analyzing the (more or less identical) cumulative final exams. As a reminder, on these exams, the spring SG participants had shown significantly better exam scores (on all measures) than nonparticipants, and had also demonstrated significantly better thinking skills on the essay questions, both lower- and higher-level, than nonparticipants. In contrast, fall participants only had outscored nonparticipants in total exam score (MC and essay scores were marginally nonsignificant), and this was mainly driven by a significant difference in lower-level (remembering and basic understanding) written essay scores. If the lack of a study group effect in fall was due to better student ability, we would expect fall students to outscore the spring students on the final exam, and we would expect fall participants to match or exceed the performance of spring participants, especially on the higher thinking levels.

Overall, the spring students ($N = 90$) had significantly higher means on all final exam and class scores than the fall students ($N = 171$; Brown Forsythe: all $p < 0.001$). These results suggest that both basic understanding and critical thinking skills were more developed in the spring SG participants, with no difference in pre-existing ability (GPA). Thus we have to reject better thinking skills as an explanation for the lack of study group effects in the more experienced fall students. Based on the estimated marginal means (controlling for GPA) of final exam scores (Table 5), it appears that the fall students (both participants and nonparticipants) performed on the level of the nonparticipants in spring, and thus the difference between semesters was due to a lack of a significant study group effect on participants in fall.

Even though this lack of a significant fall study group effect is somewhat disconcerting for any instructor trying to decide whether or not to implement a Teaching Team approach in his/her own lecture class, this situation provides us with a unique opportunity to investigate which factors lead to improved learning gains in preceptor-led study groups and what to avoid. We identified four likely and potentially interrelated variables that could explain the higher impact of study group participation in spring: differences in (1) study group logistics, (2) preceptor training, (3) student experience, and (4) student demographics.

First, fall study group preceptors tended to prepare worksheets for their study groups, which was rarely the case in spring. While many participants appreciated this effort by the preceptors (Figure 7: A3), it may have allowed some fall participants to be less active learners during study groups and to focus less on their own misconceptions and learning difficulties than spring participants. Second, while the spring preceptor training was offered throughout the semester as part of the weekly Teaching Team meetings, the fall training by Learning Center staff was centralized in a series of workshops (Table 1) for preceptors from different courses (in biology, chemistry and mathematics). The centralized

workshops decreased the relevancy of the preceptor training to the individual classes for which the preceptors were trained, and possibly prepared them less to facilitate higher-level learning in the context of this class. Third, significantly more sophomores and juniors took the class in fall than in spring, and consequently fall students had spent more time at the university (median: 3rd semester). It is possible that these generally older fall students were more resistant to a change in their memorization-based (i.e., passive) study habits (e.g., Holschuh, 2000, Weimer, 2002), as suggested by theories of cognitive growth (e.g., Piaget, 1971, Ruble, 1994, Hurtado *et al.*, 2003). Theories of cognitive growth emphasize the importance of discontinuity and uncertainty for cognitive development (e.g., Piaget, 1971, Ruble, 1994). Applied to higher education, the first year of college can play the role of such a critical period (Ruble, 1994): in this period classroom and social relationships that challenge rather than replicate the ideas and experiences students bring with them from their home environment are especially important in fostering cognitive growth (Hurtado *et al.*, 2003). Could we have missed this window of opportunity for our older students? While this is certainly a possibility, further study is needed. Finally, even though student demographics always change between spring and fall semesters, the extreme difference seen in the present study was somewhat atypical due to a curriculum change: The fall course in this study was the last 2-h course offered before being transformed into a 3-h course the following semester. As a result, many of the students who needed only two more credit hours to fulfill their science requirements rushed to take the fall course. This turned out to be mostly (77.1%) nonbiology (e.g., chemistry) majors, contributing to a large nonmajors population (86.16%) in the study groups, compared with a more even representation of nonmajors (54.72%) in spring. It is important to note that there was no difference in ability (GPA) between majors and nonmajors who participated in study groups: (spring: $MWU = 342.5$, $p = 0.921$; fall: $MWU = 232.0$, $p = 0.704$), and there were no significant differences in any of their final exam and class scores (spring: $MWU > 256.0$, $p > 0.099$; fall: $MWU > 191.0$, $p > 0.247$). Instead, the large majority of nonmajors in study groups may have influenced the agenda and the learning goals of the groups. Nonmajors likely have a different level of interest for engaging with the material and a different perception of the relevancy of the class material for their planned career paths. This most likely will affect the emphasis on active learning and higher-level thinking skills in study groups. For example, compared with nonparticipants, $\approx 18\%$ more spring participants reported being more active learners and having a good grasp of the class material at the end of class; in fall this difference was only $\approx 8\%$. This is reflected in the lack of a difference in critical thinking skills (L3+ essay questions) between study group participants and nonparticipants in fall (Table 3). Due to this somewhat atypical student population in the fall semester of our study, we expect that the potential differences between spring and fall semesters in the same course would ordinarily be much less pronounced.

We have shown consistent benefits (e.g., engagement, retention) for students in both semesters of the course and have further identified ways to improve our use of study groups in the future: Based on the successful peer-facilitated study groups in spring and lessons learned from the study

groups in the fall, we recommend to maximize the effect of peer-facilitated study groups in future implementations by (1) ensuring that study group members work together on creating their own worksheets, rather than preceptors providing them to their group, (2) offering class-centered preceptor training, (3) exposing students to active learning and higher-level thinking in study groups early in their college career (as freshmen), and (4) maintaining a relatively even match of biology and nonbiology majors in study groups (for classes that serve both demographics).

The Teaching Team Approach and Student Learning

Peer-led study groups as shown in previous and in the present study facilitate student learning. Similar to other peer-instruction models (e.g., see Crouch and Mazur, 2001; Eberlein *et al.*, 2008) our SG participants benefited more from study groups (as measured by performance on final exam questions), the more study groups they attended. It is noteworthy that a significant learning gain in the spring semester occurred after participation in only one study group, whereas other peer-instruction models (focusing on answering questions on instructor-provided worksheets) report learning gains only after at least six study group sessions were attended (e.g., Sharma *et al.*, 2005). We propose that this effect is due to a combination of metacognition, i.e., teaching students about the cognitive (Bloom) levels of learning (i.e., how to learn), and the emphasis on active learning for member-selected activities in study groups. Active cognitive learning is a prerequisite for critical thinking and thus should facilitate student performance on higher-level exam questions as well as lower-level (remembering) questions. In contrast, practice alone (without metacognition) would be expected to improve lower-level skills (re-

membering facts and explanations) without necessarily affecting higher-level thinking.

The possible drawback of this student-centered approach of the Teaching Team Program is the lack of learning gains if members (and preceptors) don't take advantage of this student-centered design. In this case (as demonstrated in the fall) study group members don't seem to study or prepare any differently than the other students in class, and no advantage is gained over studying alone or in private study groups. One of the desired consequences of reduced instructor control in peer-facilitated study groups is increased student responsibility for learning (creating more self-regulated learners; e.g., Schunk, 2001), therefore it is crucial that we impress on preceptors and students the importance of active learning, including critical-thinking activities (Bloom levels 3–6), in study groups. We suggest that preceptors make a habit of asking students to identify the learning levels of all their study group activities to bring more awareness to the level of thinking being practiced and to help them make the transition to critical thinking. A recent publication on how to implement Bloom's taxonomy in biology classes (Crowe *et al.*, 2008) provides a good overview on how this can be done.

While learning gains in some peer-instruction models seem to be especially beneficial for the weakest students in the class (Born *et al.*, 2002; Peters, 2005), we found that preceptor-led study groups, as implemented in the spring semester, facilitated student learning across all ability levels. This included improved critical-thinking skills.

Even though this study tested the Teaching Team approach in an introductory biology class, the same model is very likely beneficial for upper-level classes as well. Different versions of peer instruction have been shown to be successful for remedial purposes, for better knowledge re-

Table 7. Preceptor evaluation of the teaching team program (spring semester)

| Being a preceptor has . . . | Average rating (1 = strongly disagree to 4 = strongly agree) |
|---|--|
| Learning and performance outcomes: | |
| Kept me on top of my work for this class. | 4 |
| Improved my understanding of the course material. | 3.7 |
| Increased my confidence as a student. | 3.5 |
| Developed my time management skills. | 3 |
| Improved my performance in other classes I am taking. | 3 |
| Teaching skills and experience: | |
| Given me insight and knowledge about the teaching process. | 3.8 |
| Improved my ability to recognize when students need help. | 3.5 |
| Improved my ability to present material in an organized and understandable manner. | 3.5 |
| Improved my ability to explain complicated ideas to others. | 3.5 |
| Given me experience teaching others. | 3.5 |
| Increased my interest in teaching as a profession. | 2.6 |
| Interpersonal outcomes and skills: | |
| Allowed me to get to know the professor better. | 3.9 |
| Improved my leadership skills. | 3.7 |
| Helped me learn to work effectively with people of different background and opinions. | 3.6 |
| Improved my ability to moderate group discussions. | 3.6 |
| Allowed me to get to know more of my classmates than usual. | 3.5 |
| Improved my ability to give clear, honest, supportive feedback. | 3.5 |
| Developed my reflective listening skills; listening without making judgments about what I hear. | 3.4 |
| Improved my teamwork skills. | 3.3 |

tention, for the practice of course-specific problem-solving skills, and for critical thinking. Which peer-instruction model ultimately works best for individual classes will depend on the specific goals of the instructor and the learning objectives for the class.

The Teaching Team model of peer instruction as implemented in this study provided benefits to all team members: students, preceptors (Table 7), TAs, and the instructor. One of the biggest benefits is regular communication between preceptors and instructor (and TAs) during the weekly Teaching Team meetings (see also Platt *et al.*, 2006). Through the preceptors the instructor can reach many more students in class with regard to study strategies and possible misconceptions than would be possible through office hours alone. The preceptors also provide valuable feedback to the instructor (and TAs) on how students are learning and what difficulties they are having. This allows for a continuous assessment of teaching goals and student learning in the respective class (lecture, discussions, or labs).

Lessons Learned

When asked how study groups could be improved, some study group members suggested longer (more than 60 min) or more frequent meetings, which speaks to the success of their study groups. Some asked for a more consistent (stable) group membership and more active participation by all attendants. This could be achieved by requiring students to attend the same study group, but at the cost of reduced flexibility in student schedules, which may be counterproductive in large classes. A more desirable approach might be to ask all study groups to generate their own rules and a list of desired behaviors (such as participation). Other concerns point toward a misconception about the role of each member in our study groups: participants wished that more structure (e.g., a list of questions) was provided for each session or felt that preceptors should know the answers to all questions. We will have to work on eliminating the apparent misconception that preceptors are in charge and have to be all-knowing and answer-providing rather than being facilitating members of the study group. Ideally, the entire group should set the agenda at the beginning of each session, and the process of figuring out the answers to questions, rather than just knowing the answer, should be a group effort. These points can be addressed more specifically during preceptor training and by the preceptors themselves at the beginning of each study session.

Increasing Participation in Study Groups

Most commonly nonparticipants saw study groups as “extra” study time, a perception that could be easily addressed in the classroom. Other nonparticipants (34%) displayed a passive and unengaged attitude toward learning when explaining what it would take for them to participate in study groups: “if I have to” (e.g., due to low exam scores), “make me” (e.g., attendance as part of grade), or “reward me” (e.g., give extra credit, free food, or tips on exam questions) were common themes. Similarly, nonparticipants wanted a priori proof that study groups would indeed benefit them. These statements expose a lingering consumer attitude toward learning rather than a sense of personal responsibility

among the nonparticipants. This has been documented for other science classes as well (e.g., Lord, 2008). We will have to address this issue more directly in our classes if we want to ensure that as many students as possible receive the benefits of study groups without making them mandatory. With the present study we have addressed at least one of the concerns of the nonparticipating students: this study clearly shows that students who participate in study groups that emphasize active learning and higher-level thinking tend to score higher on exams and earn better grades than nonparticipating students.

ACKNOWLEDGMENTS

We are grateful to all the graduate students who served as TAs in this class and were willing to dedicate an additional hour of their time for our weekly Teaching Team meetings and to serve as mentors for our preceptors. We also thank Rachel Zierow, who initiated this program at the University of Texas, and the other Learning Center staff for their support of the Teaching Team Program and for providing leadership training for the preceptors. But most of all we want to thank those students who volunteered to serve as preceptors, and all the students who consented to participate in this study (IRB# 2004-09-006). Thanks also to Gary Reiness and two anonymous reviewers for constructive comments on an earlier version of this manuscript and the members of the UGA Science Education Research Group for helpful comments and suggestions.

If you would like more information on establishing peer-facilitated study groups in your course, please consult the following web resources (e.g., University of Texas: www.utexas.edu/student/utlc/study_groups/, University of Arizona: <http://teachingteams.arizona.edu/>), or contact the authors.

REFERENCES

- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., and Wittrock, M. C. (Eds.). (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, New York: Addison Wesley Longman.
- Angelo, T. A., and Cross, K. P. (1993). *Classroom Assessment Techniques* (2nd ed.), San Francisco: Jossey-Bass.
- Bloom, B. S. (Ed.). (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*, Chicago: Susan Fauer Co.
- Born, W. K., Revelle, W., and Pinto, L. H. (2002). Improving biology performance with workshop groups. *J. Sci. Educ. Tech.* 11, 347–365.
- Boyer Commission on Educating Undergraduates in the Research University (1998). *Reinventing Undergraduate Education: A Blueprint for American Research Universities*. <http://naples.cc.sunysb.edu/Pres/boyer.nsf> (accessed 9 September 2007).
- Crouch, C. H., and Mazur, E. (2001). Peer instruction: ten years of experience and results. *Am. J. Phys.* 69, 970–977.
- Crowe, A., Dirks, C., and Wenderoth, M. P. (2008). Biology in Bloom: implementing Bloom's taxonomy to enhance student learning in biology. *CBE Life Sci. Educ.* 7, 368–381.
- Eberlein, T., Kampmeier, J., Minderhout, V., Moog, R. S., Platt, T., Varma-Nelson, P., and White, H. B. (2008). Pedagogies of engagement in science. *Biochem. Mol. Bio. Educ.* 36, 262–273.
- Ebert-May, D., and Brewer, C. (1997). Innovation in large lectures-teaching for active learning. *Bioscience* 47, 601–607.
- Freeman, S., O'Connor, E., Parks, J. W., Hurley, D., Haag, D., Dirks, C., and Wenderoth, M. P. (2007). Prescribed active learning in-

- creases performance in introductory biology. *CBE Life Sci. Educ.* 6, 132–139.
- Goodlad, S. (1998). The effectiveness of peer tutoring in higher education: a typology and review of the literature. In: *Mentoring and Tutoring by Students*, ed. Sinclair Goodlad, London: Kogan Page, 49–69.
- Gosser, D. K., Cracolice, M. S., Kampmeier, J. A., Roth, V., Strozak, V. S., and Varma-Nelson, P. (2001). *Peer-Led Team Learning*, Upper Saddle River, NJ: Prentice Hall.
- Hockings, S. C., DeAngelis, K. J., and Frey, K. J. (2008). Peer-led team learning in general chemistry: implementation and evaluation. *J. Chem. Educ.* 85, 990–996.
- Holschuh, J. P. (2000). Do as I say, not as I do: high, average, and low-performing students' strategy use in biology. *J. Coll. Read. Learn.* 31, 94–108.
- Hurtado, S., Dey, E. L., Gurin, P. Y., and Gurin: G. (2003). College environments, diversity and student learning. In: *Higher Education Handbook for Theory and Research*, Vol. XVIII, ed. J. C. Smart. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Knight, J. K., and Wood, W. B. (2005). Teaching more by lecturing less. *Cell Biol. Educ.* 4, 298–310.
- Lang, S. A. (2008). *Stories of Staying and Leaving: A Mixed Methods Analysis of Biology Undergraduate Choice, Persistence, and Departure*. Ph.D. dissertation, The University of Texas at Austin, United States – Texas. Dissertations & Theses: Full Text database. (Publication No. AAT 3315093) (accessed 25 February 2009).
- Larsen, H. P., Mencke, R., Tellefson, S. J., Harrison, E., and Berman, E. (2001). The teaching teams program: a "just-in-time" model for peer assistance. In: *Student-Assisted Teaching: A Guide to Faculty-Student Teamwork* ed. J. E. Miller, J. E. Grocia, and M. S. Miller, Bolton, MA: Anker Publishing Co, 27–33.
- Libarkin, J., and Mencke, R. (2001). Developing training for undergraduate peer-teachers in an innovative program in undergraduate education. *J. Coll. Sci. Teach.* 31, 235–239.
- Lord, T. (2008). "Darn it, professor. Just tell us what we need to know to pass your course." *J. Coll. Sci. Teach.* 37, 71–73.
- Mazur, E. 1997. *Peer Instruction: A User's Manual*, Prentice Hall, Upper Saddle River, NJ.
- Miller, J. E., Grocchia, J. E., and Miller, M. S. (2001). *Student-Assisted Teaching: A Guide to Faculty-Student Teamwork*, San Francisco: Jossey-Bass.
- Moore, R. (2003). Attendance and performance: how important is it for students to attend class? *J. Coll. Sci. Teach.* 32, 367–371.
- National Center for Education Statistics (NCES) (1995). *Postsecondary Education Quick Information Survey on Higher Education Finances and Services*. Washington, D.C.: NCES. <http://nces.ed.gov/pubs95/95393.pdf> (accessed 14 March 2005).
- NCES (2003). *Short-term Enrollment in Postsecondary Education. Student Background and Institutional Differences in Reasons for Early Departure, 1996–1998*. Washington, D.C.: NCES. <http://nces.ed.gov/pubs2003/2003153.pdf> (accessed 8 July 2007).
- NCES (2008). *Mini-Digest of Education Statistics 2007*. Washington, D.C.: NCES. <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2008023> (accessed 13 March 2008).
- National Science Foundation (1996). *Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology*. Washington, DC: www.nsf.gov/pubs/1998/nsf98128/nsf98128.pdf (accessed 1 April 2008).
- Peters, A. W. (2005). Teaching biochemistry at a minority-serving institution: an evaluation of the role of collaborative learning as a tool for science mastery. *J. Chem. Educ.* 82, 571–574.
- Piaget, J. (1971). The theory of stages in cognitive development. In: *Measurement and Piaget*, ed. D. R. Green, M. P. Ford, and G. B. Flamer, New York: McGraw-Hill.
- Platt, T., Barber, E., Yoshinaka, A., and Roth, V. (2006). An innovative selection and training program for problem-based learning (PBL) workshop leaders in biochemistry. *Biochem. Mol. Biol. Educ.* 31, 132–136.
- Prince, M., and Felder, R. (2007). The many faces of inductive teaching and learning. *J. Coll. Sci. Teach.* 36, 14–20.
- Ruble, D. N. (1994). A phase model of transitions: cognitive and motivational consequences. *Adv. Exp. Soc. Psychol.* 26, 163–214.
- Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In: *Self-Regulated Learning Academic Achievement: Theoretical Perspectives* (2nd ed.), ed. B. J. Zimmerman and D. H. Schunk, Mahwah, NJ: Erlbaum, 125–151.
- Sharma, M. D., Mendez, A., and O'Byrne, J. W. (2005). The relationships between attendance in student-centered physics tutorials and performance in university examinations. *Int. J. Sci. Educ.* 27, 1375–1389.
- Sokal, F. J., and Rohlf, R. R. (1994). *Biometry: The Principles and Practice of Statistics in Biological Research*, 2nd edition, New York: W. H. Freeman and Co.
- Weimer, M. E. (2002). *Learner-Centered Teaching. Five Key Changes to Practice*, San Francisco: Jossey-Bass.