

Retention and Reform: An Evaluation of Peer-Led Team Learning

Scott E. Lewis*

Department of Chemistry and Biochemistry, Kennesaw State University, Kennesaw, Georgia 30144, United States

ABSTRACT: This study describes an evaluation of the Peer-Led Team Learning (PLTL) reform model in first-semester general chemistry. PLTL was implemented in place of one-third of the available lecture time, maintaining the same amount of structured class time under the reform. The evaluation demonstrates that classes implementing the PLTL reform at the setting featured a statistically significant improvement of 15% in the pass rate for the classes, compared to conventional, lecture-only classes at the same setting. Additionally, the PLTL classes maintained a comparable score on a comprehensive American Chemical Society final exam to that of the conventional classes. Combined, these findings suggest that the improvement in pass rates occurs while maintaining the same level of rigor in the PLTL classes. Examining student groups shows that students in underrepresented minority groups demonstrated the largest improvement in pass rates with the PLTL reform.

KEYWORDS: First-Year Undergraduate/General, Chemical Education Research, Collaborative/Cooperative Learning, Constructivism, Minorities in Chemistry, Women in Chemistry

FEATURE: Chemical Education Research

General chemistry at the university level has been described as functioning as a gatekeeper, preventing many students from successfully progressing in science and science-related careers.¹ As a result, evaluations of reform pedagogy often focus on student retention or the percent of students passing the course. This report provides an evaluation of a reform pedagogy as it relates to both student retention and student performance, which are intertwined measures necessary to consider in evaluation of pedagogical reforms. The reform pedagogy to be evaluated is peer-led team learning (PLTL), a nationally disseminated, reform teaching method that employs peer leaders as a means to incorporate active learning sessions into a course.² Peer leaders are undergraduate students who have successfully completed the target course for the PLTL implementation. Through the use of peer leaders, PLTL is naturally suited for implementation in large class sizes, which are a common occurrence with general chemistry.

REFORM EVALUATION AND STUDENT RETENTION

An important indicator on the effectiveness of reform pedagogy is to determine the impact on student retention or the percent of enrolled students who complete the course. If the reform is effective, the rate of student retention should improve. Course retention is often characterized by the percent of students passing the course, sometimes measured by DWF rate or the pass rate.³ This measure can be strongly swayed by instructor decisions. For example, the decision to offer extra credit, accept late work, or curve test scores can all promote the pass rate in a class. From a research perspective, controlling for these instructor decisions is often not possible as it removes the instructor from the important role of assigning student grades.

One resolution to this concern is to discern the impact of the reform on student retention and performance on a common exam concurrently. This approach allows for a measure of student retention

while also providing context on the extent that learning objectives have been achieved. Also, this approach offers a benefit over only comparing common exam scores, as student retention may cloud this comparison. For example, a reform pedagogy that leads to an improvement in student averages on a common exam could also have a lower rate of student retention, which could be responsible for the difference in scores observed. It may not be that the reform improved learning for the class; rather the lower rate of students not finishing the class could be responsible for the higher average as fewer students with low ability took the common exam. In short, student retention and students' academic performance are intertwined measures of reform effectiveness that cannot be considered independent of one another.

SETTING AND REFORM PEDAGOGY

The setting for this study is a large, primarily undergraduate institution located in the southeastern United States. The study focuses on the first semester of general chemistry at the institution. Class sizes for general chemistry at the setting range from 50 to 75 students, and in a semester four to eleven classes of first-semester general chemistry are offered. A common syllabus is employed that standardizes the content covered, the grading setup, and the use of a common final exam across the classes. Course content includes the first 11 chapters of a common chemistry textbook.⁴ As part of this common syllabus, the grading system used allocated the point distribution as follows: 60% midterm tests, 20% a common final exam, and 20% at the instructor's discretion. The instructors in the peer-led classes split the last category, assigning 12% for attendance to the peer-led sessions and 8% for homework or in-class quizzes or assignments. The traditional classes used the discretionary 20% for primarily assigning homework, conducting in-class quizzes or

Published: April 05, 2011

assignments, or by taking attendance during lecture. The final exam is the American Chemical Society (ACS) First-Term General Chemistry Exam.⁵ The general chemistry student population at the setting is 56.6% female, 61.9% white, 9.5% Black, 5.3% Asian, and 3.4% Hispanic, with the remaining 20.0% of students identified as mixed race, Native American, or not disclosed.

Traditionally, classes at the setting are offered in two, 75 min sessions per week, which primarily rely on lecture, in part as a result of class size. To implement the PLTL reform, classes were offered that meet for three, 50 min sessions per week, where two sessions were primarily lecture and the third was designated for the PLTL implementation. Thus, reform classes had an equal amount of structured course time to the traditional lecture only classes. Therefore any measurable benefits from the reform pedagogy could not be attributed to a change in the amount of time on task. This is different from other studies which offered PLTL sessions as additions to regular class time.^{6,7} Within the time frame of four concurrent semesters, 21 traditional classes were offered and 8 PLTL reform classes were offered. Students self-selected into the PLTL classes, though the classes were not explicitly identified as different. The PLTL sessions were modeled after the literature⁸ and are described below in terms of each major role in the setting.

Students

For the peer-led session, students were divided into groups of 12–16 students and each group was assigned a peer leader. During the peer-led session, students worked in groups of four assigned by the peer leader. The decision to employ groups of four is the one notable exception from the model in the literature, which describes one peer leader assigned to a single group of six–eight.⁸ Experience in the setting suggested that groups of four made it easier to encourage involvement from every student. Throughout the semester students met in a peer-led session once a week for 14 total sessions. During these sessions students worked on a set of problems created by their instructor. Attendance at the peer-led sessions was a mandatory part of the class. This decision was made to determine the potential for PLTL to assist all students within a class, in contrast to other studies where attendance at PLTL sessions has been optional.^{6,7,9} Attendance at the PLTL sessions counted for 12% of the students' overall grade. Students could miss two peer-led sessions without penalty, but any additional sessions missed resulted in a grade penalty. With this system, 58.0% of the students missed two or fewer sessions for full attendance credit. An additional 18.6% of students attended at least 10 of the 14 sessions.

Peer Leaders

Peer leaders enrolled in a separate, semester-long, training course for which they received upper-level elective credit. The training course was led by the author and was designed to provide pedagogical training, focusing on active learning and cooperative learning techniques, and to incorporate pedagogical techniques with the general chemistry content. To achieve these goals, the training course was modeled as a peer-led session. Peer leaders were placed into groups of four and each week were given the same worksheet their students were about to be given. Peer leaders worked through the worksheet in assigned groups and the instructor would model techniques for promoting group work and active learning. Examples of these techniques include the following:

- requesting group consensus and encouraging the group to compare answers;
- asking students to explain their work and decision-making process;
- asking students to present answers to other groups;
- encouraging students to consult references.

During their time with students, peer leaders were also regularly observed by the author and PLTL class instructors. The observer would provide constructive feedback to the peer leader upon completion of the peer leaders' current session with students. Within the training course, peer leaders were assigned a weekly reflective journal on various aspects of their sessions with students. Peer leaders were also assigned other peer-led sessions to observe and complete an observation template directed toward the extent group work is facilitated.

Instructors

Each semester, the instructors of the PLTL classes would meet weekly as a group to design a common worksheet for students to work on. Typically, the worksheet was designed to follow the material presented in lecture rather than introducing new content to students during the sessions, as prescribed in the literature.⁸ Each worksheet would generally range from 6 to 12 questions, with an effort to include both algorithmic and conceptual questions. The instructors' meetings ensured that the worksheets matched the concepts and procedures presented in class, and to account for variations in timing across the PLTL classes. It also ensured instructors were familiar with the worksheet and could refer to them in class. During the peer-led sessions instructors would regularly observe the peer leaders working with students. Effort was made to ensure that instructors did not observe their own students, as this would often provide a strong distraction to the session. Based on the observations, instructors could provide feedback to the peer leaders. Instructors would also regularly check to make sure all peer leaders were present for their own students, which provided a reinforcing image to the students that the instructor valued these sessions. On rare occasion a class instructor would fill in for a peer leader who was absent.

RESULTS

The reform was evaluated on the class level where each class was considered an observation. Analysis at the class level was chosen over the student level to ensure the independence of observations assumption in statistical models. Students within a class are likely not independent observations as it relates to chemistry performance, particular with the PLTL pedagogy that emphasizes group work.¹⁰ Student retention was measured in terms of both percentage of students who took the ACS final exam and percentage of students who received a passing grade in the class. The percentage of students taking the ACS final exam provides a picture of retention, as students who withdraw from the course or resigned themselves to not achieving a passing grade do not take the final exam. This is an important measure of student retention as it is less likely to be influenced by instructor decisions as the percent of students passing the class. Additionally, the percentage of students taking the final exam is necessary information when comparing class averages on the ACS final exam, as discussed before. The percentage of students passing the class is also an important outcome in its direct impact on student progress within a science curriculum.

Table 1. Comparison of Traditional and PLTL Classes

	traditional instruction (SD)	PLTL instruction (SD)
no. of classes	21	8
av class size	65.7 (8.1)	68.9 (4.7)
av on ACS exam	53% (3.4%)	54% (2.9%)
% taking ACS exam ^a	72% (10%)	84% (5.4%)
% passing the class ^a	53% (10%)	68% (9.2%)
SAT Mathematics av score	545 (13)	542 (11)
SAT Verbal av score	536 (12)	534 (6.4)

^a Statistically significant difference, $p < 0.05$.

Table 2. Comparison with Common Instructors

	traditional instruction (SD)	PLTL instruction (SD)
no. of classes	6	8
av class size	66.0 (9.6)	68.9 (4.7)
av on common final exam	53% (3.7)	54% (2.9%)
% taking final exam ^a	76% (8.0%)	84% (5.4%)
% passing the class ^a	58% (7.8%)	68% (9.2%)
SAT Mathematics av score	537 (10.1)	542 (11.0)
SAT Verbal av score	530 (13.3)	534 (6.4)

^a Statistically significant difference, $p < 0.05$.

The average percentage of correct responses on the ACS final exam was also measured to determine the extent to which learning objectives have been met and whether changes in the percentage of students passing the class were simply the result of instructor differences in what constituted a passing grade. Additionally, class averages for SAT Mathematics and SAT Verbal tests were collected to evaluate the possibility that one group of classes began with better-prepared students than the other. SAT subscores were chosen as they represent a well-documented predictor of student performance in general chemistry.¹¹ A test for outliers, which examined for any outcome scores that were three standard deviations or more away from the mean, showed no classes to represent an outlier.

The results in Table 1 show a 12% higher rate in the percentage of students taking the ACS exam with the PLTL classes than with the traditional format. Using an independent sample t -test, this difference was found to be statistically significant with a t -score of 3.41 ($p < 0.05$). Calculating effect size using Cohen's d the effect size was found to be 1.49, which is classified as a large effect size, or noticeable to a casual observer.¹² Similarly, the PLTL classes featured a 15% improvement in the percentage of students who received a passing grade in the class. This difference was also statistically significant with $t = 3.69$ ($p < 0.05$, Cohen's $d = 1.56$, large effect size). Both measures indicate an improvement in student retention in the course with the PLTL reform. Given that 551 students had enrolled in the eight peer-led classes, the increase in retention represents an additional 83 students passing the course who would not have passed given the percentage passing with the traditional pedagogy. The traditional and PLTL classes were found to be equivalent on their performance on the ACS Exam, using the two one-sided t -test measure for equivalence.¹³ For this test, a medium effect size was chosen to establish a range of differences $[-1.64, 1.64]$, which represents one-half of the standard deviation. The difference in ACS Exam scores was tested against the boundaries of this range with an $\alpha = 0.10$, and the results suggest a <10% chance that the difference in ACS Exam scores exceeded each end of the established range.

One challenge to the validity of the results is the possibility that students in the PLTL classes were better prepared prior to entering the class. As noted above, students self-selected into the peer-led class. To investigate student performance prior to the class, students' SAT subscores were collected from institutional records for both the PLTL classes and the traditional classes. The averages of each class's SAT Mathematics and SAT Verbal test averages are reported in Table 1. Equivalence could not be established on the SAT subscores using the same procedure as used for the ACS Exam, likely owing to greater variation among SAT subscores. Still, this result does not show evidence of a difference between the two. It is noted that the traditional class was higher on average SAT subscores than the PLTL class, so it is unlikely that the improved retention of the PLTL classes can be attributed to better incoming student preparation.

Another potential challenge to the validity of the results is the role of the instructor. It is possible that the PLTL classes were led by a subset of instructors who may have influenced the results independent of the effect of the pedagogy. Within the setting, the eight PLTL classes were led by three different instructors. During the time frame of this study, the same three instructors also led six of the 21 traditional, lecture-based classes, either during different time slots or during different semesters. To examine the possibility of instructor impact on the results, the PLTL classes were also compared to only the traditional classes led by one of the same three instructors. The results of this comparison are shown in Table 2.

The results in the common instructor comparison in Table 2 are similar to the results in the overall comparison in Table 1. Retention improves on both measures with the PLTL pedagogy, though the improvement is not as pronounced as in Table 1, but still a 10% higher pass rate is observed. The difference in retention remains statistically significant with a t -score of 2.44 ($p < 0.05$, Cohen's $d = 1.17$, large effect size) for the difference in percent taking the final exam and a t -score of 2.19 ($p < 0.05$, Cohen's $d = 1.17$, large effect

Table 3. Pass Rates of Student Groups

groups	% passing (<i>N</i>)	
	traditional instruction	PLTL instruction
Asian and white	55.0 (876)	66.0 (421)
underrepresented minorities ^a	47.0 (185)	64.7 (68)
male	55.8 (574)	68.6 (264)
female	51.4 (735)	63.7 (358)
overall	53.3 (1309)	65.8 (622)

^a Underrepresented minorities as defined by the National Science Foundation and representing Black, Hispanic, and American Indian students.¹⁴ Note: Pacific Islander was not an available option in the data.

size) for the percentage passing the course. The averages on the final exams are within 1%, indicating that the level of performance in the class was maintained. SAT subscores feature a difference of 5 points or less between the traditional and PLTL classes. It is unlikely that SAT subscores accounted for the greater course retention in the PLTL session as correlations indicate that class average SAT subscores, while correlating to performance on the final exam, did not correlate to the measures of student retention.

An additional concern to the validity of the impact of PLTL may arise from the use of 12% credit for attendance at the PLTL sessions. Credit for attendance could have provided an added incentive for students in the PLTL class to attend more regularly, leading to the improved retention. As mentioned, all classes had 20% of students' grades to be determined by techniques at the instructors' discretion. The PLTL classes used 12% of the 20% for attendance and the remaining 8% for homework or in-class assignments. The traditional classes employed the entire 20% primarily for homework, in-class quizzes or assignments, or, less commonly, for taking attendance in the course. Because the majority of homework was turned in during class, the process of assigning and collecting homework also provided an incentive to attend class, albeit an indirect incentive compared to the direct method of taking attendance. Students submitting homework outside of class meetings was an overall rare occurrence. The other techniques used, in-class assignments or taking attendance, provided a direct incentive for attendance. As a result, it can be argued that the use of the discretionary 20% in both types of classes, either directly or indirectly, provided an incentive for students to attend classes. No instructors in the setting reported using online assignments.

One area of interest with improved student retention is the impact on groups that are traditionally underrepresented in the sciences. The percentage of students passing from each student group is shown in Table 3. Because individual student groups are not well represented in each class (for example, some classes did not have any Hispanic students), pass rates for each group are presented as a percentage of the total number of students. The numbers for the traditional column represent the students from all 21 of the traditional classes. As it is probable that these are not independent observations, inferential statistics are not performed on the data in Table 3.

In Table 3 it appears that students from underrepresented minority groups (as defined by the National Science Foundation) had the largest gains in percentage of students passing, with a 17.7% higher rate of passing with the PLTL teaching model. This placed their success rate in the class approaching the rate of the Asian and white student group. Male students and female students had similar

gains of 12.8% and 12.3%, respectively. Overall, it appears that the gains in passing the course were uniform across student groups, with the strong potential for improving the performance of students from underrepresented minority groups.

DISCUSSION AND CONCLUSIONS

The results indicate a notable 15% improvement in student retention, in terms of passing general chemistry, when using PLTL versus a traditional lecture. This improvement in student retention matches the results seen in other studies,^{3,6,7,9} however this study is differentiated by several features. First, this implementation of PLTL took the place of lecture time available instead of serving as additional time. This distinction underscores the growing body of evidence that class time can be structured more effectively than through the use of only lectures.^{10,15} Second, incoming student preparation, as measured by student SAT scores, was accounted for to eliminate this as a potential confounding variable. Third, the improvement in student retention was found while maintaining performance on a common final exam, eliminating concerns that improvement in student retention may just be a result of differing standards of performance. Finally, this study used each class as an independent observation, instead of students, allowing for a more robust statistical test. Additionally, the results indicate that the PLTL pedagogy has a strong potential for improving the retention of underrepresented minority students in the sciences, as this group features the largest gain in student retention with the PLTL pedagogy.

Improving student retention in first-semester general chemistry is only one important, beginning piece of promoting retention and progression in chemistry and the sciences. Future work will examine the role of PLTL in the number of chemistry and science courses students take in following semesters. It is possible that while more students succeed with the PLTL pedagogy, when students return to more conventional pedagogies the success rates return to normal. Such a study, then, would provide information as to whether improving first-semester general chemistry removes the gatekeeper to studying science, or if it only removes the first hurdle, and curriculum-wide reform is necessary to improve progression.

Another future direction for research is determining the aspects of PLTL that are responsible for the improved gains in retention. It is possible that motivational aspects of cooperative learning, such as self-concept, are responsible for the improved gains in retention.^{16,17} Studies measuring the impact of PLTL pedagogy on student motivation or other aspects of the affective domain may offer support for this framework. Alternatively, the originators of PLTL cite the pedagogy's tendency to work within students' zone of proximal development as described by Vygotsky.¹⁸ Being able to adequately characterize interactions that occur within the zone of proximal development and relate these interactions to student success would offer strong support for this argument. Discerning an accurate theoretical explanation for the benefits of PLTL would offer important information regarding peer leader training and how best to facilitate student group work.

AUTHOR INFORMATION

Corresponding Author

*E-mail: slewis57@kennesaw.edu.

■ REFERENCES

- (1) Tai, R. H.; Sadler, P. M.; Loehr, J. F. *J. Res. Sci. Teach.* **2005**, *42*, 987–1012.
- (2) Gosser, D.; Roth, V. *J. Chem. Educ.* **1998**, *75*, 185–187.
- (3) Gafney, L.; Varma-Nelson, P. *Peer-Led Team Learning: Evaluation, Dissemination, and Institutionalization of a College Level Initiative*; Springer: Weston, MA, 2008.
- (4) Silberberg, M. S. *Chemistry: The Molecular Nature of Matter and Change*, 4th ed.; McGraw-Hill: New York, 2006.
- (5) Examinations Institute of the American Chemical Society Division of Education. *First-Term General Chemistry*; University of Wisconsin–Milwaukee: Milwaukee, WI, 2002.
- (6) Hockings, S. C.; DeAngelis, K. J.; Frey, R. F. *J. Chem. Educ.* **2008**, *85*, 990–996.
- (7) Wamser, C. C. *J. Chem. Educ.* **2006**, *83*, 1562–1566.
- (8) Gosser, D. K. The Peer-Led Team Learning Workshop Model. In *Peer-Led Team Learning: A Guidebook*; Prentice Hall: Upper Saddle River, NJ, 2001.
- (9) Lyon, D. C.; Lagowski, J. J. *J. Chem. Educ.* **2008**, *85*, 1571–1576.
- (10) Lewis, S. E.; Lewis, J. E. *J. Res. Sci. Teach.* **2008**, *45*, 794–811.
- (11) Lewis, S. E.; Lewis, J. E. *Chem. Educ. Res. Pract.* **2007**, *8*, 32–51.
- (12) Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale, 1988.
- (13) Lewis, S. E.; Lewis, J. E. *J. Chem. Educ.* **2005**, *82*, 1408–1412.
- (14) National Science Foundation. *Women, Minorities, and Persons with Disabilities in Science and Engineering*; National Science Foundation, Division of Science Resources Statistics: Arlington, VA, 2007.
- (15) Lewis, S. E.; Lewis, J. E. *J. Chem. Educ.* **2005**, *82*, 135–139.
- (16) Born, W. K.; Revelle, W.; Pinto, L. *J. Sci. Educ. Tech.* **2002**, *11*, 347–365.
- (17) Knight, G. P.; Bohlmeyer, E. M. Cooperative Learning and Achievement: Methods for Assessing Causal Mechanisms. In *Cooperative Learning: Theory and Research*; Praeger: New York, 1990.
- (18) Gafney, L.; Varma-Nelson, P. *J. Chem. Educ.* **2007**, *84*, 535–539.