JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

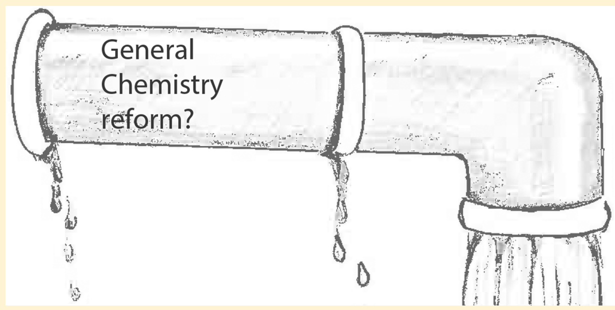
Investigating the Longitudinal Impact of a Successful Reform in General Chemistry on Student Enrollment and Academic Performance

Scott E. Lewis*

Department of Chemistry, University of South Florida, Tampa, Florida 33620, United States

S *Supporting Information*

ABSTRACT: Considerable effort in chemistry education research has been dedicated to developing and evaluating reform pedagogies designed to improve student success in general chemistry. Policy recommendations propose adoption of these techniques as a means to increase the number of science graduates, however there is the potential that the impact of these reforms is mitigated in subsequent classes. This study examines the longitudinal impact of a successful peer-led team learning reform that targets general chemistry. The study uses a quasi-experimental methodology comparing students who took general chemistry with peer-led team learning to students who experienced traditional instruction. Student enrollment and academic performance in subsequent classes were the outcome metrics evaluated. The results found that the reform impacts enrollment in the class that directly follows the target class but enrollment in subsequent classes is mitigated through student attrition within the curriculum. Additionally, no evidence was found for the reform impacting students' academic performance in subsequent classes. The results highlight the need for implementing and evaluating curricular-wide reform.



KEYWORDS: *Chemical Education Research, Collaborative/Cooperative Learning, Student Centered Learning, First-Year Undergraduate/General, Curriculum*

FEATURE: Chemical Education Research

INTRODUCTION

Evaluation of pedagogical reforms in chemistry serves to broaden our understanding of how to improve student success. The evidence generated by these efforts has led to policy recommendations to implement reform in introductory STEM courses to improve student retention and ultimately generate more science graduates. More specifically, the recent "Engage to Excel" report to improve student retention recommends that "STEM faculty learn how to use and incorporate highly effective teaching methods into their introductory STEM courses".[1] However, the evaluations of reform primarily focus on only the course within the curriculum that the reform directly targets, which are often first-year courses. This leads to the potential for an effective reform in the target course, but mitigating effects in subsequent courses and no appreciable impact on student retention within a curriculum. Thus, longitudinal studies that seek to better understand the role of pedagogical reform to influence long-term student retention are needed.[2] This study explores the longitudinal impacts of a successful reform that targets two semesters of General Chemistry.

RATIONALE AND BACKGROUND

The General Chemistry sequence, with high enrollment and low student success rates, is a logical target for efforts to improve student success. As a result, considerable research efforts have been placed on designing and evaluating curricular reforms to promote student success in General Chemistry.[3−13] This body of research includes evaluations of impacts observed within the target course, such as student performance on common exams or percent of students successfully completing the course. This approach is logical in that the first place to look for impact is as proximal as possible to the reform. However, research on the long-term impact of these reforms is much less common and needed to link these reforms to the outcome of improved student retention.

A limited number of longitudinal studies have been enacted to better understand the long-term student gains of academic interventions in the science or engineering fields. Wischusen and Wischusen examined the impact of an intensive week-long orientation prior to postsecondary enrollment.[14] Students were recruited into this program during a separate orientation session for high achieving students. The program was designed for students who intended to major in biology, and the first 60

Published: September 23, 2014

Table 1. Kendall's Tau-b Correlations

Course	General Chemistry I (N)	General Chemistry II (N)
General Chemistry II	0.52 (637)	
Organic Chemistry I	0.47 (315)	0.46 (311)
Organic Chemistry II	0.43 (210)	0.43 (218)
Biochemistry	0.32 (54)	0.13 (81)
Quantitative Analysis	0.46 (57)	0.49 (61)

students to apply were selected for this program. These students were compared to a control group of 70 students that was matched on high school GPA, college readiness tests, intended major, and gender. By narrowing into a subset of biology majors, it was found that 60% of students in the reform enrolled in three biology classes in their first three semesters, compared to 39% of students in the control.

In another longitudinal study, Sorby implemented a semester long course to improve spatial ability skills.[15] The class was available for introductory engineering students who scored low on a spatial reasoning test. In conducting a series of longitudinal investigations, students in the course were compared to a control group of eligible students who had not enrolled in the course. Those taking the course had higher GPAs in subsequent engineering courses. Female students who took the course also had a significantly improved student retention rate measured by enrollment in the university or graduation. A weighted average of the different studies leads to an 18.4% improvement in retention for female students who enrolled in the course (compared to 4.9% for male students).

In chemistry, Hall et al., initiated the Science Advancement through Group Engagement program that began with introductory chemistry and lasted through Organic Chemistry II.[16] Students were eligible for the program based on scoring in the bottom quartile of Math SAT at the research setting and 41 students participated. The program was run outside of normal class times and featured learning specialists and teaching assistants emphasizing class-related group-work. Two control groups were constructed, one based on students not enrolled in the program and one based on historical student records. The historical control group was matched to the reform group based on SAT subscores. The results showed that 68% of the students in the program completed Organic Chemistry II compared to 29% for the historical control and 27% for those not enrolled in the program.

These efforts to examine longitudinal impacts have focused on implementations that occur either outside of class or through enrollment in an elective class. As a result, each study tended to target a distinct group of students, either through intended major or at-risk status. Missing from the research literature are longitudinal studies examining the impact of reforms that target traditional gateway courses, such as General Chemistry. Conducting such investigations are important to provide a better understanding of the long-term impact of the substantial efforts that have been made to reform gateway courses and ultimately has the potential to improve student retention for a larger, more general, population of students.

Further impetus for this work comes from the observed Kendall's Tau-b correlations between student grades in General Chemistry and their grades in subsequent chemistry courses at the research setting in Table 1 (the sample for this data are identified in the Methods section).

The correlations observed describe a medium-strength relation between academic performance in General Chemistry and most follow-on courses. This suggests that a reform, which improves student performance in General Chemistry, has the potential to improve student performance throughout the chemistry curriculum. Correlations do not necessarily indicate causation, however. If there is a causal relationship, then improving student performance in General Chemistry should lead to improved performance in subsequent classes, either through content retention or the development of transferrable skills. However, there is the potential for a spurious relationship, where the correlation results from a third factor. For example, it could be that academically talented students happen to do well in both General Chemistry and in subsequent chemistry courses. In this case the correlations observed are just representative of overall academic talent. One method to investigate the possibility of causation is to enact a change that impacts General Chemistry performance, such as a pedagogical reform. If the change in General Chemistry impacts student performance in subsequent courses then there is evidence for a causal relationship. However, if a change in General Chemistry causes no change in subsequent courses, this would support a claim that the link between the courses is spurious. Thus, an investigation into longitudinal impact of a reform can provide information pertaining to the role of General Chemistry in the broader chemistry curriculum.

The Reform: Peer-Led Team Learning

The reform investigated in this study is Peer-Led Team Learning (PLTL), a nationally disseminated effort to incorporate active, socially mediated learning in the chemistry classroom. In PLTL students are engaged in problem-solving sessions as an integral part of the course.[17] Central to enacting this reform are peer leaders, undergraduate students who previously successfully completed the target course and return to lead small groups of students in the problem solving session. The peer leaders are trained and supervised by a faculty member who is involved in the target course. The problem solving sessions are designed by faculty to be challenging and promote group work.

The effectiveness of PLTL as an instructional technique in chemistry has been extensively studied and recently reviewed in the research literature.[18] The majority of past work found positive impacts on student academic performance in the target course, measured by student retention, performance on common exams or final course grades. Two studies have examined PLTL beyond one term. Wamser implemented PLTL as a voluntary course component for students across three terms of Organic Chemistry and found that persistence across the three terms was 57% for those students who opted for the workshop and 28% for those who opted not to.[19] Mitchell et al. investigated the impact of PLTL as a mandatory course component in General Chemistry I on student performance in General Chemistry II.[18] The study reported that the reform placed more students into General Chemistry II through a higher pass rate in General Chemistry I, however students who returned to traditional instruction in General

Chemistry II experienced lower pass rates in General Chemistry II. To date, no studies have examined the longitudinal impact of PLTL beyond the immediate follow-on course.

■ RESEARCH QUESTIONS

The overarching hypothesis to guide this work is that a successful PLTL reform targeting General Chemistry can have a longitudinal impact throughout a chemistry curriculum. Student enrollment and performance in subsequent courses are examined in this study, as each is relevant to the goal of promoting student retention. The following specific research questions guided this study:

1. To what extent does PLTL in General Chemistry impact enrollment in subsequent chemistry courses?
2. To what extent does the same reform impact academic performance in subsequent chemistry courses?

■ METHODS

Research Setting

The research setting is a large, primarily undergraduate institution in the southeastern United States. At the setting, only three majors required chemistry courses. The Chemistry major required the entire course sequence presented in Figure 1, where each arrow indicates a prerequisite relationship.

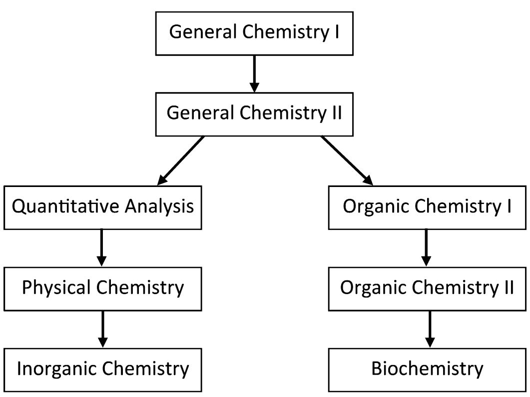


Figure 1. Required course sequence for bachelor's chemistry.

Biology required students complete the General Chemistry and Organic Chemistry sequences. Finally, Exercise and Health Science required only the General Chemistry sequence. The class sizes for General Chemistry and Organic Chemistry typically range from 55 to 75 students, whereas upper-level classes were smaller.

PLTL was implemented at the research setting beginning in Spring 2009 for General Chemistry I and Spring 2010 for General Chemistry II. The implementation occurred at the class level and the class meeting time students selected during enrollment determined whether students were in classes using PLTL. Classes were not advertised based on whether PLTL was present and it is probable that many students enrolling in General Chemistry I were not aware whether they were enrolling in a class that used PLTL. For General Chemistry II, it is believed that most students who were previously enrolled at the university were aware of which classes would use PLTL. Students self-selecting the PLTL (or traditional instruction) poses a threat to the analysis to the extent that student background is differentiated between the instructional techniques. To partially address this threat, SAT subscores were compared as discussed below. Students could still differ on other factors, such as student motivation or intended major.

PLTL was implemented at the setting as a mandatory course component with one-third of the class time set-aside for the problem-solving session. The classes not using PLTL relied on primarily lecture-based instruction owing to the class sizes. In the PLTL problem-solving sessions, students worked in assigned groups of four, with three to four groups of students assigned to a peer leader. During the problem-solving session, students were given a problem set of approximately ten questions designed by instructors to incorporate both conceptual and algorithmic type questions. Peer leaders enrolled in a training course that met weekly to prepare peer leaders to facilitate the sessions through techniques to encourage the students to actively engage the material and work together such as requesting students explain their reasoning or requesting a group reach a consensus. The training course was modeled as a peer-led session, placing peer leaders in groups to work the relevant problem set to ensure they were familiar with the content for the upcoming session. More detail on this implementation is available elsewhere in this *Journal*.[9] Past work at the research setting has shown that classes using this reform had increased pass rates in the targeted class compared to classes using traditional instruction through a decrease in student withdrawals, while maintaining student performance on a common final exam to show consistent rigor in the course.[9,18]

Participants

To investigate students' enrollment and performance in follow-on classes a longitudinal research study was employed that follows the same cohort of participants for an extended period of time. In this case there are two cohorts of interest. They are defined as Cohort 1, students who enrolled in General Chemistry I during the Fall 2010, Spring 2011, or Fall 2011 semesters at the research setting, and Cohort 2, students who enrolled in General Chemistry II during those same terms. For students who repeated General Chemistry I or II, only their last attempt was considered for eligibility in the cohort and whether they experienced the reform. During this time frame, the PLTL reform pedagogy was implemented in 10 of 26 classes of General Chemistry I offered and 8 of the 14 classes of General Chemistry II offered. There were 18 different instructors who led at least one of the 40 classes of General Chemistry I and II. Seven of the 18 instructors taught with the PLTL reform and three of those seven also taught the control classes during the same time frame.

Student enrollment from this cohort was followed through the set of subsequent chemistry courses from the Spring 2011 through the Summer 2013 semester. Student performance was measured by their final grades in the course from university records. IRB approval was obtained to conduct this study. For situations where students repeated one of the subsequent courses, only their first attempt at the course was recorded. The decision regarding repeated courses (considering only the last attempt in General Chemistry and the first attempt in subsequent courses) was made to focus on the impact of the General Chemistry experience that would be most proximal to the subsequent courses.

There were 1768 students in total who enrolled in either General Chemistry I or II during the semesters mentioned. Four of these students were auditing the class and were not considered for further analysis. Within Cohort 1, there were 1444 students enrolled in General Chemistry I (GC1). For Cohort 2, in General Chemistry II (GC2), 603 students enrolled during one of the three terms. Within the semesters of interest, 283 students had enrolled in both General Chemistry I and II representing some overlap between the two cohorts. Of the 283 students, 92 students enrolled in PLTL for both courses. The demographics and average SAT of Cohorts 1 and 2, delineated by method of instruction, are presented in Table 2.

Table 2. Demographic and Average SAT of Cohorts

	Cohort 1		Cohort 2	
Demographic Variables	PLTL	Traditional	PLTL	Traditional
Number of Students	555	889	318	285
Female Students, %	56.6	58.5	56.3	59.6
Black Students, %	17.8	13.5	14.5	15.1
Hispanic Students, %	2.5	4.4	2.2	2.5
Asian Students, %	6.7	6.1	7.5	7.4
White Students, %	65.2	68.5	67.6	68.1
Math SAT Score, Av	547	536	560	563
Verbal SAT Score, Av	553	545	537	556

■ RESULTS

Enrollment in subsequent classes among each cohort is presented in Table 3. Within each cohort, the sample sizes decreases considerably as the cohort progresses through the curriculum. This is particularly evident in Physical Chemistry and Inorganic Chemistry and is owing to the fewer numbers of students who intend to major in Chemistry (as the only major that requires those courses). Also, the time frame may not be sufficient for students to adequately progress into these courses. Extending the time frame would likely lead to only an incremental increase in the number of students in these classes particularly given that students would have to successfully complete Physical Chemistry prior to enrolling in Inorganic Chemistry. The decision was made to omit these two classes from further consideration; all of the remaining classes had at least 50 students enrolled from each cohort.

Student Enrollment in Subsequent Classes—Cohort 1

The analysis began with the 1444 General Chemistry I students in Cohort 1. The students with the PLTL reform were compared to those in the class with traditional instruction on Math and Verbal SAT scores and were found to be equivalent among General Chemistry I students using the two one-sided *t*-test procedure and ranges identified in prior work.[20] The impact of PLTL on subsequent course enrollment is determined by splitting the cohort by their instruction type in General Chemistry I. The subsequent enrollment of the split cohort is described in Table 4.

Table 4. Impact of PLTL in GC1 on Enrollment

Course Enrollment	PLTL in GC1, *N* (%)	Traditional GC1, *N* (%)	Differences, %
Initial GC1 Enrollment	555	889	
GC2	330 (59.5)	398 (44.8)	14.7
OC1	161 (29.0)	177 (19.9)	9.1
OC2	106 (19.1)	124 (13.9)	5.2
Biochem	29 (5.2)	28 (3.1)	2.1
Quant	38 (6.8)	32 (3.6)	3.2

The results in Table 4 show a substantial difference of nearly 15% in enrollment into GC2 that decreases steadily for subsequent classes. To determine if the differences in enrollment rates could be attributed to chance, a chi-square test was conducted for each subsequent class. The chi-square test evaluates the proportion of those with peer leading who enrolled in the subsequent class to the proportion of those without peer leading who enrolled. The results from the chi-square tests are presented in Table 5.

Table 5. Impact of PLTL in GC1 on Enrollment

Course Enrollment	χ^2	*p*-Value	Cohen's w
GC2	29.5	<0.001	0.1430
OC1	15.8	<0.001	0.1050
OC2	6.77	0.009	0.0685
Biochem	3.88	0.049	0.0518
Quant	7.81	0.005	0.0735

Because five related tests are conducted, the alpha value for determining statistical significance was set at a conservative 0.01. By this threshold, the PLTL in GC1 reform had a statistically significant impact on each course except Biochemistry. In addition to the chi-square test, the effect size of the difference was also calculated as Cohen's w. This metric provides an indication of the difference observed that is not dependent on sample size. Cohen prescribes a w value of 0.30 as a medium effect and a value of 0.10 as a small effect.[21] Thus, the impact of PLTL in GC1 on enrollment in GC2 and OC1 was small, and the effect approaches negligible in subsequent courses. This decrease in effect matches the trend in percent differences observed in Table 3.

Prior to making a claim that the reform had no impact on Biochemistry, statistical power needs to be considered. Failing to find a statistical difference between groups has two possibilities: there is no actual difference between the groups or there is a difference but the sample size is insufficient to show statistical significance (this latter possibility is known as

Table 3. Enrollment in Subsequent Courses

Course	Cohort 1: General Chemistry I	Cohort 2: General Chemistry II
General Chemistry II	728	
Organic Chemistry I (OC1)	338	330
Organic Chemistry II (OC2)	230	244
Biochemistry (Biochem)	57	84
Quantitative Analysis (Quant)	70	70
Physical Chemistry	11	25
Inorganic Chemistry	4	11

Type II error). Statistical power, the complement of Type II error, is a measure of the probability that given a difference is present it is determined statistically significant. In general, statistical power of 0.80 or greater is supportive of ruling out Type II error.[21] The statistical power to detect a small effect size for the chi-square tests in Table 4 is estimated to be 0.89, thus the likelihood of a Type II error in this case is low and the conclusion that the reform had no impact on Biochemistry enrollment is supported.

The observed differences in enrollment in each subsequent class can be largely explained by the impact of the reform on the pass rate in the target class. Among the 555 PLTL students in GC1, 442 students (79.6%) received a passing grade in the class. In the traditional instruction group, 564 students (63.4%) of the 889 enrolled students received a passing grade. If the number of students passing became the starting point in Table 4, the difference in enrollment rates are reduced considerably and range between 1.5% and 5.0% with the PLTL group still maintaining a higher enrollment rate throughout.

Student Enrollment in Subsequent Classes—Cohort 2

The analysis for students from the GC2 cohort proceeds in a similar fashion to GC1. The comparison of SAT scores showed the PLTL students and traditional students in GC2 had equivalent Math SAT scores, but the traditional students had statistically higher Verbal SAT scores.[20] Enrollment in subsequent courses, split by teaching style in GC2, is presented in Table 6.

Table 6. Impact of PLTL in GC2 on Enrollment

Course Enrollment	PLTL in GC2, N (%)	Traditional GC2, N (%)	Differences, %
Initial Enrollment	318	285	
OC1	192 (60.4)	138 (48.4)	12.0
OC2	144 (45.3)	100 (35.1)	10.2
Biochem	54 (17.0)	30 (10.5)	6.5
Quant	44 (13.8)	26 (9.2)	4.7

The results in this table are similar to the observed enrollment trends in Table 4, where the reform descriptively leads to a higher percent enrollment in the class immediately following the target class and the difference tapers down in subsequent courses. Chi-square tests were run for each subsequent class with the results presented in Table 7. Using the 0.01 threshold, the impact on Organic Chemistry I enrollment was significant with a small effect size. The impact on Organic Chemistry II enrollment was also a small effect size, but not statistically significant owing to the power of the chi-square test. With a sample size of 603 students from General Chemistry II, the statistical power for identifying a small effect size (w = 0.1) is 0.45 indicating a reasonable chance of a Type II error.[21]

Table 7. Impact of PLTL in GC2 on Enrollment

Course Enrollment	χ^2	p Value	Cohen's w
OC1	8.67	0.003	0.1200
OC2	6.48	0.011	0.1040
Biochem	5.22	0.022	0.0931
Quant	3.25	0.071	0.0735

Also similar to General Chemistry I, the difference in enrollment rates can be largely attributed to the higher pass rate of the reform in the target class. If the initial enrollments were replaced with the number of students who passed GC2, 271 for PLTL and 201 for traditional, the difference in percent rates would range between 2% and 5% in favor of PLTL.

Student Performance in Subsequent Classes—Cohort 1

Next, the impact of PLTL in General Chemistry I on student performance in subsequent courses was considered. To evaluate this, student grades were used as a measure of academic performance. At the research setting letter grades follow an A through F scale with no partial grades. Because letter grades represent ordinal-scale data, the nonparametric Mann–Whitney U test was used to compare the two groups. This test works by assigning each student a rank based on letter grade, where F would be the lowest rank and A the highest. Students withdrawing from the course are omitted from the analysis. The test then compares the average rank of each group and determines the significance value of the difference in average rank. The results from the Mann–Whitney test are in Table 8.

Table 8. Impact of PLTL in GC1 on Course Grades

Course	N	Av Rank PLTL in GC1	Av Rank Traditional GC1	Cliff's δ (p Value)	Power
GC1	1259	685	591	0.15 (<0.001)	0.990
GC2	638	321	318	0.07 (0.827)	0.780
OC1	315	153	162	0.06 (0.340)	0.420
OC2	210	105	106	0.02 (0.977)	0.260
Biochem	54	24	31	0.02 (0.054)	0.057
Quant	57	30	28	0.02 (0.573)	0.059

Comparing the average rank of each group in Table 8 provides a qualitative description of differences in the grade distributions between the groups. Although the difference in ranks of 94 in the GC1 target course is notable, the remaining classes have differences less than 10. As the absolute difference in rank is strongly dependent on sample sizes, effect size was also calculated which standardizes for sample size. Cliff's δ provides an indication of effect size by describing the percent of nonoverlap between the distributions.[22] In this metric for effect size, 0.15 corresponds to a small effect size and 0.33 to a medium effect size.[23] Thus, the reform had a small effect on the grade distribution in the target class of GC1, which steadily drops off in subsequent classes. Each p value in Table 8 represents the likelihood that differences observed could be attributed to chance as determined by the Mann–Whitney test. Similar to the interpretation of effect size, only the reform's impact on the target class was sufficient to rule out chance at the alpha level of 0.01.

Statistical power for this test was estimated using SAS 9.3. The effect size specified for the power analysis was the impact of the reform on the target course and the alpha level was 0.01. Interpreting statistical power for the test on GC2 can reasonably rule out Type II error and the conclusion that no difference exists between the groups can be supported. For the subsequent classes, there is the potential for Type II error where the groups are different but the sample size is insufficient to show the difference at the 0.01 alpha level cutoff. This information can be considered in the context of effect sizes, which indicates minimal evidence of any differences. Thus, for the classes after GC2, although the statistical test lacks statistical power to claim no differences, effect sizes provide little reason to believe a difference is present.

Student Performance in Subsequent Classes—Cohort 2

To examine the impact of the GC2 reform on student performance in subsequent courses, the Mann–Whitney test was conducted with the results shown in Table 9.

Table 9. Impact of PLTL in GC2 on Course Grades

Course	N	Av Rank PLTL in GC2	Av Rank Traditional GC2	Cliff's δ (p Value)	Power
GC2	565	306	256	0.25 (<0.001)	0.900
OC1	311	159	152	0.12 (0.508)	0.620
OC2	218	113	105	0.10 (0.342)	0.440
Biochem	81	41	41	0.06 (0.905)	0.140
Quant	61	30	33	0.06 (0.431)	0.098

The results are similar to the impact of the GC1 reform; however, the effect size in the target course is now in the small–medium range. The effect in subsequent courses trails off, with OC1 and OC2 just below the small effect threshold, and the remaining courses having minimal impact. Because the GC2 cohort has a smaller sample size, statistical power is less compared to the GC1 cohort. This opens the possibility for Type II error in any of the classes that failed to find a statistical difference. As before, the effect sizes provide minimal descriptive evidence that a difference exists, but it cannot be ruled out statistically.

■ DISCUSSIONS

To answer the first research question, the extent the reform impacts enrollment in subsequent chemistry courses, the reform has a statistically significant and small effect on enrollment in the class that directly succeeds the target class. As students progress through the curriculum the impact of the reform on course enrollment declines until it becomes attributable to chance. This phenomenon likely is not unique to the reform but rather is a function of student attrition between classes. For example, if a hypothetical reform had a 10% improvement in enrollment in one class, and 50% of both the reform and control students enrolled in the next class, there would only be a 5% improvement in the subsequent class. In the data presented, this is evident in examining the transition from GC2 and OC1 in Table 4, where the difference in percent metric falls 5.6%, from 14.7% to 9.1%. With those in Cohort 1 enrolled in PLTL, 48.8% of students enrolled in GC2 progressed into OC1. This can be calculated from Table 3 as 161 students enrolled in OC1 of 330 GC2 students. For those in Cohort 1 with traditional instruction, 44.5% made the same progression between the two courses. Despite the students with the reform having a higher progression rate, the differences in percent metric between the modes of instruction still fell by 5.6%.

In short, student attrition between courses will serve to mitigate the longitudinal impact of a reform that targets a gateway class. This relationship will hold true unless the reform has a positive impact on student progression throughout the curriculum. In this setting, the reform had a positive impact on the progression into the course immediately following the target, largely through the higher pass rate in the target course. In subsequent courses, the impact of the reform on progression through Biochemistry had considerable variability, though the PLTL reform had higher progression than the control in six of the seven progressions. Tables showing the progression rates between classes are available in the Supporting Information.

The role of student attrition as a mitigating factor on the impact of the reform is also a function of the majors available at the research setting. In particular, the transition from GC2 into OC1, where Exercise and Health Science majors would no longer enroll, found approximately 30% of students who pass GC2 not enrolling in OC1. Similarly, the transition from OC2 to Biochem, where Biology majors would no longer enroll, found approximately 60% of those passing OC2 not enrolling in Biochem. This attrition by design plays a sizable role in mitigating the impact of the reform in enrollment in subsequent classes.

The results observed set the current study apart from the previously reviewed studies that report longitudinal improvements of student retention between 18% and 41%.[14−16] Each of the reviewed studies investigated the impact on either an out of class program or an elective course. This allowed each program to work with a defined group of students with a single identified major such as Biology, Engineering, or Health Sciences. Thus, the attrition by design described above was not a factor for each of these programs. Instead, the results from this study speak directly toward the impact of enacting a reform within a targeted gateway class, which has a variety of student majors present. Improving pass rates within a targeted gateway class will not be sufficient to improve enrollment in courses considerably removed from the gateway, such as Biochemistry. Further, as enrollment in Biochemistry serves as a proxy for graduating with a chemistry degree at the setting, the results show that impacting the success in a gateway course had no effect on the percent graduating beyond what can be attributed to chance.

This result may call into question whether the PLTL reform is worth implementing. Such a question requires a broader consideration of costs and benefits. First, the impact of the reform on the target class clearly shows an increase in student performance in the target class and enrollment in the class that directly follows. The improvement in the target course matches results observed in numerous studies as shown by a recent meta-analysis.[24] The resulting improved academic standing for a substantial number of students is noteworthy. Additionally, the peer leaders gain experience in communicating science knowledge to students in the target class as well as the development of process skills such as time management and leadership through communicating expectations. The costs of the program to the institution involved allocating the teaching assignment of one class for a faculty member to train the peer leaders. Additionally, the peer leaders were paid a small stipend commensurate to the tuition paid to enroll in the training class. Weighing the costs to benefits at the institutional level will require an analysis of the resources available and the importance of the students' academic progress and development of upper-level students as peer leaders.

Returning to the policy recommendation to invoke reform teaching to improve the number of science graduates, the results indicate it is unlikely that targeting a gateway course alone will impact this metric at least in terms of the number of chemistry graduates. It may be possible that the reform promoted the number of graduates in the other majors that required General Chemistry, which would be ample ground for future work. Focusing again on chemistry graduates, because research has also shown that the PLTL reform improves

dx.doi.org/10.1021/ed500404q | J. Chem. Educ. 2014, 91, 2037−2044

student performance when implemented in Organic Chemistry and Biochemistry,[25] curricular-wide adoption of reform teaching has the potential to produce a very significant increase in the number of graduates. Future work would benefit by evaluating curricular-wide adoption with the metric of progression through the curriculum and ultimately number of science graduates.

For the second research question, the extent the reform impacts academic performance in subsequent courses, the impact of the reform on student performance mirrors the trend in enrollment with a significant small to small–medium effect in the target course that tapers off to a minimal effect in later courses that can be attributed to chance. In considering student grades, statistical power becomes a concern in the later courses owing to small sample sizes. Even with reduced statistical power, the minimal effect sizes provide no qualitative indications that a difference is present. Thus, no evidence was found that the mechanism by which the PLTL reform promotes success in the target class, either through skills or content learned, is transferred to other chemistry courses. Whether the relationship between student performance in General Chemistry and subsequent chemistry courses is spurious remains an open question. It is possible that another implementation of PLTL or a different reform targeting General Chemistry can lead to improved student performance in subsequent courses. In particular, both Process Oriented Guided Inquiry Learning and Problem-Based Learning are well-established reforms and have potential for long-term impacts given existing research.[26] Conversely, it may be that the content in the current General Chemistry curriculum does not lend itself to content or skills that substantially influence performance in other chemistry courses.[27] In that case, it may be necessary to alter the General Chemistry curriculum in order to observe an impact in subsequent courses.

■ LIMITATIONS

As mentioned in the discussion, the results are highly dependent on the structure of the curriculum and the majors available at the research setting. The results are primarily generalizable to those institutions with a similar curricular setup; however, the sequence of the General Chemistry through Organic Chemistry and into Biochemistry, is likely a common arrangement for many institutions. It is also expected that having a wider variety of majors in the target class would lead to a quicker amelioration of the observed benefits of the reform by the metrics described. This study also assumes that the students in the reform and traditional instruction had comparable distributions of intended majors. Finally, there may be concern about the time-length of the longitudinal work ultimately providing too small of a sample in later courses in the sequence. This concern is addressed by providing effect sizes that describe the impact independent of sample size.

■ CONCLUSIONS

In summary, the results from this work show the potential for a successful teaching reform to place more students into subsequent chemistry courses. The long-term impact of this result becomes mitigated through student attrition throughout the curriculum, both by design through intended major and through student performance in subsequent courses. That this trend was consistent when the reform was enacted in both General Chemistry I and General Chemistry II supports the reliability of these results. In terms of policy recommendations, no evidence was found that a reform that successfully targets an introductory, gateway course such as General Chemistry would lead to a notable increase in chemistry graduates. Instead, curricular wide adoption of reform teaching may be needed to achieve this goal. Further, no evidence was found that improving student performance in General Chemistry improved student performance in subsequent classes.

■ ASSOCIATED CONTENT

Supporting Information

Tables showing the progression rates described under the discussion section. This material is available via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: slewis@usf.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Partial support for this work was provided by the National Science Foundation's Course, Curriculum, and Laboratory Improvement (CCLI) program under DUE-0941976. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

■ REFERENCES

(1) President's Council of Advisors on Science and Technology. *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics*; The White House: Washington, DC, 2012.

(2) National Research Council, *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*, Singer, S. R., Nielsen, N. R., Schweingruber, H. A., Ed.; Committee on the Status, Contributions, and Future Directions of Discipline-Based Education Research: Washington, DC, 2012.

(3) Amaral, K. E.; Shank, J. D.; Shibley, I. A., Jr.; Shibley, L. R. Web-Enhanced General Chemistry Increases Student Completion Rates, Success and Satisfaction. *J. Chem. Educ.* **2013**, *90*, 296–302.

(4) Baez-Galib, R.; Colon-Cruz, H.; Rest, W.; Rubi, M. R. Chem-2-Chem: A One-to-One Supportive Learning Environment for Chemistry. *J. Chem. Educ.* **2005**, *82*, 1859–1863.

(5) Bowen, C. W. A Quantitative Literature Review of Cooperative Learning Effects on High School and College Chemistry Achievement. *J. Chem. Educ.* **2000**, *77*, 116–119.

(6) Chase, A.; Pakhira, D.; Stains, M. Implementing Process-Oriented, Guided-Inquiry Learning for the First Time: Adaptations and Short-Term Impacts on Students' Attitude and Performance. *J. Chem. Educ.* **2013**, *90*, 409–416.

(7) Dougherty, R. C.; Bowen, C. W.; Berger, T.; Rees, W.; Mellon, E. K.; Pulliam, E. Cooperative Learning and Enhanced Communication: Effects on Student Performance, Retention, and Attitudes in General Chemistry. *J. Chem. Educ.* **1995**, *72*, 793–797.

(8) Hockings, S. C.; DeAngelis, K. J.; Frey, R. F. Peer-Led Team Learning in General Chemistry: Implemenation and Evaluation. *J. Chem. Educ.* **2008**, *85*, 990–996.

(9) Lewis, S. E. Retention and Reform: An Evaluation of Peer-Led Team Learning. *J. Chem. Educ.* **2011**, *88*, 703–707.

(10) Lewis, S. E.; Lewis, J. E. Seeking Effectiveness and Equity in a Large College Chemistry Course: An HLM Investigation of Peer-Led Guided Inquiry. *J. Res. Sci. Teach.* **2008**, *45*, 794−811.
(11) Lyon, D. C.; Lagowski, J. J. Effectiveness of Facilitating Small-Group Learning in Large Lecture Classes. *J. Chem. Educ.* **2008**, *85*, 1571−1576.
(12) Oliver-Hoyo, M. T.; Allen, D.; Hunt, W. F.; Hutson, J.; Pitts, A. Effects of an Active Learning Environment: Teaching Innovations at a Research I Institutions. *J. Chem. Educ.* **2004**, *81*, 441−448.
(13) Pence, L. E.; Workman, H. J.; Haruta, M. E. A General Chemistry and Precalculus First-Year Interest Group (FIG): Effect on Retention, Skills, and Attitudes. *J. Chem. Educ.* **2005**, *82*, 65−69.
(14) Wischusen, S. M.; Wischusen, E. W. Biology Intensive Orientation for Students (BIOS): A Biology "Boot Camp". *J. Chem. Educ.* **2007**, *6*, 172−178.
(15) Sorby, S. A. Educational Research in Developing 3-D Spatial Skills for Engineering Students. *Int. J. Sci. Educ.* **2009**, *31*, 459−480.
(16) Hall, D. M.; Curtin-Soydan, A. J.; Canelas, D. A. The Science Advancement through Group Engagement Program: Leveling the Playing Field and Increasing Retention in Science. *J. Chem. Educ.* **2014**, *91*, 37−47.
(17) Tien, L. T.; Roth, V.; Kampmeier, J. A. A Course to Prepare Peer Leaders to Implement a Student-Assisted Learning Method. *J. Chem. Educ.* **2004**, *81*, 1313−1321.
(18) Mitchell, Y. D.; Ippolito, J.; Lewis, S. E. Evaluating Peer-Led Team Learning across the two semester General Chemistry sequence. *Chem. Educ. Res. Pract.* **2012**, *13*, 378−383.
(19) Wamser, C. C. Peer-Led Team Learning in Organic Chemistry: Effects on Student Performance, Success, and Persistence in the Course. *J. Chem. Educ.* **2006**, *10*, 1562−1566.
(20) Lewis, S. E.; Lewis, J. E. The Same or Not the Same: Equivalence as an Issue in Educational Research. *J. Chem. Educ.* **2005**, *82*, 1408−1412.
(21) Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale: MI, 1988.
(22) Cliff, N. Dominance Statistics: Ordinal Analyses to Answer Ordinal Questions. *Psychol. Bull.* **1993**, *114*, 494−509.
(23) Romano, J.; Kromrey, J. D.; Coraggio, J.; Skowronek, J.; Devine, L. Exploring methods for evaluating group differences on the NSSE and other surveys: Are the t-test and Cohen's d indices the most appropriate choices? *Proceedings of the Southern Association for Institutional Research*; Southern Association for Institutional Research: Arlington, VA **2006**.
(24) Freeman, S.; Eddy, S. L.; McDonough, M.; Smith, M. K.; Okoroafor, N.; Jordt, H.; Wenderoth, M. P. *Proc. Nat. Acad. Sc. U. S.* **2014**, 8410−8415.
(25) Platt, T.; Roth, V.; Kampmeier, J. A. Sustaining Change in upper level courses: peer-led workshops in organic chemistry and biochemistry. *Chem. Educ. Res. Pract.* **2008**, *9*, 144−148.
(26) Eberlein, T.; Kampmeier, J.; Minderhout, V.; Moog, R. S.; Platt, T.; Varma-Nelson, P.; White, H. B. Pedagogies of Engagement in Science; A Comparison of PBL, POGIL and PLTL. *Biochem. Mol. Biol. Educ.* **2008**, *36*, 262−273.
(27) Cooper, M. The Case for Reform of the Undergraduate General Chemistry Curriculum. *J. Chem. Educ.* **2010**, *87*, 231−232.