The dosing effect of Peer Assisted Learning in undergraduate organic chemistry

Corey Shanbrom*, Department of Mathematics and Statistics, California State University, Sacramento, California 95819, United States

Lauren Perry, Department of Mathematics and Statistics, California State University, Sacramento, California 95819, United States

*Email: corey.shanbrom@csus.edu

Abstract

Prior research has shown that supplemental peer support programs in chemistry enhance student success in the parent course. But the literature lacks studies on whether participation in multiple peer-led supplemental instruction courses is more effective than just one. Here we find a statistically significant dosing effect for Peer Assisted Learning (PAL) support in organic chemistry: students who took two sequential organic chemistry PAL courses outperformed those who only took one.

PAL is a peer-led supplemental instruction model where program enrollment is voluntary, but attendance is mandatory for enrolled students. Due to the optional nature of the program, we use propensity score weighting to mitigate academic and demographic variations between the two groups of students studied. Our dataset consisted of 77 Organic Chemistry II students who enrolled in the PAL support course between Fall 2020 and Fall 2024. Within this group, we compared those who also enrolled in the PAL support course for the prerequisite Organic Chemistry I class with those who did not. We found that students who took both PAL courses outperformed those who only took the second by 0.58 grade points (P=0.028, 95% CI 0.07-1.10) in their Organic Chemistry II course grade.

Keywords

Peer support, supplemental instruction, organic chemistry, propensity score analysis, cooperative learning, second-year undergraduate, upper-division undergraduate

Introduction

Organic chemistry is traditionally what Arendale[1] would describe as a "high-risk course," and it will surprise no readers that there is a rich literature focused on student success interventions for this content. Most relevant to our contribution is the body of research analyzing peer instruction: any of the various models of instruction using peers or near-peers interacting with students to support their success in undergraduate chemistry courses, including the organic chemistry sequence.[2-13] In addition to traditional support structures like tutoring centers, popular examples include supplemental instruction programs like Peer-Led Team Learning,[14] Learning Assistants,[15] and the University of Missouri-Kansas City model of Supplemental Instruction.[16] A large number of closely related models can be found in Arendale's extensive bibliography.[17]

Peer Assisted Learning at Sacramento State

California State University, Sacramento, also known as Sacramento State, is one of 23 campuses in the California State University (CSU) system. Sacramento State is a large, public, urban campus enrolling over 31,000 undergraduate students. It is a primarily undergraduate institution offering many Master's degrees and some doctorates but no PhDs. The student body mostly hails from the greater Sacramento region, and the institution reflects the diversity of the metropolitan area: Sacramento State is a Hispanic Serving Institution, Black Serving Institution, and Asian American Native American Pacific Islander Serving Institution. It also serves a large number of first-generation and non-traditional students.

The Peer Assisted Learning (PAL) program was established in 2011 with a National Science Foundation grant and ran its first sections in 2012 to support a beginning Chemical Calculations course. A combination of partial institutionalization and a series of large federal grants has helped the program expand to serving 17 courses in biology, chemistry, mathematics, and physics with low pass rates. In particular, the program currently serves introductory Chemical Calculations, General Chemistry I and II, and Organic Chemistry I and II.

The program is modeled on the Peer-Led Team Learning (PLTL) pedagogical structure[18] with adaptations to local needs and culture. Students in the supported parent course can opt to enroll in a 1-unit PAL section, graded Credit/No Credit, with the grade determined solely by participation and attendance (three or fewer absences required to earn credit). Students are recruited by an advertising campaign early in the semester (classroom announcements, flyers, faculty promotion) and carefully marketed to all students, not just those who may be struggling or members of certain groups. In general, PAL students reflect the composition of the parent course, but accurate assessment does require controlling for differences arising from the selection process, which is the main goal of propensity score analysis (see below).

The PAL section itself is run as a problem-solving workshop where students work in small groups (3-4 students) on supplementary worksheets written by course faculty. The workshop is led by an undergraduate instructional student assistant, called a "Peer Leader" in the PLTL literature but a "Facilitator" at Sacramento State, whose primary responsibility is ensuring that all students are working collaboratively and productively. While the Facilitator must have been successful in the course (B grade or higher) and regularly attends lecture to keep up with the content, they do not teach, tutor, or even confirm answers. When needed, they do ask scaffolding questions to lead students in the right direction, but the value of productive struggle and building confidence being “mostly sure” are core tenets of the model.

During the PAL section itself students do not work on homework assignments or study for exams, but Facilitators hold regular office hours and review sessions (open to all students, not just those enrolled in the workshop) to support these activities. In all contexts, Facilitators share best practices and study strategies as well as campus resources with their students, who often are still learning how to be a successful college student and navigate the university structure.

As Facilitators are asked to serve in many roles, such as peer mentor and academic coach as well as providing course content support, they undergo rigorous and continuous training. There is a mandatory 8-hour training before each semester, as well as optional additional trainings for new Facilitators and those holding leadership positions within the program. But since the Facilitator role demands constant growth and reflection, all Facilitators take a 2-unit upper-division letter-graded course, Honors Seminar in Peer Learning, which meets for two hours once a week. In this course Facilitators develop socio-emotional skills like leadership, professionalism, assertive communication, and cultural competency. Moreover, while success in the parent course is a necessary requirement for the position, Facilitators are hired primarily based on their demonstration of empathy and a desire to help their fellow students succeed, traits which are assessed in the interview process by program faculty as well as experienced current peer Facilitators. Interview questions are primarily open-ended hypothetical scenarios, for example: “Javier tells you he has to miss class every now and then to take care of his little sister. How do you respond?".

Further details on the pedagogical, logistic, and cultural structures of this program can be found in the works of Akhavan et al.,[19] Lundmark et al.,[20] and Shanbrom et al.[11].

Organic chemistry at Sacramento State

The organic chemistry sequence at Sacramento State is largely similar in curriculum and pedagogy to comparable institutions in the United States. With an emphasis on properties of molecules, Organic I builds the foundation for the applications in Organic II. In both courses students are assessed by a mix of exams, quizzes, and problem sets (both in-class and as homework). Organic II uses as its final exam the American Chemical Society national standardized exam; no other exams in the sequence are completely multiple choice. Attendance and participation are not assessed, but regular in-class quizzes and assignments keep attendance levels high.

Pedagogically, instructors practice active learning with interactive problems and lectures. A particular focus is placed on problem solving and analysis: every exam (except the Organic II final) requires students to problem solve and show their work step by step. Sample syllabi and exams for both Organic I and Organic II appear in our Supporting Information. The PAL worksheets, publicly available on the PAL website,[21] were written by course instructors and reinforce the problem-solving, analysis, and mechanism emphases. All 42 PAL worksheets for organic chemistry can also be found in our Supporting Information. We hope these materials provide the reader with curricular context in which to interpret our results.

However, Sacramento State's organic chemistry sequence does face certain challenges which may be, if not unique, exacerbated by local factors. As a regional Masters-granting university (Carnegie R2 classification), there are no discussion or recitation sections led by graduate student teaching assistants as might be found in a research institution. Moreover, class sizes are larger than many small private institutions, with both Organic I and II lectures capped at 75 students. As each course typically offers two lecture sections per semester, and with an additional 80 students taking these courses over summer, well over 300 non-unique students take these courses each year. In addition to the well-known challenges to student success in organic chemistry[1], Sacramento State students often work full- or part-time, or have considerable caregiving responsibilities.

Furthermore, the lab course (Chem 25) is not a part of either lecture course: Organic I is a prerequisite, while Organic II may be taken concurrently. That is, students in Organic I and II do no lab work. As the lab component is usually associated with higher grades – the DFW rate (percentage of students who did not pass the course with a C- or better) for the lab is consistently under 5%, compared to 20-40% for both lectures – disassociating the lab from the lecture effectively lowers the lecture grades compared to courses with embedded labs. All of these reasons make the organic chemistry sequence at Sacramento State even more challenging than is perhaps normal nationwide, and this situation was indeed the main impetus for implementing PAL in these courses in the first place. Large amounts of extra practice with course material is beneficial in many classes, and PAL providing dedicated, consistent, structured time to do so may be particularly helpful for students splitting their time between academic, work, and familial responsibilities. In general, PAL serves the science and math courses with the highest DFW rates at Sacramento State. In addition to Organic I and II, this includes Chemical Calculations and General Chemistry I and II.

Background

Evaluating peer support structures

Peer-led supplemental instruction has been found effective in promoting student success in the supported organic chemistry course.[3, 6-7, 9-13, 17] Guyot et al.[6] attribute at least some of this success to increased student motivation. Our results are particularly interesting in light of Rath et al.,[9] who found that supplemental instruction was more effective in Organic Chemistry I than Organic Chemistry II at an institution very similar to the one we studied.

An often-overlooked challenge when evaluating the efficacy of such programs is their voluntary nature: students choose whether or not to participate. While some peer-led supplemental instruction programs are compulsory, the voluntary programs suffer from the complication that the students who choose to participate may differ in important ways from those who do not. Thus, one cannot simply compare students who opt in with those who opt out. Other authors have recognized that voluntary PLTL programs may have stronger students opt in;[4-5] in such situations it may be the case that students who participate would have been more successful with or without the intervention.

There are a variety of statistical methods for dealing with scenarios where the control group and the treatment group are not equally likely to receive the treatment. Here, we employ propensity score weighting, which attempts to control for the likelihood of receiving the treatment. In our context, we assign to each student an inverse probability weight based on their *propensity score*, the likelihood of their opting into a particular peer-led supplemental instruction program (the treatment). Analyses are then completed using weighted least squares techniques. This approach mitigates the bias caused by students self-selecting into the treatment group. Note that the term *propensity score analysis* refers to a family of related techniques, all of which use *propensity score modeling* to assign subjects a propensity score. Various methods can then be used to account for this propensity score, including *propensity score matching* (used in Shanbrom et al.[11]) and *propensity score weighting* (used here).

More information on propensity score analysis can be found in works by Austin,[22] Bowman et al.,[3] Brookhart et al.,[23] Leite,[24] Shadish and Steiner,[25] and Zhang et al.[26] Propensity score analyses have been used to evaluate PLTL programs at the college level[11] and at the high school level,[27] as well as non-peer-led academic interventions.[28-30]

Here we use propensity score weighting to compare organic chemistry students at Sacramento State, focusing on the effect of the PAL program on two courses: lower-division Organic Chemistry I (Chem 24) and upper-division Organic Chemistry II (Chem 124).

The analysis of Shanbrom et al.[11] demonstrated the effectiveness of the PAL program in 11 STEM courses with low pass rates. This study included Organic I but not Organic II. Not only did we see that PAL was effective in Organic I, in fact the effect was stronger in this course than in any of the 10 other courses studied (see Table 2 of Shanbrom et al.[11]). Students opting into the Organic I PAL course earned a 2.10 average course GPA in Organic I, compared with an average course GPA of 1.39 for those who did not take the PAL support course (on a traditional 4-point scale). This is a massive grade increase of .71 grade points, from an average D+ to an average C. The propensity score matching controls for academic and demographic variations in the two populations, suggesting that the increase is indeed due to PAL participation; the codebase and more data is posted publicly online.[31]

Dosing effects

Since PAL was clearly very effective in supporting the parent course of Organic I, in the present contribution we sought to determine whether there was a dosing effect: whether additional

PAL exposure led to increased benefits. Explicitly, our research question was: *Among Organic Chemistry II students taking the PAL support course, do those who also took the Organic Chemistry I PAL support course outperform those who didn't?*

This research question effectively defines what we mean by the phrase "dosing effect" in this paper. We are asking whether two doses of the PAL treatment are more effective than a single dose.

Mitchell et al.[8] examine a similar dosing effect for a PLTL program in the general chemistry sequence with pass rates as the outcome of interest. See our Discussion below for a comparison of their results with ours. The meta-analytic review of Rohrbeck et al.[32] considers the "total dosage" of Peer Assisted Learning received, but their study concerns elementary school students which is obviously a dramatically different population than university organic chemistry students, and they are also interested in a different set of outcomes. They ultimately found no dosing effect in their work, and they describe mixed results in the older literature.[33-35] Three of the four authors of the review of Rohrbeck et al.[32] also published a similar meta-analytic review[36] but with a focus on psychological outcomes, and again found no dosing effect. However, all of these authors measure dosage in terms of the amount of time students received the treatment, which could vary in the programs studied. This is a vastly different measure from ours, stemming from a different programmatic structure (PAL students at Sacramento State all receive the same number of treatment hours).

Other loosely related research includes works by Wonder-McDowell,[37] who considered supplemental instruction as a "second dose" of instruction (after the primary direct instruction) for improving second grade reading levels, and Ogden et al.,[38] who studied long-term effects of Supplemental Instruction in college-level political science students.

Methods and Results

A propensity score analysis was conducted in R software[39] using packages WeightIt and Cobalt[40, 41] to assess the dosing effect of PAL courses in organic chemistry. This project was approved by the Research Integrity and Compliance Officer in the Offices of Research, Innovation, and Economic Development at Sacramento State under the Human Subjects Research Institutional Review Board 15-16-143.

As discussed in Shanbrom et al.,[11] propensity score adjustment is necessary for these data since they are observational, and because students self-select into PAL courses. A propensity score model uses confounders – variables which are related to both PAL enrollment and course grade – to model the probability of enrolling in PAL. These probabilities, or propensity scores, can then be used to mitigate self-selection bias. To avoid losing any information from these data, we performed a propensity score-weighted analysis.

We performed a propensity score weighted analysis using data from the Sacramento State institutional database for every student enrolled in NSM 12S, the PAL support course for Organic II, who had also taken Organic I at Sacramento State. The institutional database contains 95 variables on a variety of academic and demographic characteristics. The data on Organic I students spans the terms from Fall 2017 through Fall 2024, and on Organic II the terms from Spring 2020 through Fall 2024. For this analysis, we assume the impact of COVID-19 was similar across both groups. During this time, 719 students took both Organic I and Organic II. The DFW rates during this period were 26% in Organic I (*n*=1406) and 21% in Organic II (*n*=966); note that these *n*-values are larger than the 719 students who took both classes at Sacramento State as many students either took Organic I at a community college or did not progress into Organic II.

To create propensity score models, this set of variables was reduced in the following ways. First, missingness was examined and any variables with a large proportion missing were removed. Next, variables that were redundant or irrelevant to the outcome were removed (for example, race-specific indicators were removed in favor of a single variable describing ethnicity). Note that any term-specific variables correspond to the term in which the student took Organic I, since this is the term of interest for the propensity score model.

Significant subsetting was also necessary to examine a dosing effect. First, all students needed to have taken both courses at Sacramento State, so any students who took the lower-division organic chemistry course at another institution were removed from the data. Second, we limited our analysis to examining student grades for the first time a student took a course, so any course repeats were removed. After the variable selection was performed (and so variables with high missingness removed), students who were missing data from any of the remaining variables were also removed from the dataset. Finally, we were only interested in those students who attended and participated in the Organic II PAL. Because of the participation-based nature of PAL courses, we defined PAL participation in Organic II PAL as students who had both enrolled in and passed that PAL course. As attendance and participation are the only criteria for passing the PAL course, with no more than three absences allowed and Facilitators ensuring that the whiteboard marker is consistently rotating amongst group members, we assume that all the students in this group received roughly equal treatment.

After the variable selection and subsetting, all remaining variables were entered into a propensity score model using logistic regression. This model was refined first by examining variables causing complete separation in the logistic regression, meaning those variables had categories which perfectly predicted the outcome, making the model unstable. These variables were modified to collapse sparse categories into “other” in order to ensure enough data in each category. Further, input variables which are highly correlated with each other can cause problems in examining statistical significance in the model, so the model was then refined by sequentially removing variables with perfect multicollinearity or high variance inflation factors, stopping when all variance inflation factors were less than 5 – a commonly accepted threshold for acceptable levels of correlation. This resulted in the removal of several other variables. The final propensity score model includes the 15 covariates displayed in Table 1, where “term” always refers to the semester in which the student took Organic I.

We also analyzed which of these 15 covariates were most predictive of course grade in Organic II, but found that none were strongly correlated. Each predictor was used individually to construct a linear model predicting the Organic II grade, but the only two with an R^2 value over

0.1 were the student's grade in Organic I (R^2 =0.10, P<0.001) and the student's cumulative GPA (R^2=0.13, P<0.001), neither of which are particularly surprising. However, standard practice in propensity score methodology recommends including all variables that are associated with the outcome, regardless of their statistical significance for treatment assignment, to reduce the risk of omitted variable bias[22,23]. Consistent with recommendations in the literature, we used a liberal threshold of P<0.20 when assessing outcome associations to avoid excluding potential confounders, even if their individual predictive power was weak[22].

Table 1. Descriptions of the 15 covariates used in final propensity score model

Name	Description
grade.num.1	Grade in Organic I
coh	Indicates transfer or native freshman
acad.stndng.stat.desc	Academic standing
gender	Gender
eth.ipeds	Ethnicity
foreign.flg	International student status
father.ed	Father's level of education
mother.ed	Mother's level of education
pell.term.flg	Was student Pell eligible when entering Sacramento State
term.age	Student's age in term
school.zip.median.income	Median income in student's last school ZIP code
prop.pssd	Proportion of non-remedial units passed in the current term
tot.cumulative.start	Total units taken at start of term
cum.gpa.start	Cumulative GPA at start of term
data.chem.hs.gpa	High school chemistry GPA

Each propensity score represents the probability of a student having participated in NSM 12N, the PAL support course for Organic I. Using the propensity score model, each student in the dataset was assigned a weight using inverse probability weighting. See Austin and Stuart[42] for more details on inverse probability weighting in propensity score analysis. These weights are then used in subsequent analyses to calculate quantities like, for example, a weighted mean grade for each group or a weighted t-test for difference of means. After weighting, absolute mean differences were examined to confirm covariate balance. The criteria for covariate balance range from a strong criterion of standardized differences less than 0.1 to a more lenient criterion[24] of less than 0.25. In this analysis, all absolute mean differences achieved the strong criterion, with the exception of prop.pssd, which had a weighted absolute mean difference of 0.107. See Figure 1 for a summary of absolute mean differences for this model; note the first row shows the absolute mean differences in propensity scores decreasing from 0.65 before weighting to 0.02 after weighting.

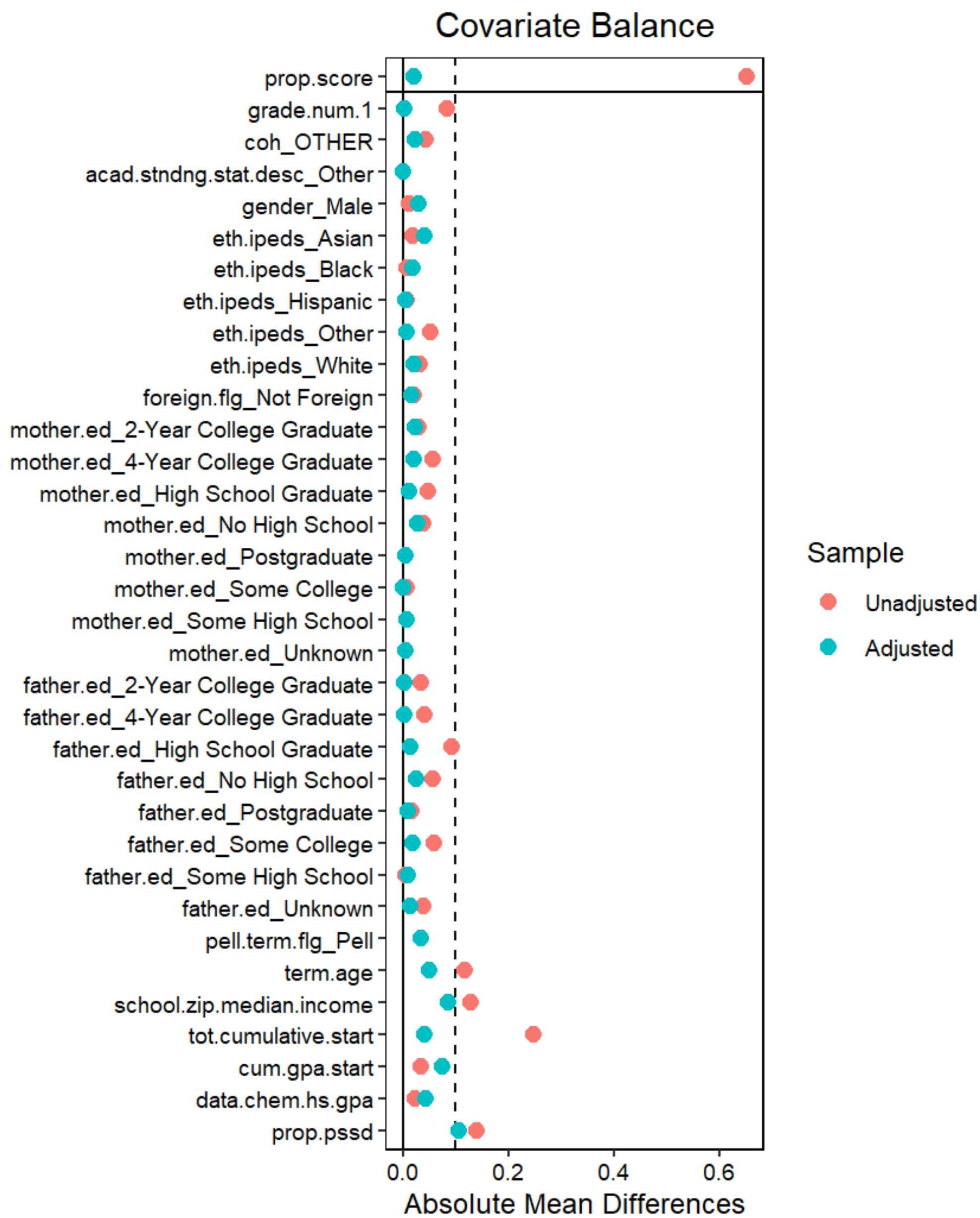

Figure 1. Covariate balance before and after propensity score weighting. The strong criterion of 0.1 is shown as a dashed line. See Table 1 for descriptions of covariates.

Finally, we examined the weighted means for each group and performed a weighted t-test (using R package weights[43]) to compare the mean grade for students who took PAL courses for both Organic I and Organic II to the mean grade for students who took the Organic II PAL only. Our main results appear in Table 2. We found that students who took both PAL courses outperformed those who took the Organic II PAL only. Taking the second PAL class resulted in an average course GPA boost of 0.58 grade points (*P*=0.028, 95% CI 0.07-1.10), a shift from a C+ to a B. The R Markdown files with the code used in our analysis appear in the Supporting Information and are posted publicly online.[44]

Table 2: Course GPAs for select groups of students; our main result is the last column, which shows a 0.58 grade point increase for students taking a second organic chemistry PAL course

Group	n	Unweighted GPA in Organic I	Unweighted GPA in Organic II	PS-Weighted GPA in Organic II
Organic II PAL only	31	2.45	2.41	2.29
PAL for Organic I and II	46	2.70	2.82	2.87

A sensitivity analysis was conducted using the EValue package in R.[45] This package translates the difference in means result into a risk ratio for interpretability of results and produces e-values, which provide a measure of how strong an unmeasured confounder would have to be to explain away the observed treatment effect. The translated risk ratio for our result is 1.63, with an e-value of 2.64. This means that, to fully explain away the treatment effect, an unmeasured confounder would need to be associated with both the treatment and the outcome by a risk ratio of at least 2.64 each, after adjusting for all measured covariates. This is a moderately strong effect.[45] However, the lower bound of the confidence interval is more sensitive, with a risk ratio of 1.07 and an e-value of 1.33. This suggests that the lower bound of the confidence interval is moderately sensitive to unmeasured covariates.[45]

Discussion

We conclude that the Peer Assisted Learning program at Sacramento State demonstrated a significant dosing effect in the organic chemistry sequence. However, there are substantial limitations to this study, and our results should be taken as an invitation to further research.

In particular, this was a relatively small sample size (*n*=77, tiny compared to our sample size of 10,333 in Shanbrom et al.[11]). Our results are statistically significant largely due to the strength of the GPA increase. While here our outcome focused on only one upper-division class which was taken by relatively few students, the large amount of subsetting necessary to examine the student groups of interest also led to a loss of approximately half of our initial data. For example, many Organic II students took Organic I at a community college and were subsequently removed from our dataset. Furthermore, we analyzed one specific program at one specific institution and were limited by the fact that PAL support for Organic II was first offered in Fall 2020 (other PAL courses at Sacramento State date back to 2012). This small sample size prohibited the examination of how GPA effects varied across demographic subgroups, but we do account for some of these variables in our model.

A different limitation is that our quantitative results fail to shed light on the question of how exactly taking the Organic I PAL helped the Organic II students. For example, we could be seeing more of a persistence effect than a dosing effect; perhaps the impact of the Organic I PAL persisting after that treatment had ended is more responsible for the observed GPA increase than the two treatments in combination. We believe a more important question here is whether the academic gains were due to specific content knowledge (a stronger foundation in Organic Chemistry I material, more time spent solving organic chemistry problems) or due to enhanced socio-emotional-psychological factors (confidence, sense of belonging, study skills, etc.). We suspect the answer is a combination of both.

An additional important limitation is the potential existence of unknown confounders: covariates which impacted the outcome but were not accounted for in our model since they do not appear in our database. The set of possible unknown confounders includes numerous variables which may be impossible to quantify reliably, or those which may fluctuate from semester to semester, or even within a single semester. Unknown confounders could be personal in nature: students with higher levels of motivation may be more likely to opt into PAL. Or they could be essentially logistic: students who have significant work or family or caregiving responsibilities may simply not have enough time for a PAL course. Other could be academic: some students, especially first-year or first-generation, may be less adept at navigating and utilizing campus resources like PAL. Moreover, not all students in this study took organic chemistry with the same instructor, and some instructors are more likely to promote PAL than others (and of course instructors' grading scales and styles can vary considerably). In all of these diverse hypothetical scenarios it is quite plausible that the observed GPA increase is not caused by PAL enrollment, but rather some other covariate which is correlated to PAL enrollment. Thus, the propensity score analysis here can only (imperfectly) account for those variables which are quantifiable, known, and present in the institutional dataset.

Therefore, the propensity score analysis certainly does not control for every difference between the two populations. Indeed, Table 2 suggests that the students opting into the Organic I PAL may be stronger in organic chemistry than those who did not, as their GPA in Organic I was already higher (2.70 versus 2.45). However, this gap widened in Organic II (2.82 versus 2.41), and widened even further after propensity score weighting (2.87 versus 2.29), which provides further evidence that the Organic I PAL participation did indeed positively affect performance in Organic II.

It is interesting to compare these results with those of Mitchell et al.[8]; see their Table 3 and surrounding discussion in particular. Like us, they consider a PLTL program at a large primarily undergraduate institution and examine combinations of PLTL dosage in a two-semester chemistry sequence. Unlike us, they measure pass rates as the outcome of interest and focus on the general chemistry sequence. Interestingly, they did not find a dosing effect in the same way that we did: students who had PLTL in both courses actually passed the second sequential course at a lower rate (79.3%) than students who had PLTL in only the second course (81.3%). They reasonably conclude that "that the improved success observed with PLTL in GC1 did not lead to transferable skills that students could then employ in GC2, and rather the effectiveness of PLTL is concentrated on the course in which it is enacted." This is starkly different from our conclusion.

However, there are several very plausible explanations for this contrast, beyond the obvious differences in measured outcome (pass rate versus GPA) and course content (general chemistry versus organic chemistry). While the PAL program in our study is based on the PLTL pedagogical model, in practice it looks very different from the PLTL program studied by Mitchell, et al.[8] In particular, their PLTL program is not voluntary for students, but rather for instructors. That is, some instructors opted to use PLTL, and their students were required to participate as part of the grade in the parent course (as opposed to a separate course enrollment and grade). While this avoids many of the self-selection problems which make evaluating the student-voluntary PAL program challenging (and thus the need for propensity score analysis), it introduces different challenges, and the authors acknowledge that "No effort was made on behalf of the research study to standardize the method of instruction, the in-class exams or assignments given, in order to model and evaluate the natural implementation of the PLTL reform." Thus, we consider our results to be complementary to those of Mitchell, et al.[8] rather than contradictory.

In conclusion, this limited study implies the possibility of dosing effects in other courses and disciplines, in other supplemental instruction programs, and at other institutions. We consider our results to be proof of concept and evidence of promising directions for further research.

Supporting information

R Markdown code for all statistical analysis (PDF). Sample syllabi, sample exams, and all PAL worksheets for both Organic I and Organic II (PDF).

Acknowledgements

We are grateful to Matthew Krauel, Jennifer Lundmark, Michelle Norris, and Vincent Pigno for their support and advice. We also thank the reviewers for suggestions which substantially improved the paper. We are especially appreciative of Cynthia Kellen-Yuen, who shared both course materials and her invaluable experience teaching organic chemistry at Sacramento State. This work was supported by Sacramento State's Assigned Time for Exceptional Levels of Service to Students Award.

Notes

The authors declare no competing financial interest.

References

1. Arendale, D. R. (1994). Understanding the supplemental instruction model. https://doi.org/10.1002/tl.37219946004
2. Barth, O. (2021). The effect of supplemental instruction on educational accomplishments and behaviors of organic chemistry scholars. *Water and Environmental Sustainability*, *1*(1), 30-36. https://doi.org/10.52293/wes.1.1.3036

3. Bowman, N. A., Preschel, S., & Martinez, D. (2023). Does supplemental instruction improve grades and retention? A propensity score analysis approach. *The Journal of experimental education*, *91*(2), 205-229. https://doi.org/10.1080/00220973.2021.1891010
4. Chan, J. Y. K. & Bauer, C. F. (2015). Effect of Peer-Led Team Learning (PLTL) on student achievement, attitude, and self-concept in college general chemistry in randomized and quasi experimental designs. *Journal of Research in Science Teaching*, 52(3), 319-346. https://doi.org/10.1002/tea.21197
5. Frey, R., Fink, A., Cahill, M. J., McDaniel, M. A., & Solomon, E.D. (2018). Peer-led team learning in general chemistry I: Interactions with identity, academic preparation, and a course-based intervention. *Journal of Chemical Education*, *95*(12), 2103-2113. https://doi.org/10.1021/acs.jchemed.8b00375
6. Guyot, M., Hsu, S., Nizomov, J., Antonenko, P., & Habenicht, S. (2024). Examining How Supplemental Instruction Impacts Students' Motivation in Organic Chemistry. *IJCER (International Journal of Chemistry Education Research)*. https://doi.org/10.20885/ijcer.vol8.iss1.art2
7. Lundeberg, M. A. (1990). Supplemental instruction in chemistry. *Journal of Research in Science Teaching*, *27*(2), 145-155. https://doi.org/10.1002/tea.3660270206
8. Mitchell, Y. D., Ippolito, J., & Lewis, S. E. (2012). Evaluating peer-led team learning across the two semester general chemistry sequence. *Chemistry Education Research and Practice*, *13*(3), 378-383. https://doi.org/10.1039/c2rp20028g
9. Rath, K. A., Peterfreund, A., Bayliss, F., Runquist, E., & Simonis, U. (2012). Impact of supplemental instruction in entry-level chemistry courses at a midsized public university. *Journal of Chemical Education*, *89*(4), 449-455. https://doi.org/10.1021/ed100337a
10. Salame, I. I. (2021). The impact of supplemental instruction on the learning achievements and attitudes of organic chemistry students. *Interdisciplinary Journal of Environmental and Science Education*, 17(2), e2232. https://doi.org/10.21601/ijese/9330
11. Shanbrom, C., Norris, M., Esgana, C., Krauel, M., Pigno, V., & Lundmark, J. (2023). Assessing student success in a Peer Assisted Learning program using propensity score matching. *Journal of College Science Teaching, 52*(7), 129–136. https://doi.org/10.1080/0047231X.2023.12315888
12. Tien, L. T., Roth, V., & Kampmeier, J. A. (2002). Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *39*(7), 606-632. https://doi.org/10.1002/tea.10038
13. Wilson, S. B., & Varma-Nelson, P. (2016). Small groups, significant impact: a review of peer-led team learning research with implications for STEM education researchers and faculty. *Journal of Chemical Education*, 93(10), 1686–1702. https://doi.org/10.1021/acs.jchemed.5b00862
14. Peer-Led Team Learning International Society. https://pltlis.org/
15. Learning Assistant Alliance. https://learningassistantalliance.org/
16. The International Center for Supplemental Instruction. https://info.umkc.edu/si/

17. Arendale, D. R. (editor). (2022). Postsecondary peer cooperative learning programs: Annotated bibliography. [Unpublished manuscript]. Department of Curriculum and Instruction, University of Minnesota. https://z.umn.edu/peerbib
18. Gosser, D. K., Cracolice, M. S., Kampmeier, V. R., Strozak, V. S., & Varma-Nelson, P. (2001). *Peer-led Team Learning: A Guidebook*. Prentice Hall.
19. Akhavan, N., Davison, J., Taylor, E., Hill, J., Mosely, V., & Salimo, K. (2024) The Community and Structure of Sacramento State's Peer Assisted Learning Program. *Advances in Peer-Led Learning,* 4, 19-27. https://doi.org/10.54935/apll2024-01-03-19
20. Lundmark, J., Paradis, J., Kapp, M., Lowe, E., & Tashiro, L. (2017). Development and impact of a training program for undergraduate facilitators of peer-assisted learning. *Journal of College Science Teaching*, *46*(6), 50-54. https://doi.org/10.2505/4/jcst17_046_06_50`
21. Peer Assisted Learning program (PAL). https://www.csus.edu/pal
22. Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786
23. Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, *163*(12), 1149–1156. https://doi.org/10.1093/aje/kwj149
24. Leite, W. (2016). *Practical Propensity Score Methods Using R*. Sage Publications.
25. Shadish, W. R., & Steiner, P. M. (2010). A primer on propensity score analysis. *Newborn and Infant Nursing Reviews*, *10*(1), 19-26. https://doi.org/10.1053/j.nainr.2009.12.010
26. Zhang, Z., Kim, H. J., Lonjon, G., & Zhu, Y. (2019). Balance diagnostics after propensity score matching. *Annals of Translational Medicine*, *7*(1). https://doi.org/10.21037/atm.2018.12.10
27. Thomas, A. S., Bonner, S. M., Everson, H. T., & Somers, J. (2015). Leveraging the power of peer-led learning: Investigating effects on STEM performance in urban high schools. *Educational Research and Evaluation*, 21(7-8), 537-557. https://doi.org/10.1080/13803611.2016.1158657
28. Hodara, M. (2013). *Improving students' college math readiness: A review of the evidence on postsecondary interventions and reforms*. Community College Research Center, Teachers College, Columbia University. https://doi.org/10.7916/D8M32SS7
29. Rose, S. J. (2013). The effectiveness of pre-course and concurrent course interventions on at-risk college physics students' mechanics performance. (Ph.D. dissertation), University of Illinois at Urbana-Champaign, Urbana-Champaign, IL.
30. Windsor, A., Bargagliotti, A., Best, R., Franceschetti, D., Haddock, J., Ivey, S., & Russomanno, D. (2015). Increasing retention in STEM: Results from a STEM talent expansion program at the University of Memphis. *Journal of STEM Education*, 16(2), 11-19. https://www.jstem.org/jstem/index.php/JSTEM/article/view/1886/1660
31. Shanbrom, C. (2021). PAL Effectiveness Analysis for Chem 24. https://webpages.csus.edu/Corey.Shanbrom/PALdata/chem24-PAL-Analysis.html
32. Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic

review. *Journal of educational Psychology*, *95*(2), 240. https://doi.org/10.1037/0022-0663.95.2.240

33. Cohen, P. A., Kulik, J. A., & Kulik, C. L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*(2), 237-248. https://doi.org/10.3102/00028312019002237
34. Cook, S. B., Scruggs, T. E., Mastropieri, M. A., & Casto, G. C. (1985). Handicapped students as tutors. *The Journal of Special Education*, *19*(4), 483-492. https://doi.org/10.1177/002246698501900410
35. Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin*, 89, 47–62. https://doi.org/10.1037/0033-2909.89.1.47
36. Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, *98*(4), 732. https://doi.org/10.1037/0022-0663.98.4.732
37. Wonder-McDowell, C. (2008). *The effects of aligning supplemental and core reading instruction on second-grade students' reading achievement*. Utah State University.
38. Ogden, P., Thompson, D., Russell, A., & Simons, C. (2003). Supplemental instruction: Short-and long-term impact. *Journal of Developmental Education*, *26*(3), 2.
39. R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
40. Greifer, N. (2025). *WeightIt: Weighting for Covariate Balance in Observational Studies*. R package version 1.4.0, https://CRAN.R-project.org/package=WeightIt
41. Greifer, N. (2024). *cobalt: Covariate Balance Tables and Plots*. R package version 4.5.5, https://CRAN.R-project.org/package=cobalt
42. Austin, P. C., and Stuart E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine,* 34(28), 3661–3679. https://doi.org/10.1002/sim.6607
43. Pasek, J., Tahk, A., Culter, F., & Schwemmle, M. (2021). *weights: Weighting and Weighted Statistics*. R package version 1.0.4, https://CRAN.R-project.org/package=weights
44. Shanbrom, C. (2025). Organic Chemistry PAL Dosing Effects. https://webpages.csus.edu/Corey.Shanbrom/PALdata/OChemAnalysis.pdf
45. VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: introducing the E-value. *Annals of internal medicine*, *167*(4), 268-274. https://doi.org/10.7326/m16-2607