Advances in Engineering Education

2025: VOLUME 13 ISSUE 1

# The Impact of Peer-led Study Groups on Student Achievement in a Gateway Engineering Thermodynamics Course

RYAN J. MILCAREK
Arizona State University

SAMANTHA R. BRUNHAVER
Arizona State University

THIEN TA
Quality Evaluation Designs

CODY D. JENKINS
Arizona State University

GARY LICHTENSTEIN
Quality Evaluation Designs

KARL A. SMITH
University of Minnesota and Purdue University

## ABSTRACT

Certain introductory engineering courses have been termed gateway courses, due to high failure rates resulting in many students leaving engineering altogether. In one such course, thermodynamics, instructors intentionally formed peer-led study groups (PLSGs) with an undergraduate teaching aide (UGTA) playing a supportive facilitator role instead of providing direct instruction. PLSG students met one hour per week for eight weeks, over and above the standard teaching assistant-led recitation. The intervention resulted in statistically significant improvements in students' course grades, pass rates, and graduation rates compared to students who participated in the standard recitation only. PLSG earned statistically almost one full letter grade higher than students in the no-treatment group; they were also statistically more likely to have passed the course and to have graduated with their degree approximately one year after taking it. PLSG students were also compared to students in a control group who participated in an additional hour of the standard teaching assistant-led recitation each week. Students in the PLSGs performed at similar levels to students in the control despite receiving far less instruction and guidance from the teaching aide. Notably, no differences were found in the effectiveness of the intervention for transfer students, women, and racial/ethnic minoritized students, indicating that the intervention provides at least the same educational advantages to underrepresented groups as to non-underrepresented groups. These results support the implementation of peer-led student groups in thermodynamics and other upper-level gateway engineering courses. Preparation and training videos are available for those interested in implementing this intervention in their own courses: https://lth.engineering.asu.edu/reference-guide/peer-led-study-groups/.

**Key words:** Cooperative learning, Peer instruction, Study groups

## INTRODUCTION

Middle-year engineering gateway courses covering engineering fundamentals are well-known obstacles in engineering programs (Lord and Chen 2014). These courses–which include statics, dynamics, mechanics of materials, thermodynamics, material science, and electrical circuits–tend to have high student enrollment and rely on lectures as the primary mode of instruction (MacGregor et al. 2000; Stains et al. 2018). The curriculum shifts from focusing on broad technical and professional skill development to learning challenging disciplinary-specific content; assessments become more individualized, usually with a heavy emphasis on problem sets; and one-on-one interaction with peers and instructors typically decreases as student-to-instructor ratios increase (Lord and Chen 2014). Consequently, whereas first-year courses are traditionally designed to build student identity, sense of belonging, and retention, second and third-year courses contribute to student frustration, lack of motivation, and attrition from engineering programs, all of which disproportionately affect historically marginalized students in engineering (Lichtenstein et al. 2014).

Class-based interventions to counteract these issues have included the use of concept inventories, concept maps, model-eliciting activities, design contests, case-based instruction, and rapid feedback assessment methods (e.g., clickers, flashcards) (Lord and Chen 2014), all of which center on improving student curricular engagement and conceptual understanding. Fewer interventions have addressed the perceived loss of social support and social isolation experienced by many students during the middle years of their undergraduate programs (Lord and Chen 2014).

This paper explores the impact of peer and near-peer support in a gateway engineering thermodynamics course on student achievement. We examine the effects of participating in a peer-led study group (PLSG) for middle-year engineering students enrolled in a thermodynamics course with an average failure rate of 47%. As the name implies, PLSGs are student-instructed, with an undergraduate teaching aide (UGTA) playing a supportive facilitator role rather than providing direct instruction. The intervention builds on the work of Treisman and colleagues (e.g., Fullilove and Treisman 1990; Hsu, Murphy, and

Treisman 2011; Treisman 1983) and is designed to promote social and academic support among students through cooperative problem-solving. While Treisman's model has been previously adopted in general first-year engineering courses (Flores et al. 2010; Minin et al. 2016; Pazos et al. 2007; Viera et al. 2019), our study uniquely utilizes PLSGs in a gateway engineering thermodynamics course typically taken at the end of the second year or beginning of the third year in an undergraduate engineering curriculum. In this paper, we describe the effects of pilot efforts to implement PLSGs, which increased course grades, pass rates, and graduation rates compared to conventional recitations led by a graduate teaching assistant (GTA). Our research supports PLSGs as a promising practice that promotes social interaction, is easy to implement, and has demonstrated effectiveness on academic achievement.

## LITERATURE REVIEW

Taking high-challenge foundation courses can amount to stressful experiences for engineering students. One critical consideration in high-stress performance situations is to ensure that there is a balance between high expectations and high social support. Pelz and Andrews (1966) argued that high levels of challenge (expectations) and security (support) are essential for high performance in achievement-related settings. Edmonson and Roloff (2008) echoed this sentiment, finding that optimal learning occurs when high accountability for meeting demanding goals meets high "psychological safety" among learners. Cooperative learning, especially cooperative-based group learning similar to what occurs in PLSGs, provides students with academic and social support in high-challenge learning environments (Johnson et al. 1991; Smith et al. 2005; Streveler and Smith 2020).

Several meta-analyses of cooperative learning among college students have underscored the importance of providing students with social support within academic settings. Johnson, Johnson, and Smith (1991) found in a meta-analysis of 305 studies that cooperative effort promotes greater learning among college students than competing with others (effect size=0.68) or working alone (effect size=0.55). This finding held irrespective of students' gender, ethnicity, cultural background, language, social class, or ability and was tied to greater perceived social support (both academically and personally) from peers and instructors. In another meta-analysis, Springer, Stanne, and Donovan (1999) reported mean effect sizes for students' achievement and persistence of .51 and .46 for cooperative learning in first-year science, technology, engineering, and mathematics (STEM) courses. The authors observed, "The .51 effect size of small-group learning on achievement reported in the study would move a student from the 50th percentile to the 70th on a standardized test [and that] a .46 effect size on students' persistence [would be] enough to reduce attrition in STEM courses and programs by 22%." Other studies have likewise shown the positive effects of social support promoted by cooperative learning on undergraduate students, specifically, that it improves social adjustment and integration into college life, increases student sense of belonging, and reduces incongruencies between students' interests and college curricula (Smith et al. 2005), all of which ultimately improves academic achievement and persistence to graduation (Arendale 2004; Cámara-Zapata and Morales 2020; Felder and Brent 2007). The following section describes the details and benefits of peer-led learning, a specific type of cooperative learning driving the present work.

**Peer-Led Learning**

The PEER-led, Student-Instructed STudy group (PEERSIST) model promotes academic and social support among students through cooperative problem-solving. Lazar (1995) supports this approach, concluding that study groups profoundly serve students' intellectual, social, and emotional needs by allowing them a safe space to take risks, experiment with ideas, and challenge each other's thinking.

PEERSIST builds on a model initially developed by Philip Uri Treisman in 1977 at the University of California, Berkeley, based on observations that Asian-American students in self-formed study groups achieved higher grades than their peers in a challenging gateway mathematics course (Fullilove and Treisman 1990; Hsu et al. 2011; Treisman 1983). Treisman adapted this study group model into the Emerging Scholars Program (ESP) for African American students, significantly increasing their course achievement and persistence in the major (Fullilove and Treisman 1990).

A similar method called Peer-Led Team Learning (PLTL) was implemented in large-enrollment general chemistry courses at the City College of New York in the early 1990s. Students were assigned to small, peer-led groups to solve problems collaboratively (Gosser, Kampmeier, and Varma-Nelson 2010; Wilson and Varma-Nelson 2016). Funding from the National Science Foundation (NSF) has contributed to the growth of PLTL literature, including a PLTL guidebook, manuals with examples, and training for peer leaders (Wilson and Varma-Nelson 2016). The national PTLT office also maintains a website that covers PLTL-related studies and information (Arendale 2004).

ESP and PLTL are nearly identical. In both approaches, students work in small groups of four to eight peers to solve course-related problems of moderate to extreme difficulty developed by the instructor. In addition, graduate (GTA) or undergraduate (UTA) teaching aides are trained to facilitate the small group session and encourage every student to collaborate with their peers and participate in the discussion. GTAs and UTAs only intervene with well-planted hints or suggestions if students get very stuck or are pursuing an approach that will not lead to a correct solution (Dreyfuss et al. 2015; Eberlein et al. 2008; Fullilove and Treisman 1990; Hsu et al. 2011; Streitwieser and Light 2010; Wilson and Varma-Nelson 2016). A difference between the approaches is that, in ESP, students are grouped by achievement level since Treisman discovered that higher-achieving students tend to dominate mixed-ability groups (Tough 2021). This feature has not been mentioned in PLTL literature.

The ESP and PLTL approaches are not remedial. Problems that students work on are designed to challenge them in order to promote both collaboration and perseverance. During the session, students are expected to help one another solve the session problems by sharing their ideas and critiquing their peers' work (Asera 2001; Fullilove and Treisman 1990; Gafney and Varma-Nelson 2008; Tien, Roth, and Kampmeier 2002). The purpose of peer-to-peer interaction in these small groups is for students to discover discipline-based methods and concepts through trial and error so that they learn to think like experts. Ensuring that each student acquires that capability is more important than the group getting the correct answer to each problem (Fullilove and Treisman 1990; Philip Uri Treisman, personal communication, March 23, 2020).

ESP and PLTL have been implemented in STEM courses for decades (Duncan and Dick 2000; Murphy and Treisman 2008; Ta et al. 2021; Tenney and Houck 2003; Treisman 1983). Research related to both ESP and PLTL has consistently shown improvement in students' academic performance, success, motivation, satisfaction, and persistence in the major (Hsu et al. 2011; Gafney and Varma-Nelson 2008). Some researchers have cited the ESP model in studies of PLTL implementation (Liou-Mark, Dreyfuss, and Younge 2010; Powell et al. 2012; Snyder et al. 2016; Tien et al. 2002). However, up to now, most literature related to one approach has not referenced the other despite sharing nearly identical features and advancing similar outcomes.

This paper builds on the literature base in three ways. First, the study of peer-led learning in engineering is rare relative to other disciplines (see Flores et al. 2010; Minin et al. 2016; Pazos et al. 2007; Viera et al. 2019, for exceptions). Second, whereas previous research has concentrated primarily on first-year courses, our study uniquely utilizes the method in a gateway engineering thermodynamics course typically taken in students' second or third year. Third, our study compares student achievement obtained through PLSGs with that obtained through conventional GTA-led recitation sessions. We explore two main research questions:

1. To what extent do PLSGs promote student achievement (i.e., course grades, course pass rates, and graduation rates) in a gateway engineering thermodynamics course relative to standard TA-led recitations?
2. Do the effects of PLSGs, if any, accrue differently for transfer students, women students, and racial/ethnic minoritized students?

## METHODOLOGY

### Design-Based Research Approach

Our research on the effectiveness of PLSGs in the introductory thermodynamics course has been incremental and iterative. Development of the PLSGs has followed a Design-Based Research (DBR) approach, which emerged from design experiments introduced by Brown (1992) and Collins (1992). Collins, Joseph, and Bielaczyc (2004) described design experiments as “a way to carry out formative research to test and refine educational designs based on principles derived from prior research,” which is an excellent description of the intention of this research. DBR is often used to improve instructional processes and refine educational interventions (Brown 1992; Cobb et al. 2003; Collins 1992). DBR follows a data-driven process wherein educators (1) design instructional tools based on learning theories to address a learning problem, (2) implement and test instructional tools in the classroom, (3) analyze student learning to evaluate the tools, and (4) reflect and critique the tools and implementation processes and revise the tools for continuous improvement (Sandoval 2014). In this study, we developed and implemented PLSGs over three academic semesters. We iterated on the PLSGs in the second and third semesters, incorporating findings from and improving upon the methods used in the first.

**Research Team Positionality**

The research team brings a range of experience and expertise to this study. Author 1 is an assistant professor of mechanical engineering and the principal investigator for this study. Author 4 is a mechanical engineering Ph.D. student. Authors 1 and 4 are directly involved in teaching the introductory thermodynamics course described in this study as an instructor and GTA. Authors 2 and 6 are professors at the associate and emeritus levels who teach and conduct research in both engineering and engineering education. Authors 3 and 5 are engineering education researchers and education-based program evaluators. All team members are committed to student success and have worked on various projects related to this area, with experience ranging from two to more than fifty years. In addition, Authors 1, 2, 3, 4, and 6 have taken thermodynamics courses as part of completing undergraduate engineering degrees, have encountered the challenging concepts in these courses that make the subject difficult to learn, and are committed to identifying better ways to teach these concepts. The research team's interdisciplinary and diverse backgrounds enabled rich dialogue from multiple perspectives, strengthening our DBR process. Notably, we all have some degree of largely positive experience with cooperative learning as researchers, instructional staff, and/or students. We recognize that this experience influences our views of PLSGs as a lever for institutional change regarding how gateway engineering courses are taught broadly.

**Institutional and Course Context**

Thermodynamics (MAE 241) is situated within Arizona State University (ASU), a large comprehensive public research university that is open-access and “measured not by whom it excludes, but by whom it includes and how they succeed” (ASU 2023). Thermodynamics is a required course for

ASU undergraduates in mechanical and aerospace engineering and an elective for civil engineering students. Some students enroll from other engineering majors, and very few from non-engineering majors. The fifteen-week course enrolls between 300-400 students each semester and has an average failure rate of 47%. The course represents most students' first exposure to thermodynamics and, for transfer students, their first engineering course in a four-year university setting.

MAE 241 consists of a three-unit instructional component that meets for 75 minutes twice per week and a recitation component that meets for 50 minutes once per week. Student participation in the recitation component is required, with attendance taken each week. Recitations are typically led by a GTA and consist of approximately 25 students. During the recitation, the GTA solves two to three problems on a whiteboard while the students observe, similar to a lecture format. Students are encouraged to ask questions and take photographs of the whiteboard during the session and solutions are posted online at the end of the week. In addition, no peer interaction is encouraged in these settings, with most engagement occurring between students and the GTA (Jenkins et al. 2023). Approval to conduct our intervention in the course was obtained from ASU's review board for human subjects research.

### Intervention Design

PLSGs in MAE 241 consist of four to five students, grouped based on having similar scores on the first exam in the course. PLSG students work together to solve instructor-selected problems ranging from moderate to extreme difficulty. These problems are not necessarily ones encountered in homework or exams and, instead, are meant to help students develop a schema for solving thermodynamics problems and cultivate a discipline-based mindset co-constructed with peers through dialogue and interaction. The teaching aide in this design is an undergraduate student (UGTA) who observes students in the PLSGs and encourages group discussion. The UGTA only assists students with the problem-solving process if necessary. Rather than immediately intervening when students make mistakes, the UGTA allows members of the PLSG to notice and correct the group's mistakes. UGTA intervention only occurs when students make severe errors that other members cannot correct. As in the required recitation sessions, PLSG students are encouraged to take photographs of their work to refer to after the session is completed.

### First Iteration (Spring 2020)

The first iteration of the PLSG model was implemented in the spring of 2020 with four students from one section of MAE 241, each of whom scored below 60% on the first exam (Exam 1) and accepted the instructor's invitation to participate in an extra study group after receiving their first exam scores. The intervention lasted eight weeks, beginning after the first exam. Although the intervention was available to students of all achievement levels, the research team targeted recruitment toward those who earned less than 60% on the first exam based on historical data predicting that these students were likely to fail the course. The PLSG students met virtually via Zoom while in-person instruction was suspended during the COVID-19 pandemic. A UGTA initially hired to teach a recitation section of thermodynamics was selected to run the PLSGs based on their previous success in achieving course outcomes. The UGTA was trained to implement PLSGs with this small, self-selecting cohort. Training consisted of reading a brief description of Treisman's model (Tough 2021) and, following the reading, meeting with the course instructor for an hour-long training to discuss the purpose and role of the UGTA in the PLSGs. The course instructor selected 2-3 problems for students to solve during the weekly PLSG sessions. The UGTA solved the problems independently each week prior to conducting the PLSGs. The UGTA piloted the first iteration of the PLSGs and subsequently led the second iteration, described later in greater detail.

The PLSG students were matched with a comparison group of eight students who had similar demographics and Exam 1 scores and received no treatment. Data collected included four exam scores and a final grade for all twelve students, as well as interviews with two PLSG members. Despite having only seven fifty-minute PLSG sessions during the semester, three of four PLSG members passed the course, compared with three of eight comparison students. The research team was both surprised by and skeptical of these results. Even though there were only four students in the treatment condition and the design had minimal controls built-in, the results seemed notable. Could PLSGs generate such a dramatic improvement in student performance? We were encouraged by the exam results and comments from the two PLSG students interviewed, who said that they enjoyed the PLSG, the PLSG helped them pass the course, and the experience helped them adjust to ASU as transfer students. Both students recommended that PLSGs be continued in the future. Over the summer of 2020, we strengthened the design and prepared to run a more robust study the following semester.

### Second Iteration (Fall 2020 and Spring 2021)

We ran a second iteration of the PLSG model in the fall of 2020, increasing student participation and adding a comparison group for the eight-week intervention following the first exam. Whereas the previous iteration was supported by one of three faculty who taught the course, the second iteration was supported by all three instructors, each of whom invited the students in their sections to participate in the study. In addition, the Director of the ASU Fulton Schools of Engineering Office of Undergraduate Success & Engagement, encouraged by results from the first iteration, provided funding for an additional UGTA to help manage the larger sample. The UGTA followed a similar training procedure, reading a description of the Treisman model (Tough 2021) and meeting with the course instructor and research team to discuss the purpose of the PLSGs. The new UGTA also observed a PLSG led by the previously trained UGTA before leading a session. Following the conclusion of the fall of 2020, the second iteration of the PLSG model repeated in the spring of 2021.

The PLSGs operated the same way in the second iteration as in the first. The comparison group consisted of students who participated in an additional hour of standard TA-led recitation (TAR) over and above that required of all students. This setup ensured that students in the PLSG and comparison groups received the same amount of exposure (2 hours) to solving thermodynamics problems outside of class each week. Like the PLSGs, participation in the additional TAR session was voluntary and focused on the same problems students worked on in the PLSGs. However, unlike the PLSGs, the TA leading the session presented these problems in lecture style. While students could ask the TA questions, there was limited student-student discussion, which we hypothesized was a critical feature of the PLSG model.

Students were recruited to the intervention after receiving their first exam scores and assigned to either a PLSG or TAR based on their time availability. In the fall of 2020, fifty of 299 (17%) students were recruited to participate in either a PLSG or TAR; at the end of the term, 22 PLSG students and 5 TAR students had attended at least four out of six of their respective sessions. In the spring of 2021, 62 of 375 (17%) students were recruited to participate in either a PLSG or TAR; at the end of the term, 26 PLSG and 10 TAR students had attended at least four out of six of their respective sections. Combining the two semesters, a total of 48 PLSG students and 15 TAR students were included in the study. Attrition between the PLSG and TAR groups was similar in both semesters. We compared the final grades and pass rates of students who started the PLSG or TAR but did not attend a minimum of four sessions to those of the no-treatment group and, upon finding no statistically significant differences, included these students in the no-treatment group.

**Data Source**

Data for this paper includes the exam scores, course pass rates, and four- or five-year graduation rates for all students enrolled in MAE 241 during the fall of 2020 and spring of 2021. Appendix A summarizes the course and student demographic information for these two semesters, as collected from the ASU university registrar. As shown in Table 6, students with similar demographics and educational backgrounds were enrolled each semester, with a few exceptions. Higher percentages of seniors and aerospace engineering majors were enrolled in the fall of 2020, while higher percentages of sophomores and mechanical engineering majors were enrolled in the spring of 2021 (academic level: $\chi^2(2) = 11.590$, $p = .009$; major: $\chi^2(2) = 39.789$, $p < .001$). Whereas mechanical engineering majors at ASU typically take the introductory thermodynamics course in the fourth semester of their undergraduate degree program, aerospace engineering majors typically take the course in their fifth semester or later. Further analysis revealed that students who took the course in the spring of 2021 tended to have lower scores on exams 1-3, lower final grades, a lower course pass rate, and (as expected) a lower graduation rate upon one-year follow-up, all at a significance level of $p < 0.05$, compared to students who took the course in the fall of 2020 (Appendix A, Table 7). Because of these semester differences, we collapsed the two semesters' worth of data into a single dataset but included semester as a control variable in our analyses used to address each research question. All analyses were performed in Stata 15 (Stata 2017).

## RESULTS

### Descriptive Statistics

Tables 1 and 2 compare students in the PLSG, TAR, and no-treatment groups based on their demographic characteristics and course outcomes. Compared to the no-treatment group, the PLSG group had a statistically higher proportion of racial/ethnic minoritized students ($\chi^2(2) = 7.381$, $p = .025$; post-hoc $p = .007$). Students in the PLSG group also had a statistically lower average score on the first exam than the no-treatment group (Table 2, post-hoc $p = .007$). This result makes sense, given participation was voluntary, higher scoring students who did not need additional instruction did not sign up, and the research team's efforts to target recruitment in the PLSGs toward students who earned less than 60% on their first exam.

Kruskal-Wallis H tests revealed non-significant differences among the groups' final exam scores or final course grades (Table 2). However, these tests do not account for potential baseline differences in student performance. To address this, we also conducted Quade tests (non-parametric ANCOVAs that allow for the inclusion of covariates) using students' first exam scores as a covariate to control for prior academic performance. The Quade tests revealed that PLSG students earned both significantly higher final exam scores ($F(2, 558) = 4.528$, $p = .011$; post-hoc $p = .010$) and higher final grades ($F(2, 564) = 11.659$, $p < .001$; post-hoc $p < .001$) than their no-treatment peers, controlling for first exam scores. Students in the PLSG group also had a higher course pass rate (post-hoc $p < .001$) and a higher graduation rate approximately one year post-taking the course (post-hoc $p < .001$) than students in the no-treatment group.

### Linear Regression Model: Final Course Grade

Table 3 summarizes the results of the multivariate linear regression model in which students' final course grade was predicted as a function of treatment condition (reference group: PLSG), controlling for the semester they took MAE 241, the instructor with whom they took the course, their score on the first exam, their statuses as a transfer, woman, racial/ethnic minoritized, and/or engineering student, and their age and cumulative grade point average (GPA) at the beginning of the course. Students who withdrew from the course did not receive a final grade and, therefore, were excluded

*Table 1. Course and student information: PLSG vs. TAR vs. No-treatment.*

| | PLSG | TAR | No-treatment |
|---|---|---|---|
| **Course Information** | | | |
| Number of groups | 20 | 5 | - |
| Number of students | 48 | 15 | 611 |
| **Student Information** | | | |
| Semester | | | |
| Fall 2020 | 45.8% | 33.3% | 44.5% |
| Spring 2021 | 54.2% | 66.7% | 55.5% |
| Gender | | | |
| Male | 77.1% | 86.7% | 83.1% |
| Female | 22.9% | 13.3% | 16.9% |
| Race/ethnicity | | | |
| American Indian/Alaska Native | 0.0% | 0.0% | 0.5% |
| Asian/Asian American | 18.8% | 13.3% | 7.0% |
| Black/African American | 6.3% | 0.0% | 2.6% |
| Hispanic/Latino | 41.7% | 26.7% | 25.9% |
| International | 0.0% | 0.0% | 6.2% |
| Native Hawaiian/Pacific Islander | 0.0% | 0.0% | 0.3% |
| White/European American | 29.2% | 53.3% | 52.9% |
| Multiracial | 4.2% | 6.7% | 3.8% |
| Not Available | 0.0% | 0.0% | 0.8% |
| Black/Latinx/Indigenous/Pacific Islander[1] | | | |
| Yes | 47.9% | 26.7% | 29.3% |
| All others | 52.1% | 73.3% | 70.7% |
| U.S. Citizenship Status | | | |
| Yes | 91.7% | 93.3% | 90.7% |
| All others | 8.3% | 6.7% | 9.3% |
| Admit Type | | | |
| First time student | 60.4% | 66.7% | 75.6% |
| Transfer student | 39.6% | 33.3% | 24.4% |
| Academic Level | | | |
| Sophomore | 16.7% | 13.3% | 13.7% |
| Junior | 58.3% | 60.0% | 57.8% |
| Senior | 22.9% | 20.0% | 26.7% |
| Post-bachelor's | 2.1% | 6.7% | 1.8% |
| Academic Major | | | |
| Aerospace engineering | 22.9% | 0.0% | 27.0% |
| Civil engineering | 18.8% | 6.7% | 13.6% |
| Mechanical engineering | 54.2% | 86.7% | 52.7% |
| Other engineering | 2.1% | 0.0% | 4.1% |
| Non-engineering | 2.1% | 6.7% | 2.6% |
| Mean Age (*SD*) | 21.5 (3.3) | 21.9 (5.4) | 21.3 (3.3) |
| Mean Cumulative GPA (*SD*)[2] | 3.45 (.37) | 3.37 (1.01) | 3.31 (.64) |

[1] We shorten American Indian/Alaska Native, Black/African American, Hispanic/Latin American, and Native Hawaiian/Pacific Islander to "Black/Latinx/Indigenous/Pacific Islander" in this paper.
[2] Cumulative GPA was reported by the registrar on a 0-4 scale.

***Table 2. Exam scores, final grades, and pass rates: PLSG vs. TAR vs. no-treatment.***

| | PLSG Group (n = 48) | TAR Group (n = 15) | No-treatment Group (n = 508) | Comparison Results |
|---|---|---|---|---|
| Mean Exam 1 Score (*SD*)[1,2] | 49.2 (13.8) | 56.3 (20.9) | 59.2 (21.6) | $H = 12.468$, $p = .002$ |
| Mean Exam 2 Score (*SD*) | 67.1 (15.6) | 62.6 (13.9) | 65.6 (21.8) | $H = 1.325$, $p = .516$ |
| Mean Exam 3 Score (*SD*) | 65.5 (14.1) | 68.8 (16.4) | 62.3 (22.5) | $H = 1.224$, $p = .542$ |
| Mean Final Exam Score (*SD*) | 68.6 (17.2) | 67.6 (23.6) | 64.7 (25.4) | $H = .265$, $p = .876$ |
| Mean Final Grade (*SD*) | 80.5 (8.2) | 80.5 (8.7) | 76.5 (16.7) | $H = 1.009$, $p = .604$ |
| Course Pass Rate[3] | 91.7% | 86.7% | 65.6% | $\chi^2(2) = 16.321$, $p < .001$ |
| Graduation Rate[3] | 89.6% | 73.3% | 63.0% | $\chi^2(2) = 14.291$, $p < .001$ |

[1]Exam scores and final grades were measured on a 0-100 scale.
[2]Exam scores and final grades were compared using Kruskal Wallis H tests.
[3]Pass rates and graduation rates were compared using a chi-square differences test.

from this analysis. As shown in Table 3 (Model 2), students in the no-treatment obtained significantly lower average final course grades than their peers in the PLSG group (B = -9.97, $p < .001$); however, no significant difference was seen between students in the PLSG and TAR groups. We also tested the interaction effects between transfer, woman, and racial/ethnic minoritized student status and each treatment condition on final course grade. No interaction terms for any of these models were significant, indicating no differences in the effect of the intervention on final course grade between underrepresented and non-underrepresented students.

**Logistic Regression Model: Course Pass Status**

Table 4 displays the results of a multivariate logistic regression model in which students' pass/fail status (1 = pass, 0= fail) was predicted as a function of treatment condition (reference group: PLSG), controlling for the semester they took MAE 241, the instructor with whom they took the course, their statuses as a transfer, woman, racial/ethnic minoritized, and/or engineering student, and their age and cumulative GPA at the beginning of the course. This analysis includes all students, including those who withdrew from the course before taking the first exam; therefore, students' first exam scores were excluded from the model. As shown in Table 4, students in the no-treatment group were significantly less likely to pass the course relative to students in the PLSG group (log odds = -1.78, $p = .001$); however, as with final course grade, no significant difference was seen between students in the PLSG and TAR groups. Similar to the linear regression model predicting students' final course grades, we tested the interaction effects between transfer, woman, and racial/ethnic minoritized student status and each treatment condition on course pass status, respectively. In addition, we found no significant interaction effects, indicating no differences in the effect of the intervention on course pass status between underrepresented and non-underrepresented students.

***Table 3. Linear regression results for final course grade (n=560).***

| | Unstandardized Beta Coefficients | | | | |
|---|---|---|---|---|---|
| **Variable** | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** |
| (Intercept) | 68.60 (2.46) *** | 8.04 (9.60) | 6.61 (9.82) | 8.24 (9.67) | 7.54 (9.59) |
| Condition: No-treatment group[1] | –3.86 (2.71) | –9.97 (2.24) *** | –8.50 (2.61) ** | –10.28 (2.45) *** | –9.44 (2.57) *** |
| Condition: TAR group[1] | –.99 (6.38) | –6.10 (4.64) | –4.93 (6.03) | –6.36 (5.33) | –5.32 (5.77) |
| Semester: Spring 2021[2] | - | 4.54 (2.72) | 4.62 (2.74) | 4.56 (2.73) | 4.60 (2.75) |
| Instructor: Instructor 1[3] | - | 9.27 (3.65)* | 9.30 (3.67)* | 9.25 (3.66)* | 9.24 (3.68)* |
| Instructor: Instructor 2[3] | - | –.70 (2.69) | –.68 (2.71) | –.71 (2.70) | –.71 (2.70) |
| Instructor: Instructor 3[3] | - | -4.69 (4.35) | –4.77 (4.36) | –4.67 (4.36) | –4.73 (4.37) |
| Instructor: Instructor 4[3] | - | –.53 (2.75) | –.56 (2.77) | –.53 (2.76) | –.53 (2.76) |
| Exam 1 Score (0-100) | | .65 (.05)*** | .65 (.05)*** | .64 (.05)*** | .65 (.05)*** |
| Transfer student: Yes[4] | - | 2.26 (2.23) | 5.72 (4.03) | 2.30 (2.25) | 2.22 (2.25) |
| Gender: Women[5] | - | .81 (1.98) | .71 (1.99) | –.44 (4.76) | .81 (1.99) |
| Black/Latinx/Indigenous/Pacific Islander: Yes[4] | - | –2.25 (2.04) | –2.31 (2.04) | –2.24 (2.05) | –1.19 (3.82) |
| Engineering student: Yes[4] | - | 6.71 (5.75) | 6.55 (5.81) | 6.70 (5.76) | 6.73 (5.75) |
| Age (in years) | - | –.17 (.27) | –.16 (.27) | –.17 (.27) | –.17 (.27) |
| Cumulative GPA (0-4) | - | 8.47 (2.21)*** | 8.41 (2.24)*** | 8.49 (2.21)*** | 8.44 (2.22)*** |
| Condition: No-treatment * Transfer student: Yes | - | - | –3.99 (4.48) | - | - |
| Condition: TAR * Transfer student: Yes | - | - | –2.98 (9.58) | - | - |
| Condition: No-treatment * Gender: Women | - | - | - | 1.41 (5.20) | - |
| Condition: TAR * Gender: Women | - | - | - | 1.06 (7.36) | - |
| Condition: No-treatment * Black/Latinx/Indigenous/Pacific Islander: Yes | - | - | - | - | –1.17 (4.43) |
| Condition: TAR* Black/Latinx/Indigenous/Pacific Islander: Yes | - | - | - | - | –2.12 (9.42) |

Robust standard errors are reported in parentheses. *, **, *** indicate significance at the 90%, 95%, and 99% level, respectively.
[1] Reference group: PLSG group
[2] Reference group: Fall 2020
[3] Reference group: Instructor 5
[4] Reference group: All others
[5] Reference group: Men

## Logistic Regression Model: Graduation Status

Table 5 shows the results of the multivariate logistic regression model predicting students' graduation status approximately one year post-taking the course, given similar controls to the logistic regression model predicting students' pass status. Most engineering majors enrolled in MAE 241 take the course in the spring of their second year or the fall of their third year; thus, we expected that most students

***Table 4. Logistic regression results for course pass status (n=672).***

| | Log-odds Coefficients | | | | |
|---|---|---|---|---|---|
| **Variable** | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** |
| (Intercept) | 2.40 (.52) *** | 2.45 (1.34) | 1.49 (1.18) | 2.26 (1.13)* | 2.20 (1.16) |
| Condition: No-treatment group[1] | –1.75 (.53)** | –1.78 (.56)** | –1.23 (.59)* | –1.56 (.58)** | –1.43 (.68)* |
| Condition: TAR group[1] | –.53 (.92) | –.03 (1.21) | –.54 (.98) | –.17 (1.04) | –.48 (1.12) |
| Semester: Spring 2021[2] | - | –.70 (.30)* | –.68 (.29)* | –.68 (.29)* | –.69 (.29)* |
| Instructor: Instructor 1[3] | - | –.02 (.50) | –.01 (.47) | –.03 (.47) | –.01 (.47) |
| Instructor: Instructor 2[3] | - | –.68 (.38) | –.62 (.37) | –.67 (.37) | –.66 (.37) |
| Instructor: Instructor 3[3] | - | –.52 (.47) | –.52 (.46) | –.51 (.46) | –.50 (.46) |
| Instructor: Instructor 4[3] | - | –.37 (.38) | –.35 (.37) | –.36 (.37) | –.36 (.37) |
| Transfer student: Yes[4] | - | .37 (.27) | 1.10 (1.05) | .34 (.26) | .36 (.26) |
| Gender: Women[5] | - | –.18 (.25) | –.19 (.24) | –.17 (1.06) | –.18 (.24) |
| Black/Latinx/Indigenous/Pacific Islander: Yes[4] | - | –.47 (.20)* | –.47 (.19)* | –.46 (.19)* | –.22 (.96) |
| Engineering student: Yes[4] | - | –.39 (.56) | –.32 (.51) | –.39 (.51) | –.36 (.50) |
| Age (in years) | - | –.10 (.03)** | –.09 (.03)** | –.09 (.03)** | –.10 (.03)** |
| Cumulative GPA (0-4) | - | .92 (.22)*** | .96 (.19)*** | –.88 (.17)*** | .89 (.17)*** |
| Condition: No-treatment * Transfer student: Yes | - | | –.85 (1.06) | - | - |
| Condition: TAR * Transfer student: Yes | - | - | 2.41 (2.23) | - | - |
| Condition: No-treatment * Gender: Women | - | - | - | –.03 (1.09) | - |
| Condition: TAR * Gender: Women | - | - | - | –.44 (2.12) | - |
| Condition: No-treatment * Black/Latinx/Indigenous/Pacific Islander: Yes | - | - | - | - | –.26 (.98) |
| Condition: TAR* Black/Latinx/Indigenous/Pacific Islander: Yes | - | - | - | - | .87 (1.78) |

Robust standard errors are reported in parentheses. *, **, *** indicate significance at the 90%, 95%, and 99% level, respectively.
[1] Reference group: PLSG group
[2] Reference group: Fall 2020
[3] Reference group: Instructor 5
[4] Reference group: All others
[5] Reference group: Men

taking the course in the fall of 2020 or the spring of 2021 would have graduated with their degree by the fall of 2023, assuming a four to five-year graduation rate. Similar to the previously run models for final course grade and pass status, we determined that students in the no-treatment group were significantly less likely to have graduated with their degree than PLSG students (log odds = -1.46, $p$ = .003); however, no significant difference was observed in the TAR group relative to the PLSG group. Finally, we found no significant interaction effects, indicating no differences in the effect of the intervention on graduation status between underrepresented and non-underrepresented students.

***Table 5. Logistic regression results for graduation status (n=672).***

| | Log-odds Coefficients | | | | |
|---|---|---|---|---|---|
| **Variable** | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** |
| (Intercept) | 2.15 (.47)*** | –.04 (1.25) | .30 (1.18) | .27 (1.13) | –.50 (1.12) |
| Condition: No-treatment group[1] | –1.62 (.48)** | –1.46 (.49)** | –1.73 (.67)* | –1.64 (.58)** | –.88 (.54) |
| Condition: TAR group[1] | –1.14 (.75) | –.94 (.81) | –1.61 (.94) | –.90 (.90) | –.32 (.89) |
| Semester: Spring 2021[2] | - | –.75 (.29)* | –.73 (.29)* | –.72 (.29)* | –.70 (.29)* |
| Instructor: Instructor 1[3] | - | –.01 (.45) | –.03 (.44) | –.02 (.45) | –.03 (.45) |
| Instructor: Instructor 2[3] | - | –.22 (.36) | –.22 (.35) | –.22 (.35) | –.22 (.35) |
| Instructor: Instructor 3[3] | - | –.54 (.45) | –.52 (.44) | –.53 (.44) | –.55 (.44) |
| Instructor: Instructor 4[3] | - | –.28 (.35) | –.28 (.34) | –.28 (.34) | –.27 (.34) |
| Transfer student: Yes[4] | - | .31 (.24) | –.61 (.90) | .32 (.25) | .27 (.25) |
| Gender: Women[5] | - | .30 (.25) | .31 (.24) | –.82 (.93) | .29 (.24) |
| Black/Latinx/Indigenous/Pacific Islander: Yes[4] | - | –.14 (.19) | –.14 (.19) | –.13 (.19) | .98 (1.00) |
| Engineering student: Yes[4] | - | –.19 (.52) | –.16 (.50) | –.20 (.51) | –.19 (.50) |
| Age (in years) | - | –.03 (.03) | –.03 (.03) | –.03 (.03) | –.03 (.03) |
| Cumulative GPA (0-4) | - | .97 (.22)*** | .94 (.17)*** | .94 (.17)*** | .93 (.18)*** |
| Condition: No-treatment * Transfer student: Yes | - | - | .92 (.92) | - | - |
| Condition: TAR * Transfer student: Yes | - | - | 2.16 (2.15) | - | - |
| Condition: No-treatment * Gender: Women | - | - | - | 1.21 (.96) | - |
| Condition: TAR * Gender: Women | - | - | - | –.70 (1.66) | - |
| Condition: No-treatment * Black/Latinx/Indigenous/ Pacific Islander: Yes | - | - | - | - | –1.16 (1.02) |
| Condition: TAR* Black/Latinx/Indigenous/ Pacific Islander: Yes | - | - | - | - | –1.86 (1.66) |

Robust standard errors are reported in parentheses. *, **, *** indicate significance at the 90%, 95%, and 99% level, respectively.
[1] Reference group: PLSG group
[2] Reference group: Fall 2020
[3] Reference group: Instructor 5
[4] Reference group: All others
[5] Reference group: Men

## DISCUSSION

Our statistical analyses demonstrated that students who participated in the PLSGs not only performed better in the introductory thermodynamics course but had a higher course pass rate and a higher graduation rate approximately one year post-taking the course compared to their no-treatment counterparts. These were the only significant differences observed between the PLSG, TAR (control), and no-treatment groups. No significant differences were seen between students in the PLSG and TAR groups, both of whom received an additional hour of thermodynamics education outside the usual course each week. There were also no differences in the effectiveness of the intervention by transfer, woman, and/or racial/ethnic minoritized student status, indicating that the intervention provides at least the same educational advantages to underrepresented groups as to non-underrepresented groups.

At first glance, the lack of significant differences between the PLSG and TAR groups might suggest that driving the PLSGs' success was the additional hour of instruction each week; to some extent, this additional time may have been a factor. However, it is worth noting that the TA leading the TAR sessions was highly trained, with multiple semesters of experience teaching thermodynamics, and met with the faculty teaching the course weekly for additional training during the study. TAR groups were also similarly sized to PLSG groups (4 students), allowing for very personalized instruction similar to tutoring in small groups. Students in the PLSGs were given the same problem set each week as students in the TARs, for the same time duration, but with comparatively little guidance from their facilitator, only when stuck or pursuing an incorrect path. Thus, our results indicate that students learning cooperatively with little assistance from a facilitator performed just as well in the course as those in similarly sized groups taught directly by a highly trained and motivated TA. This finding is very promising for future expansion into the course, as it demonstrates that cooperative learning is just as helpful as recitation sessions run by a highly trained and motivated TA for the course.

Despite no differences between the PLSG and TAR groups, the PLSG intervention has been met with significant faculty and institutional support and positive student feedback. (Word among students in the course has been that to pass thermodynamics, sign up for a PLSG). As a result, the course has now fully transitioned to a PLSG model. This third iteration addresses earlier challenges by embedding PLSGs directly into the course structure. Each class session of one hundred students is split into four recitations, with two each assigned to the PLSG and TAR (control) formats, respectively. This change ensures that all students receive just one hour of additional thermodynamics education as part of their regularly scheduled recitations, as opposed to PLSG students needing to participate in an additional hour outside usual course hours. Notably, each PLSG session has three to four facilitators present in this new iteration. Yet, scaling has not been a problem due to the use of UGTA facilitators, who are less costly to hire than graduate teaching assistants. Facilitators are prepared using newly developed instructional videos to ensure quality and consistency across groups (modules describing best practices in implementation, https://lth.engineering.asu.edu/reference-guide/peer-led-study-groups/).

Importantly, while the PLSG approach involves peer-led group work, it differs from typical group work due to the oversight of the facilitator and, therefore, is less prone to barriers that can limit peer groups, such as conformity, decision inertia, or groupthink (Reader, 2017). The facilitators' extensive training and continual oversight ensure that the groups remain focused on the task while providing an environment to deeply engage in problem formulation and solution. Students can take risks and make mistakes with the assurance that they will not get too far off track before the facilitator redirects the problem-solving. Facilitators and students have noted an increase in their problem-solving confidence over the course of the sessions. Research on the PLSG model is ongoing, with current studies focused on long-term student outcomes, the impact of facilitation practices, and strategies for sustainable scaling across additional courses and disciplines.

## LIMITATIONS

This study contained limitations. Participation in the PLSGs was quite low relative to overall student enrollment in the introductory thermodynamics course, which means that our conclusions about the intervention were drawn from very small group sizes. In addition, since the research team targeted recruitment in the PLSG to those who earned less than 60% on the first exam, membership in the PLSG group was imbalanced. Sessions were also held as an additional hour outside of class, which affected student attendance. Further research is needed to determine whether the outcomes in the present study generalize to all students in the course. Self-selection of the study participants is another potential limitation, as the participants in the PLSGs and/or TARs may represent a more initiative-taking group relative to their peers. This limitation is being addressed in the next iteration of the study, where all students will participate in either the intervention (PLSG) or control (TAR) group. We also tested the PLSGs at a single institution and in a single gateway course. More work is needed to assess the effectiveness of PLSGs in different classroom and university settings.

## CONCLUSIONS

PLSGs benefitted a gateway engineering thermodynamics course by increasing students' academic achievement. Students in the PLSGs performed at similar levels to those in the control–in which an experienced and motivated TA gave small-group tutoring to students–despite receiving far less instruction and guidance. Both groups performed statistically similarly in both final course grade, course pass rate, and graduation rate approximately one year post-taking the course. However, only students in the PLSG group outperformed similar students in the no-treatment group. These results support implementation of Treisman's model of peer-led study groups (Fullilove and Treisman 1990; Hsu et al. 2011; Treisman 1983) in upper-level gateway engineering courses. Future iterations of the PEERSIST model will implement the PLSG style of recitation directly into the thermodynamics course.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr. Tami Coronella, Director of Engineering Student Success and Engagement, and the Ira A. Fulton Schools of Engineering Office of Academic and Student Affairs at Arizona State University for their support. This material is based upon work supported by NSF Grant No. 2205001. Any opinions, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

Arendale, David R. 2004. *Pathways of Persistence: A Review of Postsecondary Peer Cooperative Learning Programs.* University Digital Conservancy. https://hdl.handle.net/11299/200463.

Asera, Rose. 2001. *Calculus and Community: A History of the Emerging Scholars Program. A Report of the National Task Force on Minority High Achievement.* College Entrance Board. https://eric.ed.gov/?id=ED562582.

Arizona State University. n.d. "ASU Charter." https://www.asu.edu/about/charter-mission.

Brown, Ann L. 1992. "Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings." *Journal of the Learning Sciences* 2, no. 2: 141-178. https://doi.org/10.1207/s15327809jls0202_2.

Collins, Allan. 1992. "Toward a Design Science of Education." In *New Directions in Educational Technology*, edited by Eileen Scanlon and Tim O'Shea, 15-22. Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-77750-9_2.

Collins, Allan, Diana Joseph, and Katerine Bielaczyc. 2016. "Design Research: Theoretical and Methodological Issues." *Journal of the Learning Sciences* 13, no. 1: 15-42. https://doi.org/10.1207/s15327809jls1301_2.

Dreyfuss, A. E., Melanie Villatoro, Michael C. Loui, James E. Becvar, Geoffrey B. Saupe, and Wayne Johnson. 2015. "Getting Past the First Year: Retaining Engineering Majors." In *Proceedings of the Frontiers in Education Conference*, El Paso, Texas, October 21-25. https://doi.org/10.1109/FIE.2015.7344363.

Duncan, Hollis, and Thomas Dick. 2000. "Collaborative Workshops and Student Academic Performance in Introductory College Mathematics Courses: A Study of a Treisman Model Math Excel Program." *School Science and Mathematics* 100, no. 7: 365-373. https://doi.org/10.1111/j.1949-8594.2000.tb18178.x.

Eberlein, Thomas, Jack Kampmeier, Vicky Minderhout, Richard S. Moog, Terry Platt, Pratibha Varma–Nelson, and Harold B. White. 2008. "Pedagogies of Engagement in Science: A Comparison of PBL, POGIL, and PLTL." *Biochemistry and Molecular Biology Education* 36, no. 4: 262-273. https://doi.org/10.1002/bmb.20204.

Edmondson, Amy C., and Kathryn S. Roloff. 2008. "Overcoming Barriers to Collaboration: Psychological Safety and Learning in Diverse Teams." In *Team Effectiveness in Complex Organizations: Cross-Disciplinary Perspectives and Approaches*, edited by Eduardo Salas, Gerald F. Goodwin, and C. Shawn Burke, 217-242. Routledge. https://doi.org/10.4324/9780203889312.

Felder, Richard M., and Rebecca Brent. 2007. "Cooperative Learning." In *Active Learning: Models from the Analytical Sciences (ACS Symposium Series 970, Chapter 4)*, edited by Patricia Ann Mabrouk. American Chemical Society. https://doi.org/10.1021/bk-2007-0970.

Flores, Benjamin, James Becvar, Ann Darnell, Helmut Knaust, Jorge Lopez, and Josefina Tinajero. 2010. "Implementing Peer Led Team Learning in Gateway Science and Mathematics Courses for Engineering Majors." In *Proceedings of the American Society for Engineering Education (ASEE) Annual Conference and Exposition*, Louisville, Kentucky, June 20-23. https://peer.asee.org/16650.

Fullilove, Robert E., and Philip Uri Treisman. 1990. "Mathematics Achievement Among African American Undergraduates at the University of California, Berkeley: An Evaluation of the Mathematics Workshop Program." *The Journal of Negro Education* 59, no. 3: 463-478. https://doi.org/10.2307/2295577.

Gafney, Leo, and Pratibha Varma-Nelson. 2008. "Development and Implementation of Peer-Led Team Learning," In *Peer-led Team Learning: Evaluation, Dissemination, and Institutionalization of a College Level Initiative*, edited by Leo Gafney and Pratibha Varma-Nelson, 9-27. Springer. https://doi.org/10.1007/978-1-4020-6186-8_2.

Gosser Jr., David K., Jack A. Kampmeier, and Pratibha Varma-Nelson. 2010. "Peer-led Team Learning: 2008 James Flack Norris Award Address." *Journal of Chemical Education* 87, no. 4: 374-380. https://doi.org/10.1021/ed800132w.

Hsu, Eric, Teri J. Murphy, and Uri Treisman. 2008. "Supporting High Achievement in Introductory Mathematics Courses: What We Have Learned from 30 Years of Emerging Scholars Program." In *Making the Connection: Research and Teaching in Undergraduate Mathematics Education*, edited by Marilyn Paula Carlson and Chris Rasmussen, 205-220. Mathematical Association of America. https://doi.org/10.5948/UPO9780883859759.017.

Jenkins, Cody, D., Thien N. Y. Ta, Ryan J. Milcarek, Gary Lichtenstein, Samantha R. Brunhaver, and Karl A. Smith. 2023. "Work in Progress: PEERSIST – A Formation of Engineers Framework for Understanding Self-Efficacy and Persistence among Transfer Students." In *Proceedings of the American Society for Engineering Education (ASEE) Annual Conference and Exposition*, Baltimore, Maryland, June 25-28. https://peer.asee.org/44325.

Johnson, David W., Roger T. Johnson, and Karl A. Smith. 1991. *Active Learning. Cooperation in the College Classroom.* Interaction Book Company. ISBN 978-0939603145.

Lazar, Althier M. 1995. "Who is Studying in Groups and Why? Peer Collaboration Outside the Classroom." *College Teaching* 43, no. 2: 61-65. https://doi.org/10.1080/87567555.1995.9925517.

Lichtenstein, G., Helen L. Chen, Karl A. Smith, and Theresa A. Maldonado. 2014. "Retention and Persistence of Women and Minorities Along the Engineering Pathway in the United States." In *Cambridge Handbook of Engineering Education Research*, edited by Aditya Johri and Barbara M. Olds, 311-334. Cambridge University Press. https://doi.org/10.1017/CBO9781139013451.021.

Liou-Mark, Janet, A. E. Dreyfuss, and Lori Younge. 2010. "Peer Assisted Learning Workshops in Precalculus: An Approach to Increasing Student Success." *Mathematics and Computer Education* 44, no. 3: 249-259. https://www.semanticscholar.org/paper/Peer-Assisted-Learning-Workshops-in-Precalculus%3A-An-Liou-Mark-Dreyfuss/73ba015533121c9094647e7c8a2a0a8be71df5d5.

Lord, Susan M., and John C. Chen. 2014. "Curriculum Design in the Middle Years." In *Cambridge Handbook of Engineering Education Research*, edited by Aditya Johri and Barbara M. Olds, 181-200. Cambridge University Press. https://doi.org/10.1017/CBO9781139013451.014.

MacGregor, Jean, James L. Cooper, Karl A. Smith, and Pamela Robinson. 2000. *Strategies for Energizing Large Classes: From Small Groups to Learning Communities. The Jossey-Bass Higher and Adult Education Series.* Jossey-Bass. ISBN 978-0787953379.

Minin, Serge, David Varodayan, Christopher Schmitz, Brian Faulkner, Dong San Choi, and Geoffrey L. Herman. 2016. "Minority Merit: Improving Retention with Cooperative Learning in a First-Year Electronics Course." In *Proceedings of the IEEE Frontiers in Education Conference*, Erie, Pennsylvania, October 12-15. https://doi.org/10.1109/FIE.2016.7757611.

Pazos, Pilar, Denise Drane, Gregory Light, and Annette Munkeby. 2007. "A Peer Led Team Learning Program for Freshman Engineering Students: Impact on Retention." In *Proceedings of the American Society for Engineering Education (ASEE) Annual Conference and Exposition*, Honolulu, Hawaii, June 24-27. https://peer.asee.org/1926.

Pelz, Donald C., and Frank M. Andrews. 1966. "Autonomy, Coordination, and Stimulation, in Relation to Scientific Achievement." *Behavioral Science* 11, no. 2: 89-97. https://doi.org/10.1002/bs.3830110202.

Powell, Rita M., Christian Murphy, Adam Cannon, Joshua Gordon, and Arthi Ramachandran. 2012. *Emerging Scholars Program—A PLTL-CS Program that Increases Recruitment and Retention of Women in the Major.* University of Pennsylvania. https://repository.upenn.edu/cis_reports/974.

Reader, Tom W. 2017. "Team Decision Making." In *The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes*, edited by Eduardo Salas, Ramón Rico, and Jonathan Passmore, 271-296. John Wiley & Sons. https://doi.org/10.1002/9781118909997.ch12.

Sandoval, William. 2014. "Conjecture Mapping: An Approach to Systematic Design Research." *Journal of the Learning Sciences* 23, no. 1: 18-36. https://doi.org/10.1080/10508406.2013.778204.

Smith, Karl A., Sheri D. Sheppard, David W. Johnson, and Roger T. Johnson. 2005. "Pedagogies of Engagement: Classroom-based Practices." *Journal of Engineering Education* 94, no. 1: 87-101. https://doi.org/10.1002/j.2168-9830.2005.tb00831.x.

Snyder, Julia J., Jeremy D. Sloane, Ryan D. Dunk, and Jason R. Wiles. 2016. "Peer-led Team Learning Helps Minority Students Succeed." *PloS Biology* 14, no. 3: e1002398. https://doi.org/10.1371/journal.pbio.1002398.

Springer, Leonard, Mary Elizabeth Stanne, and Samuel S. Donovan. 1999. "Effects of Small-group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-analysis." *Review of Educational Research* 69, no. 1: 21-51. https://doi.org/10.3102/00346543069001021

Stains, Marilyne, Jordan Harshman, Megan K. Barker, Stephanie V. Chasteen, Renee Cole, Sue Ellen DeChenne-Peters, M. Kevin Eagan Jr. et al. 2018. "Anatomy of STEM Teaching in North American Universities." *Science* 359, no. 6383, 1468-1470. https://doi.org/10.1126/science.aap8892/

StataCorp. 2017. *Stata statistical software: Release 15.* https://www.stata.com/company/.

Streitwieser, Bernhard, and Gregory Light. 2010. "When Undergraduates Teach Undergraduates: Conceptions of and Approaches to Teaching in a Peer Led Team Learning Intervention in the STEM Disciplines—Results of a Two-year Study." *International Journal of Teaching and Learning in Higher Education* 22, no. 3: 346-356. https://eric.ed.gov/?id=EJ938570.

Streveler, Ruth A., and Karl A. Smith. 2020. "Opinion: Course Design in the Time of Coronavirus: Put on Your Designer's CAP." *Advances in Engineering Education* 8, no. 4: 1-19. https://advances.asee.org/opinion-course-design-in-the-time-of-coronavirus-put-on-your-designers-cap/.

Ta, Thien N. Y., Gary Lichtenstein, Cody D. Jenkins, Karl A. Smith, Ryan J. Milcarek, and Samantha R. Brunhaver. 2021. "The PEERSIST Project: Promoting Engineering Persistence through Peer-led Study Groups." In *Proceedings of the American Society for Engineering Education (ASEE) Virtual Conference*, June 22-26. https://peer.asee.org/37881.

Tenney, A., and B. Houck. 2003. "Peer-led Team Learning in Introductory Biology and Chemistry Courses: A Parallel Approach." *Journal of Mathematics and Science: Collaborative Explorations* 6, no. 1: 11-20. https://doi.org/10.25891/6PGZ-J157.

Tien, Lydia, T., Vicki Roth, and J. A. Kampmeier. 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, no. 7: 606-632. https://doi.org/10.1002/tea.10038.

Tough, Paul. 2021. *The Inequality Machine: How Universities are Creating a More Unequal World—And What to Do About It.* Random House. ISBN 978-1473589148.

Treisman, Philip Uri. 1983. "Improving the Performance of Minority Students in College-level Mathematics." *Innovation Abstracts* 5, no. 17: 1-4. https://eric.ed.gov/?id=ED234874.

Viera, Julian, Christina Convertino, Erika Mein, Elsa Q. Villa, and Sarah Hug. 2019. "Shifting Pre-calculus from a Gatekeeper to a Gateway Course." In *Proceedings of the Collaborative Network for Engineering and Computing Diversity Conference*, Crystal City, Virginia, April 14-22. https://peer.asee.org/31790.

Wilson, Sarah Beth, and Pratibha Varma-Nelson. 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, no. 10: 1686-1702. https://doi.org/10.1021/acs.jchemed.5b00862.

## APPENDIX

*Table 6. Course and student information: Fall 2020 vs. Spring 2021.*

| | Fall 2020 | Spring 2021 |
|---|---|---|
| **Course Information** | | |
| Number of sections | 3 | 4 |
| Number of instructors | 3 | 3 |
| Number of students | 299 | 375 |
| **Student Information** | | |
| Group Type | | |
| Peer Led Study Group (PLSG) | 7.4% | 6.9% |
| TA Led Recitation (TAR) | 1.7% | 2.7% |
| No treatment | 91.0% | 90.4% |
| Gender | | |
| Male | 83.3% | 82.4% |
| Female | 16.7% | 17.6% |
| Race/ethnicity | | |
| American Indian/Alaska Native | 0.3% | 0.5% |
| Asian/Asian American | 9.0% | 7.2% |
| Black/African American | 3.3% | 2.4% |
| Hispanic/Latin American | 27.1% | 26.9% |
| International | 3.7% | 7.2% |
| Native Hawaiian/Pacific Islander | 0.7% | 0.0% |
| White/European American | 50.8% | 51.5% |
| Multiracial | 4.3% | 3.5% |
| Not Available | 0.7% | 0.8% |
| Black/Latinx/Indigenous/Pacific Islander[1] | | |
| Yes | 31.4% | 29.8% |
| All others | 68.6% | 70.2% |
| U.S. Citizenship Status | | |
| Yes | 92.0% | 89.9% |
| All others | 8.0% | 10.1% |
| Admit Type | | |
| First time student | 74.6% | 74.1% |
| Transfer student | 25.4% | 25.9% |
| Academic Level | | |
| Sophomore | 9.4% | 17.6% |
| Junior | 58.2% | 57.6% |
| Senior | 30.1% | 23.2% |
| Post-bachelor's | 2.3% | 1.6% |
| Academic Major | | |
| Aerospace engineering | 37.1% | 17.3% |
| Civil engineering | 14.7% | 13.1% |
| Mechanical engineering | 42.5% | 62.4% |
| Other engineering | 2.3% | 5.1% |
| Non-engineering | 3.3% | 2.1% |
| Mean Age (*SD*) | 21.5 (3.4) | 21.2 (3.2) |
| Mean Cumulative GPA (*SD*)[2] | 3.34 (.60) | 3.31 (.66) |

[1] We shorten American Indian/Alaska Native, Black/African American, Hispanic/Latin American, and Native Hawaiian/Pacific Islander to "Black/Latinx/Indigenous/Pacific Islander" in this paper.
[2] Cumulative GPA was reported by the registrar on a 0-4 scale.

***Table 7. Exam scores, final grades, pass rates, and graduation rates: Fall 2020 vs. Spring 2021.***

| | Fall 2020 | Spring 2021 | Comparison Results |
|---|---|---|---|
| Mean Exam 1 Score (*SD*)[1,2] | 65.2 (20.1) | 52.6 (20.4) | $U = 26{,}292$, $p < .001$ |
| Mean Exam 2 Score (*SD*) | 72.1 (18.4) | 60.3 (21.8) | $U = 26{,}949$, $p < .001$ |
| Mean Exam 3 Score (*SD*) | 69.8 (21.2) | 56.7 (20.5) | $U = 24{,}246$, $p < .001$ |
| Mean Final Exam Score (*SD*) | 65.8 (22.3) | 64.6 (26.7) | $U = 39{,}586$, $p = .803$ |
| Mean Final Grade (*SD*) | 79.6 (14.8) | 74.7 (16.7) | $U = 32{,}361$, $p < .001$ |
| Course Pass Rate[3] | 75.9% | 61.6% | $\chi^2(1) = 15.664$, $p < .001$ |
| Graduation Rate[3] | 74.6% | 57.6% | $\chi^2(1) = 21.125$, $p < .001$ |

[1]Exam scores and final grades were measured on a 0-100 scale.
[2]Exam scores and final grades were compared using Mann Whitney U tests.
[3]Pass rates and graduation rates were compared using a chi-square differences test.

## AUTHORS

**Ryan J. Milcarek** (he, his) is an Associate Professor of Mechanical Engineering in the School for Engineering of Matter, Transport and Energy at Arizona State University. He obtained his bachelor's, master's, and doctoral degrees in the Mechanical and Aerospace Engineering Department at Syracuse University. He also obtained a master's in Energy Systems Engineering, Certificate of Advanced Study in Sustainable Enterprise, and Certificate in University Teaching while completing his graduate studies. His current research is focused on engineering education, solid oxide fuel cells, generation fuels and chemicals, microcombustion, manufacturing of ceramic materials, energy storage for demand response, and thermochemical energy storage. Email: ryan.milcarek@asu.edu.

**Samantha R. Brunhaver** (she, her) is an Associate Professor of Engineering in The Polytechnic School within the Ira A. Fulton Schools of Engineering at Arizona State University. She earned her bachelor's degree in mechanical engineering from Northeastern University, and her master's and doctoral degrees in Mechanical Engineering from Stanford University. Her current research interests include engineering career pathways, undergraduate student persistence, professional engineering practice, and faculty mentorship. Her 2020 NSF Faculty Early Career (CAREER) Award focuses on increasing the adaptability of undergraduate engineering students. Email: samantha.brunhaver@asu.edu.

**Thien N. Y. Ta** is an Educational Research Evaluator with Quality Evaluation Designs, a firm specializing in educational research and program evaluation. She obtained a bachelor's and master's degrees in mechanical engineering from HCMC University of Technology and Education, and HCMC University of Technology in Vietnam, and master's degrees in mechanical engineering and engineering education from Arizona State University. Her current research interests include STEM persistence, Learning Assistance in STEM, and educational evaluation. Email: thien@qualityevaluationdesigns.com.

**Cody D. Jenkins** (he, his) is a mechanical engineering doctorate student in the School for Engineering of Matter, Transport and Energy at Arizona State University. He obtained his bachelor's degree in mechanical engineering from Arizona State University. His current research specializes in building energy modeling and technoeconomic analysis. Email: cdjenki6@asu.edu.

**Gary Lichtenstein** (he, they) is founder and principal of Quality Evaluation Designs, a firm that conducts academic research, program evaluation, and program development for K-12 schools, colleges, and universities, non-profits, and government agencies nationwide. His research interests include STEM persistence and higher education transformation. Dr. Lichtenstein has twice been awarded the William Elgin Wickenden Award from ASEE for the *Journal of Engineering Education* (JEE) article each year that exemplifies the highest standards of scholarly research. Email: gary@qualityevaluationdesigns.com.

**Karl A. Smith** (he, him) is Emeritus Cooperative Learning Professor of Engineering Education at Purdue, which he joined in 2006 to help start the engineering education Ph.D. program. He is also Morse-Alumni Distinguished University Teaching Professor and Emeritus Professor of Civil, Environmental, and Geo-Engineering at the University of Minnesota. He has received numerous awards, including Fellow, Hall of Fame, and Lifetime Achievement Award from the

American Society for Engineering Education. Dr. Smith obtained his bachelor's and master's degrees in metallurgical engineering from Michigan Technological University and his doctorate in Educational Psychology from the University of Minnesota. His research focuses on engineering education, cooperative learning and structured controversy, and project management and teamwork. Email: ksmith@umn.edu.