A "Light Touch" Intervention to Improve Students' Perception of Faculty Caring, Belongingness and Academic Achievement

Franchesca Garcia, California Polytechnic State University, San Luis Obispo
Zoe L Rehnberg, California Polytechnic State University, San Luis Obispo
Leily Majidi, California Polytechnic State University, San Luis Obispo
Prof. Eric Ocegueda, California Polytechnic State University, San Luis Obispo
Alan Zhang, California Polytechnic State University, San Luis Obispo

Alan Zhang is an assistant professor of mechanical engineering at California Polytechnic State University, San Luis Obispo. His engineering education interests are on student non-cognitive factors, autonomy, and motivation. His prior work includes exploring the effectiveness of conceptual modeling kits in the engineering statics curriculum.

Dr. John Chen P.E., California Polytechnic State University, San Luis Obispo

John Chen is a professor of mechanical engineering. His interests in engineering education include conceptual learning, conceptual change, student autonomy and motivation, lifelong learning skills and behaviors, and non-cognitive factors that lead to stu

A "Light Touch" Intervention to Improve Students' Perception of Faculty Caring, Belongingness and Academic Achievement

Abstract

This methods full paper adapts a previous study in the literature to investigate transferability to different classroom contexts. Barriers to student success are complex: noncognitive and affective (NCA) factors are not related to talent or intelligence but are still shown to impact academic achievement. A previous study on a relatively low-effort intervention (and hence "light touch") – sending strategically timed and personalized emails – demonstrated several achievement outcomes for underrepresented minority students in high-enrollment, introductory-level courses at a large public university. This project adapts and tests this intervention in the small class sizes (35 students) of the mechanical engineering department at a mid-size public university to investigate the effects on two NCA factors of interest: students' perception of faculty caring and sense of belonging. Four instructors applied the intervention across a variety of second- and third-year classes in the Spring 2025 quarter (N = 291 total students). After the first major assessment (a quiz or midterm around the third or fourth week), low-performing students were randomly assigned to the treatment or control group. The treatment group (n = 41) received a personalized email about their individual score and methods to improve, while all other students received a generic email about overall class performance and available resources. The treatment was administered again after the second major assessment. The students' NCA factors were assessed through surveys given at the beginning and end of the 10-week quarter. Final course grades were collected as another outcome of interest. For almost all student demographic subgroups and outcomes of interest, 90% confidence intervals for the treatment effect included the null result, indicating that the treatment may have had less effect than the original study. We present a reflection on these results and on context-specific study parameters such as smaller, upper-division courses to form recommendations for future work.

Keywords: noncognitive factors, belonging, caring, underrepresented students.

Introduction

Studies with the aim of improving engineering students' academic achievement typically focus on cognitive factors, such as building study and learning skills, alternative learning aides for specific topics, or tutoring and coaching. However, a body of research has demonstrated the powerful impact of noncognitive and affective (NCA) factors – factors not related to innate talent or intelligence – on student performance [1]-[6]. These include, for example, motivation, engineering identity, and growth mindset. In this study, we examine whether a previously tested intervention [7] thought to improve two NCA factors can be adapted to a different learning environment and whether it impacts student performance across a variety of engineering courses.

Background and Motivation

This project seeks to adapt a previous study by Carrell and Kurlaender [7], to a different classroom setting and examine its impact on two noncognitive traits (perception of faculty caring and sense of belonging) and a broad set of academic outcomes. Carrell and Kurlaender [7]

sought to develop a low-effort intervention to address the lack of interaction between students and faculty on the presumption that other educators would be more likely to adapt such a "light touch" intervention. The intervention is theoretically grounded from extensive research on the role of student-faculty interactions and their importance to a host of college outcomes, including academic achievement, persistence and degree completion [8]-[11]. The cited studies theorize that such interactions help students socialize into college and gain validation from faculty on their ability to succeed [12], [13]. In addition, these interactions help reveal the unintended "hidden curriculum" [14]-[16] in college, such as learning how to effectively use faculty office hours and adopting key strategies for success in a course.

Carrell and Kurlaender [7] first conducted a series of exploratory qualitative studies with African-American and Latinx students. These two student populations have a lower persistence and graduation rate at the college of study, which is a four-year, broad-access university in the US West. Two findings emerged from these focus groups: Students reported that they were more successful in classes when they (a) felt more connected to their instructor and (b) understood class expectations. According to the researchers, these findings are consistent with social-psychological theory on how students' uncertainty about college can impact their success [4].

Based on these findings, the researchers developed what they considered a 'light touch' intervention consisting of a few individualized and strategically timed emails. Each email would convey three key messages: (1) knowledge of the student's current class performance, (2) keys to success in the class, and (3) reminder of available help (e.g., office hours or tutoring). The intervention was pilot-tested in a 420-person introductory economics course. Based on historic data, students who either performed poorly or missed the first (of five) required homework assignments scored about 10 percentage points lower in the final course grade. Thus, the intervention was triggered by poor performance in or missing this assignment. This resulted in 69 students who were randomly selected to either receive the first email (treatment; n = 35) or not (control; n = 34). Of these, an eventual 16 dropped the course, leaving a study sample of 53 students.

The first email acknowledged the treatment students' performance while emphasizing that recovery was possible. They were reminded of key behaviors necessary to be successful in the course, including attending lectures and discussion sections, completing the practice and required assignments, and attending office hours. A second email was sent to this treatment group after the first midterm examination, with feedback individualized based on their performance. Strong-performing students (better than B+ score) were commended on their performance and reminded of the professor's office hours. Middle-performing students (between C- and B) were informed that their grade in the course is likely to be strongly influenced by this midterm and that they still could improve their grade. They were told of key behaviors that will help them improve and additional resources available. Students that received less than a C- were told that they are at risk of failing the course, but also that they could recover and pass. They were further provided with behaviors needed to recover, additional help available, and the professor's office hours.

The results were highly encouraging. The treatment group scored 14 percentage points higher on the second midterm examination and achieved a final course grade that was 8 percentage points (approximately one-half letter-grade) higher than the control group. A detailed analysis of their

work showed that, compared to the control group, the treatment group scored 15 percentage points higher on homework, spent 2+ hours more on homework, and attended more office hours.

Based on the encouraging finding from the pilot test, the researchers designed a wide-scale test of the intervention. Based at a different large, broad-access, four-year university, they recruited 22 faculty members who collectively taught 20 different courses ranging from art to engineering with over 3,000 distinct students at all grade levels. Due to the logistical difficulties of standardizing the study across a diverse mix of instructors and courses, the previous trigger event was replaced with a random assignment of about 1500 students to be either in the treatment or control group. The large sample size allowed for stratifying the data for analysis across demographic groups. The treatment was also changed from two to three email contacts because of the longer academic term at the new test university (a 16-week semester rather than a 10-week quarter). The faculty were instructed to send the first email to the treatment students after the first "meaningful assignment" of the term.

Short-term outcomes found no significant effect in the overall treatment population in course grade, the percentage of total points earned in the class, the pass/fail rate or whether students earned a high grade (i.e., A or B) vs. low grade. For underrepresented minority (URM) students, however, there were differences. These students earned a better course grade, were 4.9 percentage points more likely to get an A or B in the course, were less likely to drop the course and less likely to "give up," which the researchers defined as earning less than 25% of the total available course points. More importantly, URM students in the treatment group experienced positive "spillover" effects, meaning their grades in courses other than the ones in the experiment were higher. This implied that the positive treatment effects not only improved the experiment course, but it did not come at the expense of a reallocation of effort from other courses. In fact, the treatment may have positively impacted the participants' other courses as well.

Long-term outcomes measured included college persistence, credit accumulation, and graduation at a defined time point of three years out. The treatment group overall showed better persistence one-semester later, but this effect faded between treatment and control groups beyond that. For URM students, however, the researchers found a larger and longer lasting persistence effect. The overall treatment group and the URM treatment group both had higher credit accumulation compared to the control group. Graduation showed no effect for overall students but a positive and significant effect for URM students. Importantly, the race or ethnicity of the instructor had no significant impact on any outcomes.

Because of the different findings between the pilot and wide-scale tests, the researchers repeated the pilot test and were able to replicate those results. This seems to suggest that it was difficult to duplicate the pilot test at scale, but that the intervention is robust and effective. This study sought to adapt the intervention for testing in the investigators' home department.

Methods

We implemented our own light touch intervention in four mechanical engineering courses (a total of 8 sections with 291 students) at a mid-size, teaching-focused public university in Spring 2025. During the first week of the 10-week quarter, instructors asked students to give informed

consent for the study and complete a pre-survey containing questions about demographic information, personality traits, and two NCA factors: perception of faculty caring and sense of belonging in the field of engineering. Students completed the same survey again at the end of the term.

To identify students eligible for participation in the study, instructors aimed to predict which students were at risk of performing poorly in the class. As such, the trigger event was chosen as poor performance on an early quiz or midterm exam. This took place during the third and fourth weeks of the quarter. Based on these criteria, there were 78 students enrolled in the study. These students were randomly assigned to either treatment (n = 41) or control (n = 37).

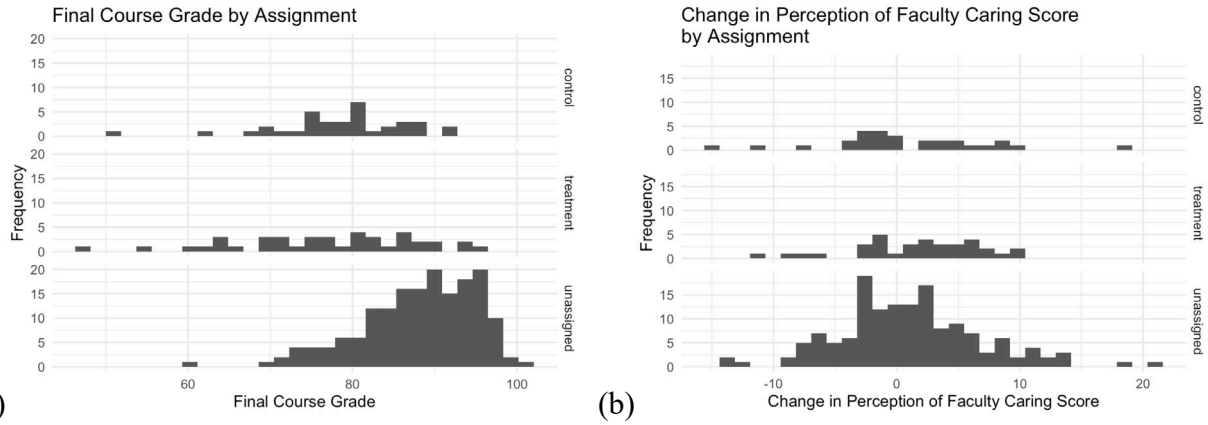
The first round of the intervention was rolled out following the trigger event. Treatment students received an email indicating their personal performance on the exam and outlining steps they could take to improve. A template email was personalized by each instructor to match their individual communication style. To avoid interference between treatment and control students (i.e., non-study or control students who did not receive an email from their instructor talking to treatment students who did), instructors also sent out a "placebo email" to all non-treatment students. This email only indicated the overall average grade on the exam. After the second midterm exam, in week eight or nine, another round of the intervention was rolled out, with students in the treatment group receiving a customized email depending on their midterm performance and the remaining students receiving another placebo. Sample emails are available in Appendix A. At the end of the quarter, pre- and post-survey and grade data were collected for all consenting students, including treatment, control, and non-study participants. To prepare for statistical analysis, all data was pre-processed (see details in Appendix B).

In the analysis, we were interested in the effect of the intervention on several different response variables, including course performance metrics like final course grade and final exam grade, as well as the NCA factors change in perception of faculty caring and sense of belonging. For each of these variables, we performed both descriptive and inferential statistics. First, we constructed side-by-side histograms to visualize any differences in outcomes between treatment students, control students, and non-study students. We also constructed a series of t-confidence intervals to estimate the effect of the intervention on each response variable. We constructed confidence intervals to estimate the treatment effect for all study participants, as well as for subgroups of interest, including each course, year of study, gender, and Pell Grant status.

Results

In our analysis, we compared outcomes between treatment students (n = 41), control students (n = 37), and non-study students (n = 171), finding few differences. For example, median final course grade was 77.69 for treatment students and 79.59 for control students (IQR of 16.33 and 8.69, respectively). Median change in perception of faculty caring was 2 points for treatment students, 0 points for control students, and 1 point for non-study students (IQR of 6.5, 7, and 7, respectively). We also visualized the distribution of each response variable across groups. Differences in the distributions would indicate a potential treatment effect. Figure 1 shows the distributions of (a) final course grade and (b) change in perception of faculty caring across these three groups. In neither plot do we see major differences in the distribution of response between

treatment and control. This indicates that the intervention did not have a noticeable effect on these outcomes. Similarity between treatment and control outcomes was consistent across all response variables we investigated (see supplemental results in Appendix C).



(a) (b)

Figure 1: The distributions of (a) final course grade and (b) change in perception of faculty caring for control (n = 37), treatment (n = 41), and non-study (n = 171) students. There are no major differences between histograms for control and treatment students, indicating the intervention did not have a noticeable effect on these outcomes.

When comparing the distribution of study students (both treatment and control) to the distribution of non-study students, we do notice a difference for final course grade (Figure 1(a)). In particular, final course grades are higher, on average, for non-study students (median of 79.4 for study students vs. median of 89.13 for non-study students – IQR of 11.62 and 9.50, respectively). This is as expected, however, as students were selected for the study if they were determined to be at risk of poor performance in the course. It appears that the study eligibility criteria did have some success in that the study enrolled students who needed more support.

To verify the descriptive findings, we also constructed 90% confidence intervals to estimate the effect of the intervention on each outcome of interest. Each confidence interval provides us with a range of plausible values for the difference in average response between the treatment group and the control group. A confidence interval that contains 0, therefore, indicates there is no significant difference in response between these two groups. In addition to considering the intervention effect for all students, we also looked for subgroup-specific effects (e.g., does the intervention have a different effect for White students than for URM students?). Estimates of the treatment effect on final course grade and perception of faculty caring can be found in Figure 2.

Overall, almost every confidence interval, for every response variable, and every subgroup, contains 0, indicating no significant treatment effect. Similar findings for the treatment effect on Midterm 2 score and change in students' sense of belonging can be found in Appendix C.

Discussion

In contrast to the previous study by Carrell and Kurlaender [7], we found that the treatment had no statistically significant effect on course performance (final exam score or final course grade)

or on NCA factors (perception of faculty caring or sense of belonging). These findings were consistent across all subgroups examined, including students from URM backgrounds.

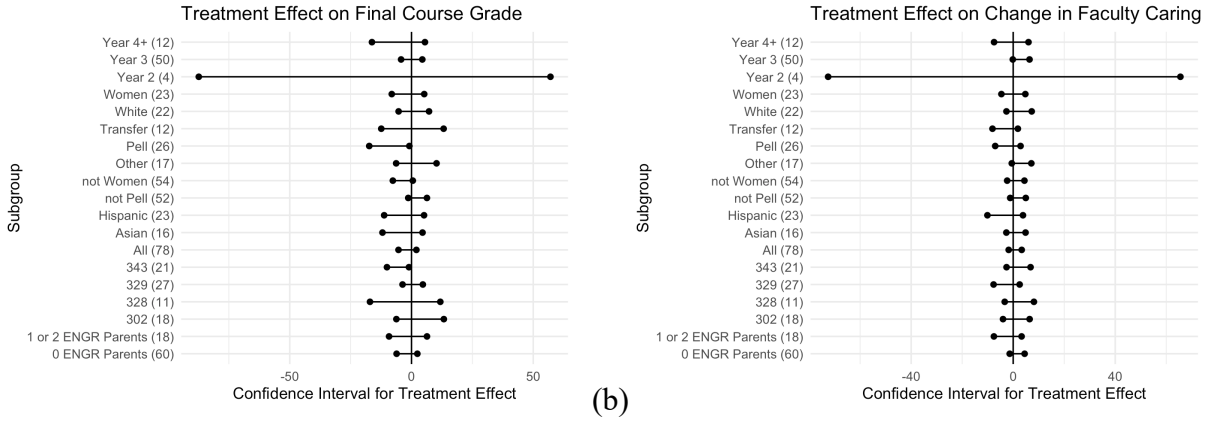


Figure 2: A series of 90% t-confidence intervals estimating the treatment effect on (a) final course grade and (b) change in perception of faculty caring for various student subgroups. Subgroup sample sizes are provided in parentheses. The interval labeled 'All' was constructed using all study students; the other intervals were constructed using the labeled subgroup (e.g., 'Women' or 'Transfer' students). The four subgroups labeled with a number represent the four individual courses in the study. Each plot contains a vertical line at 0. Intervals including 0 indicate that we cannot conclude a treatment effect on the response variable, intervals completely below 0 indicate the treatment group performed worse on the response variable than the control group, and intervals completely above 0 indicate the treatment group performed better on the response variable than the control group.

The null results for the NCA factors may be attributable to several differences between the classroom environment of the two studies. First, all courses in the present study were offered at an undergraduate teaching-focused university with class sizes typically capped at 35 students (as was the case for all classes in this study), while the prior work was conducted in large-enrollment courses (e.g., 420-student class for the pilot study). Smaller class sizes naturally facilitate more frequent and personalized student-faculty interactions. In fact, the instructors in this study already actively leveraged the smaller classroom environment to promote interactions with all students. The instructors provide further comments on this effect in the Reflections section. Second, all courses examined here were third- and fourth-year courses, while Carrell's and Kurlaender's [7] study included courses across all undergraduate levels. Upper-division students may have a more developed understanding of engineering and a more firmly established engineering identity. Together, these differences may have resulted in higher baseline NCA factors at the start of study, limiting the perceived change with a potential ceiling effect. Additionally, it is possible that the low-effort intervention has reduced efficacy in this context and does not meet the threshold necessary to produce a measurable impact, requiring a higher-intensity intervention. This, for example, could take the form of a higher frequency in personalized emails or one-on-one intervention meetings. However, this higher intervention intensity could increase the workload required from instructors for this 'light touch' intervention, thus future work parameterizing the relationship between baseline NCA factors, intervention intensity, and measured impact will be crucial for determining the necessary threshold.

The null results for academic measures may instead be related to the timing of the first trigger event. Our study took place in a quarter system, with the trigger event (a second quiz or first midterm) occurring in the third or fourth week of the term. In contrast, the pilot study by Carrell's and Kurlaender's [7] used the first homework assignment as a trigger event in a quarter system, while the larger scale study was conducted in a semester system. The later trigger event in our study may have reduced the time available for students to recover academically prior to the final exam. This limitation is further compounded by the cumulative nature of many engineering courses, where mastery of earlier material is necessary to successfully engage with subsequent content.

Despite the lack of statistically significant effects, many instructors reported positive qualitative responses from some students in the treatment group. These responses included expressions of appreciation for the email, acknowledgements of plans to improve performance, and increased engagement through attending office hours or asking questions in lecture. While these factors were not systematically tracked or quantified, they indicate that the intervention may have influenced student effort or help seeking behaviors, despite not resulting in a grade improvement. Additional instructor reflections are provided in the Reflections section.

Reflections

This section presents reflections written by the participating instructors based on their experience implementing the light touch intervention and considering the research findings.

[Instructor 1] This light touch intervention complemented my existing style of interacting with my students, which may have influenced the effectiveness of the treatment. For context, I make a point to learn and address all my students by name each quarter, and send follow-up messages in Canvas for students consistently missing homework submissions or scoring below certain thresholds on exams. However, these messages are not usually personalized because of the limitations of Canvas. When I took the time to address each student directly by name, the number of students replying to and acknowledging my messages increased, with a few following up in office hours.

[Instructor 2] I normally would not make a special effort to contact students who start off my class struggling because I had thought that, at this point in their academic career and their maturity, they would come and ask for my help. However, I now understand that maybe I need to reconsider this belief. Several of the students I contacted (in the treatment group) replied to my email and expressed sincere appreciation for my gesture. My effort in outreach to them was quite low, but the impact seemed to be meaningful. And although our data did not show significant differences between the treatment and control groups in all measures, the impact on the few students made me think that this was a worthwhile effort.

[Instructor 3] Prior to this study, I occasionally reached out to struggling students by writing a brief handwritten note on their returned midterm exams, encouraging them to come see me if they needed support. While this approach felt personal, it was informal and inconsistent. The light touch intervention provided a more organized and intentional way to personalize outreach to students who were at risk, while still requiring relatively low effort. Although the quantitative

results of our study did not show a statistically significant impact of the intervention, I observed a noticeable increase in the number of students attending office hours after the emails were sent. In addition, several of these students began asking questions more actively during office hours, suggesting increased comfort and engagement. These observations indicate that, even if the effects were not captured by our measured outcomes, the intervention had a meaningful impact on how some students interacted with me and sought academic support.

[Instructor 4] Before the study, I would typically reach out to struggling students using Canvas's automated messaging system. I would try to make these messages personalized; however, due to limitations in Canvas, I could not insert student names and other individualized details. Through this light touch intervention, I was able to increase the personalized details in the emails, while streamlining the process to keep the time required comparable to before. After sending the first email, a few treatment students replied with appreciation for my words of support and acknowledged planned changes to their study strategies based on my suggestions. For a couple of those students, I also observed an increase in course engagement through attending office hours and asking questions during lectures. These noticeable behavioral changes indicated to me that the personalized emails may have improved these students' comfort in approaching me, despite the statistical results for all treatment students showing no measured effect.

Future Work

The null results that conflict with Carrell's and Kurlaender's [7] studies indicate a need for further investigation into contextual factors that may diminish the effect of the study. Specifically, it will be beneficial to implement the study in a large-size, lower division course, where baseline levels of faculty interaction and academic confidence may differ substantially from this study and align closer to the prior work. Furthermore, we can adjust the intensity of the treatment to quantify its relationship with the baseline NCA factors. Through such a study, we can determine if factors such as academic year and/or class size affect the efficacy of the intervention. These additional studies can also further explore the experience of students from URM backgrounds by increasing sample size. Crucially, our study found no statistically significant result even for the URM subgroup, though our results are limited by the small sample size.

Additionally, the study can be further improved by expanding the variables collected and further standardizing the trigger event. For the data, a measure of student effort or course interaction could be collected by tracking office hour attendance or frequency of questions asked. This could help determine whether students are reacting to the intervention at all. For the trigger event, we could prioritize an earlier intervention by utilizing an introductory quiz (during week one or two) to ensure students have sufficient time to incorporate feedback.

Conclusion

In this study we investigated whether a previously successful light touch intervention—personalized, strategically timed faculty emails—could be adapted to small, upper-division mechanical engineering courses at a teaching-focused public university. In contrast to prior findings by Carrell and Kurlaender [7], we found no statistically significant effects of the

intervention on academic outcomes or on the targeted NCA factors of interest: students' perception of faculty caring and sense of belonging. These null results were consistent across student subgroups, including URM students, although the sample sizes of several subgroups were small.

The lack of measurable effects likely reflects important contextual differences between this study and prior work. Small class sizes, frequent instructor-student interactions, and more established engineering identities among upper-division students may have resulted in higher baseline levels of caring and belonging, limiting the potential impact of the intervention. Additionally, the compressed timeline of a 10-week quarter and later trigger events, compared to prior work, may have reduced opportunities for academic recovery.

Despite the null quantitative findings, instructor reflections suggest that personalized outreach was positively received by some students in treatment groups and may have encouraged seeking help in office hours and engagement in the classroom which are not captured by the study's outcome measures. These findings highlight the importance of instructional context when adapting student-success interventions and suggest several possible future studies to explore the transferability of the original intervention. These potential future works include examining the effect of classroom environment (such as small versus large class sizes and first-year versus upper-division courses) or the intervention intensity on NCA and academic outcomes, as well as broader measures of student engagement and support-seeking behaviors.

References

[1] Scheidt, M., Senkpeil, R., Chen, J., Godwin, A., & Berger, E. (2018). SAT does not spell SUCCESS: How non-cognitive factors can explain variance in the GPA of undergraduate engineering and computer science students. In *Proceedings of the Frontiers in Education Annual Conference* (pp. 1–5). San Jose, CA.

[2] Scheidt, M., Godwin, A., Berger, E., Chen, J., Self, B., Widmann, J., and Gates, A. (2021) "Engineering Students' Noncognitive and Affective Factors: Group Differences from Cluster Analysis," *J. of Engineering Education*, 110(2), pp. 343-370.

[3] Broda, M., Yun, J., Schneider, B., Yeager, D., Walton, G., & Diemer, M. (2018) "Reducing Inequality in Academic Success for Incoming College Students: A Randomized Trial of Growth Mindset and Belonging Interventions." *Journal of Research on Educational Effectiveness,* 11 (3): 317–38.

[4] Murphy, M. C., Gopalan, M., Carter, E.R., Emerson, K., Bottoms, B. & Walton, G. (2020) "A Customized Belonging Intervention Improves Retention of Socially Disadvantaged Students at a Broad-Access University." *Science Advances* 6 (29).

[5] Walton, G. M., & Cohen, G.L. (2011) "A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students." *Science* 331 (6023): 1447 –51.

[6] Yeager, D.S., Walton, G.M., Brady, S.T., Akcinar, E.N., Paunesku, D., Keane, L., Kamentz, et al. (2016) "Teaching a Lay Theory before College Narrows Achievement Gaps at Scale." *Proceedings of the National Academy of Sciences* 113 (24): E3341–E48.

[7] Carrell, S. E., & Kurlaender, M. (2023) My Professor Cares: Experimental Evidence on the Role of Faculty Engagement, *American Economic Journal: Economic Policy*, 15 (4): 113-41. doi: 10.1257/pol.20210699

[8] Tinto, Vincent. 1993. *Leaving College: Rethinking the Causes and Curses of Student Attrition*. 2nd ed. Chicago, IL: University of Chicago Press.

[9] Astin, Alexander W. 1993. *What Matters in College? Four Critical Years Revisited*. San Francisco, CA: Jossey-Bass.

[10] Astin, Alexander W. 1999. Student Involvement: A Developmental Theory for Higher Education. *Journal of College Student Development* 40 (5): 518–29.

[11] Kuh, George D., and Shouping Hu. 2001. "The Effects of Student-Faculty Interaction in the 1990s." *Review of Higher Education* 24 (3): 309–32.

[12] Rendón, L.I. 1994. "Validating Culturally Diverse Students: Toward a New Model of Learning and Student Development." *Innovative Higher Education* 19 (1): 33–51.

[13] Rendón, Laura I. 2002. "Community College Puente: A Validating Model of Education. *Educational Policy* 16 (4): 642–67.

[14] Winter, J. & Cotton, D. (2012) Making the hidden curriculum visible: sustainability literacy in higher education, Environmental Education Research, 18:6, 783-796, DOI: 10.1080/13504622.2012.670207

[15] Orón Semper, J.V., Blasco, M. (2018) Revealing the Hidden Curriculum in Higher Education. *Stud Philos Educ* **37**, 481–498. https://doi.org/10.1007/s11217-018-9608-5

[16] Koutsouris, G., Mountford-Zimdars, A. & Dingwall, K. (2021) The 'ideal' higher education student: understanding the hidden curriculum to enable institutional change, Research in Post-Compulsory Education, 26:2, 131-147, DOI: 10.1080/13596748.2021.1909921

Appendix A

Treatment Email Template: After Trigger Event
Below is the template email used, with each faculty slightly modifying the email to fit their own communication style and tone.

> Hello, XXX,
>
> The midterm was just completed, and your score was OO/PP. I suspect that you are disappointed with this result, and I believe you can do better and improve your grade in my class. Based on my experience, the keys to improving performance in this class is to [list 2-3 key strategies, such as diligently completing the homework, work in a small group with others on homework, or ask questions via email or office hours]. As a reminder, my office hours are: [List office hours]; please come and see me with any questions.
>
> Prof. WWWW

Treatment Email Template: After Second Major Assessment

> Hello, XXX,
>
> I reviewed the midterm results and noticed your score of OO/PPP. [Use option 1, 2, or 3 below depending on the student's performance]
>
> 1) You've done a great job at improving in this class! Please keep it up. As a reminder, the keys to success in this class are [repeat the key strategies you listed from the first email]. Also, here are my office hours; please come and see me with any questions: [List office hours].
>
> 2) You did pretty well, and I want to encourage you to keep working hard; you can still improve on the final exam. As a reminder, the keys to success in this class are [repeat the key strategies you listed from the first email]. Also, here are my office hours; please come and see me with any questions: [List office hours].
>
> 3) I know you can do better and I encourage you to not give up and keep working on improving. You'll get another chance on the Final Exam. As a reminder, the keys to success in this class are [repeat the key strategies you listed from the first email]. Also, here are my office hours; please come and see me with any questions: [List office hours].
>
> Prof. WWWW

Placebo Email Template

> Hello,

> We just completed the first of two midterms in the class, and I wanted to let you know the overall results. The mean score on it across all sections was YY/ZZZ. The next midterm will be on DATE or WEEK #.
>
> Prof. WWWW

Appendix B:

Data Processing

We made many data cleaning decisions while processing student grades and survey responses.

Race. In the pre-survey, students selected which racial and ethnic group(s) they identify with. The options included 'White', 'Asian', 'Hispanic, Latino, or Spanish Origin', 'Black or African American', 'Middle Eastern or North African', 'Native Hawaiian or Other Pacific Islander', and 'American Indian or Alaska Native'.

In our cleaning, the categories 'White', 'Asian', and 'Hispanic, Latino, or Spanish Origin' were maintained. There were insufficient numbers of students identifying with any other identity, or combination of identities. Therefore, all other races and ethnicities were collapsed into an 'Other' category.

Pell Grant. In the pre-survey, students were asked if they received a Pell Grant as part of their financial aid package. The response options provided were 'Yes', 'No', and 'Not sure'. For our analysis, responses were recoded to TRUE, indicating that a student did receive a Pell Grant, or FALSE, indicating that a student either did not receive a Pell Grant or was unsure.

Years. In the pre-survey, students were asked how many years (including the current year) they had been formally enrolled as a student at a college-level institution. The options were '1 Year', '2 Years', '3 Years', '4 Years', '5 Years', and '6 Years or more'. Separately, students were asked if they transferred to their current university.

Most transfer students enrolled in the study gave inconsistent answers to the Years question, reporting different numbers of years between the pre- and post-surveys. Therefore, the value for the Years variable for transfer students was replaced with "Transfer" while non-transfer students retained thar Years response.

Gender. In the pre-survey, students were asked to select which gender(s) they identify with. The options provided were 'Female', 'Male', and 'Non-binary'.

All students selected either only 'Female' or only 'Male', with the exception of two students who selected both 'Non-binary' and 'Male'. This was an insufficient number for a third gender category, so these students were categorized as 'Male' for the analysis.

Family Engineering Expertise. In the pre-survey, students were asked to select which area(s) of expertise ('Engineering', 'Computer Science', 'Other STEM', 'Non-STEM') apply to their

family members ('Parent/Guardian #1', 'Parent/Guardian #2', 'Siblings', 'Other Relative(s)'). Responses describing 'Siblings' and 'Other Relative(s)' were dropped from our analysis for lack of clarity and relevance. Responses describing parents/guardians were summarized into a single count (0, 1, or 2), indicating the numbers of parents/guardians with expertise in 'Engineering' and/or 'Computer Science'.

Personality Traits. In the pre-survey, students were asked to rate how well a series of personality statements described them. Students rated each statement on a scale from 1 (Very Inaccurately) to 7 (Very Accurately). These statements spanned five categories, with three statements per category.

Openness	Conscientiousness	Extroversion	Agreeableness	Neuroticism
I have a vivid imagination.	I leave my belongings around.	I don't talk a lot.	I sympathize with others' feelings.	I get upset easily.
I have excellent ideas.	I make a mess of things.	I keep in the background.	I have a kind heart.	I change my mood a lot.
I am full of ideas.	I often forget to put things back in their proper places.	I am quiet around strangers.	I feel others' emotions.	I have frequent mood swings.

The ratings for conscientiousness and extroversion were reversed. Student responses were then averaged to create five personality scores.

Change in Sense of Belonging. Students were asked to rate their agreement with each of the following statements regarding how much they feel they belong in engineering and the engineering community.

1. I feel comfortable in engineering.
2. I feel I belong in engineering.
3. I enjoy being in engineering.
4. I feel comfortable in my engineering classes.

The individual ratings, on a scale from 1 (Not at All) to 7 (Very Much So), were summed to create a belongingness score ranging from 4 to 28. The belongingness score from the pre-survey was subtracted from the post-survey score to quantify the change in sense of belonging.

Chage in Perception of Faculty Caring. Students were asked to rate their agreement with each of the following statements regarding their perception of faculty caring.

1. I feel comfortable asking a faculty member for help with a personal problem.
2. Faculty connect relevant topics to my major.
3. I feel that a faculty member would be sensitive to my difficulties if I shared them.
4. I see faculty members as role models.
5. I feel comfortable socializing with a faculty member outside of class.

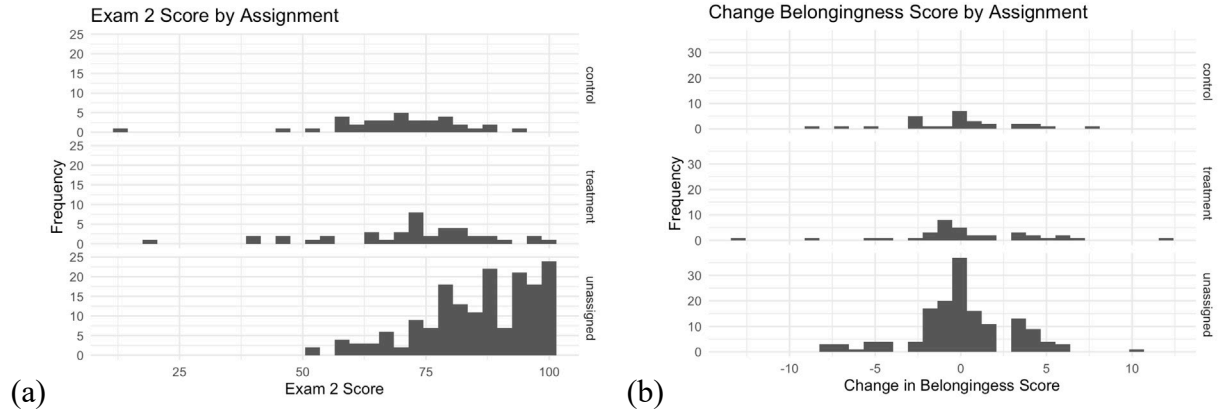
6. I feel comfortable asking a faculty member for help if I do not understand course-related material.
7. I feel comfortable seeking help from a faculty member before or after class.
8. I feel that a faculty member would take the time to talk to me if I needed help.

The individual ratings, on a scale from 1 (Not at All) to 7 (Very Much So), were summed to create a faculty caring score ranging from 8 to 56. The faculty caring score from the pre-survey was subtracted from the post-survey score to quantify the change in perception of faculty caring.

Appendix C:

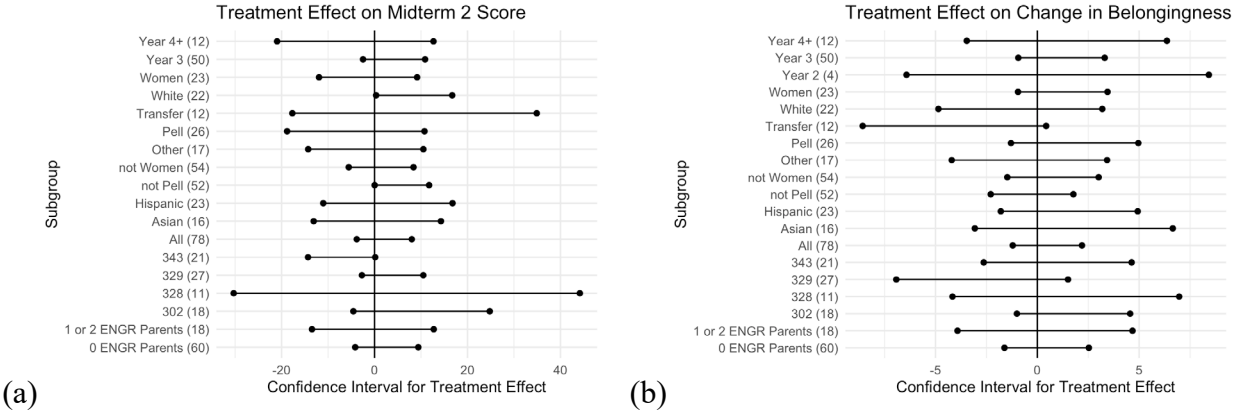
Additional Results

The distribution of (a) midterm 2 score and (b) change in sense of belonging, comparing outcomes for treatment students (n = 41), control students (n = 37), and non-study students (n = 171).



(a) (b)

A series of 90% confidence intervals estimating the treatment effect on (a) midterm 2 score and (b) change in sense of belonging. We estimated both the overall treatment effect and subgroup-specific effects.



(a) (b)