2026 ASEE
ANNUAL CONFERENCE
June 21–24, 2026 | Charlotte
Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #51887

Learning Outcomes for Calculus I Students Grouped by Uniformly Low Pre-calculus Scores

Dr. Diana D Morris, University of Virginia

Dr. Diana Morris is an assistant professor at the University of Virginia, and enjoys teaching a variety of math courses to undergraduates.

Dr. Hui Ma, University of Virginia

Dr. Hui Ma is an associate professor of applied math at the University of Virginia. She holds a Ph.D. in applied mathematics from the University of Alabama at Birmingham. Her current research interest focuses on mathematics education and STEM education, in particular student-focused instruction and learning-based grading.

Dr. Deepyaman Maiti, University of Virginia

Deepyaman Maiti is currently an Assistant Professor in the Department of Engineering and Society in the School of Engineering and Applied Science at the University of Virginia. He teaches Introduction to Engineering and Applied Mathematics classes, and se

Learning Outcomes for Calculus I Students Grouped by Uniformly Low Precalculus Scores

Abstract

This study examines the effects of grouping students with the lowest pre-calculus placement scores into one section of Calculus I. By comparing to similar students within other sections, we explored whether this grouping affected student performance academically in terms of increased precalculus proficiency throughout the semester, the number of calculus "targets" achieved, and the cumulative final exam. The results indicated almost no statistically significant difference in any of these outcomes. One exception was that students with the lowest preparation levels in the target group outperformed their peers from other sections in terms of calculus targets mastered by the end of the semester. Students in the target section were also much likelier to indicate satisfaction with their calculus placement, although this may be linked to their initial selection for the intervention group.

Keywords: homogeneous grouping, Calculus, academic readiness, pre-calculus background

Introduction

Many students in introductory calculus courses struggle with background concepts, including algebra, knowledge of functions, and limits. Moreover, while some high schools routinely channel college-bound students into calculus, others do not even offer the subject [1]. All of this can lead to a freshman engineering class with widely varying preparation levels [2].

Grouping by math performance, or "tracking", at the K-12 level has often been viewed unfavorably, owing to possible stigma associated with being in a lower set and concern about locking students out of college-bound trajectories. [3] In the college setting, the sorting of students by background generally occurs through their placement into remedial or lower-level math classes, where concerns about stigma and lack of progression to higher-level courses persist [2]. One study found that the prerequisite math classes designed to prepare students for calculus are so unsuccessful at that goal that simply placing students directly into calculus appeared to be the single most effective strategy for getting them through the subject [4], [5]. Students often enter remedial courses with the goal of entering calculus but never emerge on the other side [2].

The research into grouping college students by prior knowledge into different sections of a single course is limited, with one study of a statistics course showing no difference in overall performance [6]. One example of an effective alternative from the computer science field involved splitting students by programming experience into three courses that nevertheless all provide the same credit and meet the same prerequisites, effectively not disadvantaging students who enter without any programming background. [7]

Our research differs from others in that the students targeted for intervention took the same class as others but were grouped together in a single section. In other words, students with a limited background were not funneled into a remedial course or even a different course serving the same function as a more advanced one, but exactly the same course and our work investigates how, if at all, their performance differed from similarly-prepared peers not in a homogeneous group.

Context

This IRB-approved study involves first year students in an engineering school at a major research university that does not offer math courses below Calculus I. Enrollment in this first course thus includes a wide variation in background, from students with almost no pre-calculus to students with at least one year of high school calculus. Observationally, students with weaker precalculus skills can feel left behind in lectures and group work, leading them to participate less and ask fewer questions than their more experienced peers. Mastery grading was introduced in Fall 2022, in part to support those with a weaker background and to address student anxiety. Through several iterations of this grading scheme in Fall 2023 and 2024, we found improvements in student grades and reductions in anxiety although these benefits did not necessarily persist in future courses [8].

In Fall 2025, Calculus I became a required course for entering engineering students; prior to this, students could elect to begin the sequence in Calculus II regardless of official credit. In other words, AP or transfer credit in Calculus I became required for entry into Calculus II. Students who did not have this credit, even with a strong high school course in the subject, were required to take Calculus I. Before the change in Fall 2024, just 169 out of 716 entering first year engineering students (23.6%) entered the sequence in Calculus I. Following the change in 2025, enrollment in Calculus I rose to 228 of 725 first year engineering students (31.4%) and the number of sections offered increased from five to seven, taught by five different instructors.

Given the increased proportion of students with prior calculus experience, this Calculus I cohort exhibited an even wider variation in student background than before. Out of concern for class dynamics, the decision was made to create a special section of the course reserved for students with the most limited prior exposure to the underlying calculus concepts. Our motivation was two-fold:

1. With a class composed uniformly of students in need of review, the lecture pace could be slowed and focused attention could be placed on background material without frustrating students who did not need such review.
2. Students with very little calculus background would not be as likely to end up surrounded by those who had already taken the course in its entirety.

Selecting students for the target section

To select students for the target section of Calculus, scores on the pre-calculus portion of a summer placement test were used. This placement test consisted of 30 multiple choice questions in each of pre-calculus, Calculus I and Calculus II and was required of all incoming engineering students. It was one of several factors (including AP/transfer credit, high school courses, and student's self-assessment) that determined where a student entered the calculus sequence.

The 36 students with the lowest pre-calculus scores were placed into a single Calculus I section. The remaining 192 students were placed randomly in six other sections of the course, taught by four other instructors. Enrollment in the other sections ranged between 27 and 35 students each. The target section was indistinguishable from the other six sections in course listings, and students were not informed of any differences.

Importantly, the test on which selection was based was not proctored. Although we could be confident that low scores corresponded to a weaker level of background knowledge (or at least a weaker level of recall), there is no guarantee that higher scores accurately predicted a strong background, given that students could in theory look up the answers. In other words, one limitation of this work is that the method for selecting students could not guarantee that the students most in need of intervention were in the target section.

Class Structure

Each section met for four fifty-minute classes per week and used a mixture of lectures and daily worksheets. Assessments were bi-weekly with one built-in review day before each.

Grading Scheme

All sections used mastery grading, with the course divided into 20 learning targets (LT's), 10 of them "core" targets. LT's included such topics as interpreting the meaning of a derivative, related rates, and the Fundamental Theorem of Calculus. A student "mastered" a given target by demonstrating completely correct work on the corresponding problem(s) on one of the bi-weekly in-class assessments. Students had three opportunities to master each learning target. The grade was based on the number of learning targets mastered (55%, with core targets weighted slightly more), a 30-question multiple choice pre-calculus assessment given 3 times during the semester (10%), pre-class and in-class worksheets and online homework completion (20%), and a comprehensive final exam (15%).

Students scoring 80% on the pre-calculus assessment received full credit for that category, and were not required to take it again after reaching the 80% threshold.

PreCalculus Testing

Although pre-calculus testing was introduced as a component of the Calculus I grade in Fall 2024, a new set of tests was developed in Fall 2025 with more emphasis on the skills needed for

Calculus I, in particular algebraic simplifications. The test questions fell into four categories: algebra, functions, trigonometry, and exponentials/logs. Rather than a single pre-calculus unit at the beginning, topics were spread throughout the semester "just-in-time" for their use before the corresponding calculus topics.

Targeted review materials were provided for each of the four topics and students reported that they found these materials useful. The pre-calculus component of the course served not only to help students plug gaps in their high school preparation, but because the first assessment was given on the third day of class, it served as a **baseline measurement of student preparation.** Because the initial test, on which placement in the target section was determined, was not proctored, we relied on this first pre-calculus test for a more accurate measure of student preparation.

Class Structure for the Target Section

Students in the target section took the same (mastery graded) assessments, precalculus assessments, online homework, and final exam as every other section, and on the same schedule. Class instruction in this special section was similar to that of other sections, with a few differences. Lecture notes and in-class worksheets were combined into one document graded daily, to reduce complexity and encourage students to view taking class notes as essential. Lectures were shorter and explicitly addressed the basic algebra steps needed to solve problems rather than those steps being assumed. Students worked in groups for most of class time with two in-class TA's and the instructor circulating.

Although the structure of the special section had minor differences from the main group, the intervention being evaluated in this paper is the grouping of uniformly under-prepared students into one section and our interest is how, if at all, being in this environment affected these students academically.

Research Questions

1) How do students with low pre-calculus placement scores grouped together in a Calculus I section perform on a pre-Calculus knowledge test with multiple attempts throughout the semester, compared to similar students in other sections where students of various backgrounds are mixed?
2) How do the students in this special section perform on the regular Calculus assessments given throughout the course, compared to similar students in other sections?
3) How do these at-risk students perform compared to similar students in other sections in terms of final exam performance?
4) How, if at all, do these students' perceptions of the course vary from similar students in other sections?

Data Collection and Analysis

To address Research Question 1 (RQ1), we divided students in the main group and the target section into four subgroups based on their scores on the first round of the precalculus assessment. We then compared averages, total improvement and overall pass rate within each group from the target and control sections using two-sided t-tests where normality assumptions were appropriate (as measured by the Shapiro-Wilk test). Where they were not, we looked for different means between the groups using the (non-parametric) permutation test. Because of the small size of the samples, we used the Fisher exact test to look for significant differences between the proportions of students who passed in each group.

For Research Question 2 (RQ2), we compared the total number of learning targets mastered and the number of attempts required to pass, again by subgroup within the target and control groups. Again, we used a two-sided t-test where appropriate and a permutation test where the Shapiro-Wilk test determined they were not.

For Research Question 3 (RQ3), we compared the final exam scores overall and by group. As before, because the normality assumptions were not met for the smaller subgroups, we used a mixture of two-sided t-tests and permutation tests to assess whether there was any differences between comparison groups.

For Research Question 4 (RQ4), we used data from an end-of-semester survey, in which students indicated whether they felt that their calculus placement was correct (i.e. if their Calculus I course was above their level, at their level, or below their level). Student responses from the target section were compared to students in other sections. Because of the presentation of the data, we were not able to split into comparison groups.

Results

Research Question 1: Precalculus Performance

Overall Performance on Precalculus Assessment 1

We viewed the results of the first precalculus assessment (PC1), given on the third day of class, as a baseline for students' background knowledge. As expected, the target section had a lower average (12.25) on PC1 than did the other sections (14.64) on this 30-point test. Using Welch's t-test to compare sample means, this revealed a statistically significant difference ($p = 0.025$). Moreover, 13 students in the main group (around 7% of the total) reached the 80% threshold (score = 24) to pass the test on this first attempt. None of the students in the target section reached it.

Establishing comparison groups

The results of PC1 were also used to find students in other sections similar to those in the target section. Students were divided into four "comparison groups" within the main group and target section, determined by score. The bins for each group were as follows: scores 1-7 (A), scores 8-13 (B), scores 14-18 (C) and scores 19-22 (D). These ranges were carefully chosen because the means of the target and control groups were similar. Because no students in the target section scored higher than 22, the 16 students in the main group whose initial scores exceeded 22 were excluded from the analysis as not comparable.

The analysis thus proceeds with all 36 students from the target section and 172 students in other sections who both took PC1 and scored less than 23. First, we investigated whether the groups were in fact comparable by testing for significantly different means between the groups (e.g. Main Group A and Target Group A). Because several of the samples were small and normality assumptions as determined by the Shapiro-Wilk test were not met, we used a permutation test, which indicated that the means between comparison groups were plausibly the same.

Comparison Groups					
Group based on PC1 Score	**Main n (%)**	**Target n (%)**	**Average (Main)**	**Average (Target)**	**p-values**
1-7 (Group A)	14 (8.2%)	9 (25%)	5.79	4.78	p=0.184
8-13 (Group B)	73 (42.4%)	11 (30.6%)	10.45	10.36	p=0.9241
14-18 (Group C)	50 (29.1%)	10 (27.8%)	16.3	16.14	p=0.8100
19-22 (Group D)	35 (20.3%)	6 (16.7%)	20.34	20	p=0.5941

Table 1. Comparison Groups Based on the First Precalculus Assessment (PC1)

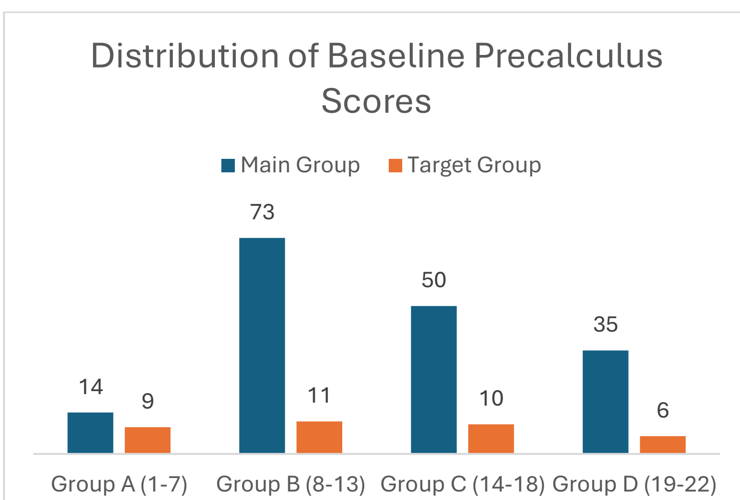

Figure 1. Counts within each comparison group

It is of note that the students in the target section are overrepresented in the very lowest group (Group A) with 25% vs. 8.2% from the main group. This reflects the selection process, which although not perfect, seems to have succeeded in selecting students with a weaker background.

Total Improvement

Students were given two further opportunities to take the precalculus test with two comparable versions of the test (PC2 and PC3) given in-class in October and December. To assess students' overall improvement in precalculus, we measured the absolute difference between their highest score of the three assessments and their score on PC1, as seen in Table 2. These numbers require some caution in that students who reached 24 out of 30 points on the second round did not take the test again (hence the improvements do not necessarily represent their true gains in the class) and students with high initial scores had less "room" for improvement.

Given that the test was 30 points total, the numbers indicate that on average, students in every subgroup improved their precalculus over the semester. Using a paired t-test for the initial and highest performance on the Precalculus assessment in subgroups with n greater than 30 and the Wilcoxon Test for subgroups with n less than 30, we see that *all subgroups made significant gains*.

Total Improvement on Precalculus Assessment					
Category	**Comparison Group**	**n**	**test**	**mean difference**	**p-value**
Main	**A**	14	Wilcoxon	8.571	**0.0016**
	B	73	paired t	9.056	**<2.2e-16**
	C	50	paired t	6.7	**<2.2e-16**
	D	35	paired t	4.857	**8.31e-13**
Target	**A**	9	Wilcoxon	9	**0.0039**
	B	11	Wilcoxon	8.2	**0.009**
	C	10	Wilcoxon	7.1	**0.0059**
	D	6	Wilcoxon	4.5	**0.035**

Table 2. Total Improvement on Precalculus Assessments by Group

Table 3 shows improvement on the precalculus test by comparison group. As all subgroups A, B, and D had roughly normal distributions, a t-test with unequal variance was used to compare. Subgroups C were compared for similar means using a permutation test. *No differences in means between paired groups were significant.* Similar students in the target group and in the main group seemed to improve in similar ways.

Average Precalculus Improvement by Group			
Group	**Main**	**Target**	**p-value**
A	8.57	9	0.8604
B	9.06	8.2	0.6083
C	6.7	7.1	0.7388
D	4.86	4.5	0.6008

Table 3. Average Total Improvement by Comparison Group

Pass Rate

Overall, 46.3% of students in the main group passed the precalculus test by achieving at least 80% on one of their attempts. If we exclude the 16 students who scored too high initially to be within any comparison group, however, this reduces to 41.3%. Within the target group, 36.1% passed. The results are broken down by comparison category in Table 3. Within the groups that were strongest on PC1, the pass rates are very similar.

Because of the small sample sizes, Fisher's exact test was used to compare the proportions between groups: the p-values exceeded 0.6 in all cases; *no differences were statistically significant*.

Precalculus Test Pass Rates by PC1 Scores

	Main	Target
A	1 of 14 (7.1%)	0 of 9 (0.0%)
B	14 of 73 (19.2%)	3 of 11 (27.3%)
C	28 of 50 (56.0%)	6 of 10 (60%)
D	28 of 35 (80%)	5 of 6 (83.3%)

Table 4. Pass Rates by Comparison Group

Research Question 2 (RQ2): Calculus I assessments

Overview

The majority of the overall course grade consisted of passing (or "mastering") a certain number of learning targets (LT's) out of the possible 20. Students had three opportunities to pass each target. In order to see whether students in the target group performed differently on Calculus I topics, we compared how many LT's they passed overall along with their average number of attempts both as a whole and compared to similar students in the other sections.

Total Targets Passed

Comparing the target section to all others revealed a slightly lower total number of LT's passed by the intervention group, but it was not statistically significant. *On this measure, the target section performed as well as the main group on the whole.*

Total Learning Targets (out of 20)		
Main	**Target**	**p-value**
17.1	15.9	p=0.09

Table 4. Total Learning Targets

The next table shows the total number of targets passed within the target section and the main class, again split by comparison group. The two groups that met normality assumptions as determined by the Shapiro-Wilk test were compared using a t-test and those that did not were compared using a permutation test. *Group A (the lowest performers) passed significantly more learning targets on average in the target group than in the main group*. The differences between the other groups were not statistically significant.

Average Learning Targets (out of 20) by Comparison Groups				
Group	**Main**	**Target**	**Test**	**p-value**
A	8.46	13.67	t-test	**0.0214**
B	15.28	14.55	permutation	0.6255
C	18.6	17.6	permutation	0.1195
D	19.37	19.17	t-test	0.7758

Table 5. Average Learning Targets by Comparison Group

As an alternate visualization, the chart below shows the similarities between groups.

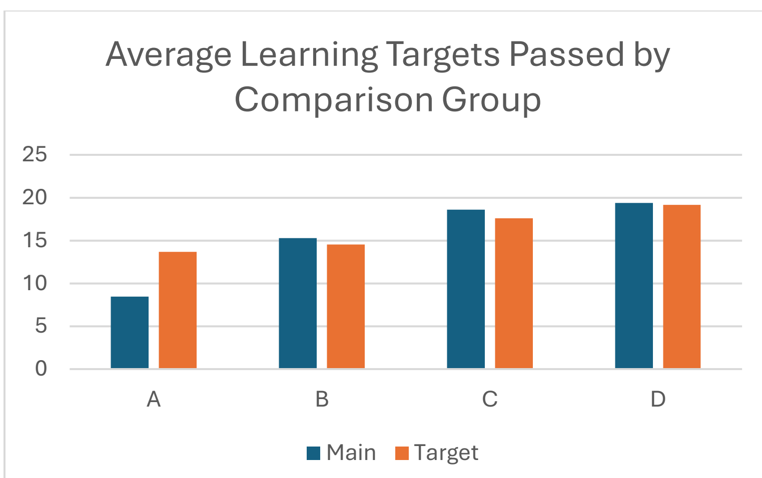

Figure 2. Average Number of LT's by Comparison Group

Total Number of Attempts

Here we compared the average number of attempts to pass a learning target between groups, i.e. the number of attempts required for each LT successfully attempted. The number could range between 1 and 3: for a student who mastered each LT on the first attempt, this would be 1 and for a student who required all three attempts for each LT passed, it would be 3.

By group, the results are summarized in the following table. We see that students within all groups in the target section required on average fewer attempts than comparable students in the other sections but *the differences were not statistically significant*.

Average Number of Attempts by Comparison Group				
Group	**Main**	**Target**	**Test**	**p-value**
A	2.18	1.97	t-test	0.2974
B	1.82	1.64	permutation	0.2877
C	1.48	1.43	permutation	0.6652
D	1.29	1.28	t-test	0.9757

Table 6. Average Number of Attempts per Learning Target

Research Question 3 (RQ3): Final Exam Performance

Here we analyze final exam scores from the 169 students from the main class and the 35 students from the target section who took both PC1 and the final exam. *Overall, students in the target section performed significantly worse on average than students in the main group*. Given their starting point, this is unsurprising. Welch's t-test to detect differences in population means yielded a p-value of 0.020.

Final Exam Scores (out of 150)		
Main	**Target**	**p-value**
123.8	111.5	p=0.0203

Table 7. Overall Final Exam Scores

Testing the eight comparison groups for normality using the Shapiro Welk test, we ran either the t-test if normality assumptions were met or the permutation test if not. Test results are given in the table.

Although the target group performed better in the lowest category, this difference was not significant. *Students in Group C (the second highest performing on PC1) performed significantly worse on the final exam*.

Final Exam Score by Comparison Groups				
Group	**Main**	**Target**	**test**	**p-value**
A	76.3	90.36	t-test	0.2939
B	117.06	110.65	permutation	0.394
C	132.9	118.5	permutation	**9e-04**
D	136.3	133.04	t-test	0.3445

Table 8. Final Exam Scores by Comparison Group

Research Question 4 (RQ4): Student Attitudes

We were not successful in collecting much useable data on this question due to survey design flaws and data-matching challenges. However, in an end-of-semester student survey, we found that 77% of respondents in the target section felt that they had been placed at the correct calculus level vs. 58% of students in the main group. At least some of that difference might be attributed to the fact that more students in the intervention group had limited calculus background and were clear candidates for Calculus I. Because the survey data was not matched to class data, however, we were not able to break the responses into comparison groups.

Conclusions and Further Work

As expected, the overall performance of the target section was worse than the overall performance of the main group on average by every measure except mastery of Calculus topics. This is likely linked to their selection as low-scorers on the precalculus placement test. However, when compared to students with similar scores on the baseline precalculus test, there was almost no difference in performance between any subgroups. The lowest group in the target section passed significantly more calculus targets than their peers in other sections, and one of the higher performing groups in the target section performed worse on the final exam.

While the results of this study seemed to show little difference in outcomes, we found that in itself a worthwhile piece of information. It allayed some of our fears that sorting students in this way could be harmful. Grouping by background seemed not to disadvantage the less-prepared students who were selected for the target section. It also did not disadvantage the relatively few better-prepared students who ended up in the section as well, by our imperfect selection process. Perhaps if the special section had not been tied so carefully to the schedule of the main class, more time could have been devoted to review and a slower progression of topics that would have benefited the lower performing groups within the special section.

Possibly the most important question for future work is that of student attitudes. Were students in the target section less intimidated since they were in a more homogeneous group? Or would these students have felt more motivated if surrounded by students with more experience in calculus? Carefully constructed surveys would be instrumental in this regard in future studies.

In one final note, we find it interesting that had students been required to pass the precalculus assessment in order to enter calculus, over half of all our students, including many who passed Calculus I comfortably, would never have entered.

References

[1] P. Sadler and G. Sonnert, “The Path to College Calculus: The Impact of High School Mathematics Coursework,” *Journal for Research in Mathematics Education*, vol. 49, no. 3, p. 292, 2018, doi: https://doi.org/10.5951/jresematheduc.49.3.0292.

[2] M.A. Rios and P. Burdman. “Staying the Course: Examining College Students' Paths to Calculus”. 2023 *Just Equations*.

[3] L. M. Mulkey, S. Catsambis, L. C. Steelman, and R. L. Crain, “The long-term effects of ability grouping in mathematics: A national investigation,” *Social Psychology of Education*, vol. 8, no. 2, pp. 137–177, Jun. 2005, doi: https://doi.org/10.1007/s11218-005-4014-6.

[4] MAA Communications, “Calculus as Preparation for Calculus – Mathematical Association of America,” *Maa.org*, Jul. 16, 2024. https://maa.org/math-values/calculus-as-preparation-4-calculus/

[5] California Community Colleges in partnership with The RP Group. *Preparatory Pathways and STEM Calculus Completion: Implication of the AB 1705 Standards.* 2024. https://rpgroup.org/Portals/0/Documents/Projects/MultipleMeasures/AB705_Workshops/PreparatoryPathwaysSTEMCalcCompletion_February2024.pdf

[6] D. Turner, J. Pleuss, and C. Collins, “The Impact of Academically Homogeneous Classrooms for Undergraduate Statistics,” *PRIMUS*, pp. 1–11, Jan. 2020, doi: https://doi.org/10.1080/10511970.2019.1710629.

[7] J. P. Cohoon, J. M. Cohoon, and M. L. Soffa, “Educating Diverse Computing Students at the University of Virginia,” *Computer*, vol. 46, no. 3, pp. 52–55, Mar. 2013, doi: https://doi.org/10.1109/mc.2013.39.

[8] H. Ma, D. D. Morris, and L. Wheeler, “Mastery Grading in Calculus: effects on performance and perception across demographics,” *Asee.org*, Jun. 22, 2025. https://peer.asee.org/56957