

How do introductory physics and mathematics courses impact engineering students' performance in subsequent engineering courses?

Kyle M. Whitcomb,¹ Z. Yasemin Kalender,¹ Timothy Nokes-Malach,² Christian Schunn,² and Chandralekha Singh¹

¹*Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, 15260*

²*Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, 15260*

In collegiate engineering curricula in the US, physics and mathematics are treated as foundational with all students taking physics and mathematics in both semesters of freshman year and additional mathematics courses in later semesters. Using academic data from the cohorts of students in introductory physics since 2009, we investigated the correlation between the performance of undergraduate engineering majors in introductory physics and mathematics courses and their performance in subsequent engineering courses. We find an interesting relationship between the best predictors of performance, advanced mathematics courses, and the physics sequence.

I. INTRODUCTION AND GOALS

In a well-designed curriculum, different courses build effectively on each other so that student learning can be maximized. It is of particular importance to an interdisciplinary field such as engineering to strive toward improving the structure of their curriculum [1–4]. Educational data mining and learning analytics can help inform these curricular changes.

Some engineering schools are considering how later courses are impacted by foundational courses taken in the first year which include physics and mathematics courses. For example, some departments within our engineering school are trying to take a closer look at their curricula for possible changes and were interested in the impact of the foundational courses on later engineering courses in order to suggest ways to improve the structure of their curricula. Here we focus on using “big” institutional data to investigate how physics and mathematics courses in and beyond the first year impact performance in subsequent engineering courses.

Sadler and Tai, using a linear regression analysis on the years of instruction (YOI) in STEM subjects, found that student performance in target college science courses was best predicted by the YOI of mathematics in high school followed by the YOI of the target subject in high school, while there was no statistically significant cross-disciplinary correlation [5]. We use “big” institutional data on 5546 engineering students since Fall 2009 to investigate the correlation between grades earned in the foundational courses and in advanced engineering courses controlling for SAT Math scores and high school grade point average. Comparing the correlations of different foundational courses with later courses can provide guidelines for structuring engineering curricula.

Similar to many other engineering programs, the School of Engineering at the University of Pittsburgh (Pitt) requires all engineering students to take a common first-year curriculum followed by curricula set by individual departments. The common first-year engineering curriculum includes two semesters each of introductory engineering, physics, chemistry, and calculus since these courses are perceived to be foundational for all engineering disciplines. The physics, mathematics, and chemistry courses follow the material typical of these courses across the nation while the introductory engineering sequence includes programming in C++ and

TABLE I. The relevant portions of the first two years in the Mechanical Engineering and Materials Science (MEMS) curriculum. The target advanced engineering courses are bolded.

Term	Full Course Name	Short Name
1	General Chemistry for Engineers 1	Chem 1
	Intro to Engineering Analysis	Engr 1
	Basic Physics for Science and Engineering 1	Phys 1
	Analytic Geometry and Calculus 1	Calc 1
2	General Chemistry for Engineers 2	Chem 2
	Intro to Engineering Computing	Engr 2
	Basic Physics for Science and Engineering 2	Phys 2
	Analytic Geometry and Calculus 2	Calc 2
3	Analytic Geometry and Calculus 3	Calc 3
	Intro to Matrices and Linear Algebra	Lin Alg
	Materials Structure and Properties	Mat. Structure
	Statics and Mechanics of Materials 1	Statics 1
4	Differential Equations	Diff Eq
	Statics and Mechanics of Materials 2	Statics 2

MATLAB in the context of solving engineering problems. Beyond that, each department sets its own curriculum which often includes advanced mathematics courses. In this paper, we focus on the curriculum of the Department of Mechanical Engineering and Materials Science (MEMS). Table I shows the relevant courses from this curriculum along with shortened names used in this paper.

It is important to examine the relationships between courses to ensure that they build effectively on each other. In this study, we investigate the following research questions:

- R.1. To what extent does student performance in foundational courses (in physics, mathematics, and chemistry) predict performance in subsequent engineering courses?
- R.2. Are these predictive relationships linear or non-linear?
- R.3. Which foundational courses are the best predictors of course achievement in target advanced engineering courses according to a linear regression analysis?

II. METHODOLOGY

The university provided de-identified data on engineering students from Fall 2009 through Fall 2017. The data includes demographic information, high school GPA, SAT scores, ACT scores, Advanced Placement (AP) scores, and course grade information while at the university. Grade points G earned in all courses by a student were converted to a standardized “z-score” z using the mean μ and standard deviation σ of grade points in that particular class by computing $z = (G - \mu)/\sigma$, which is in units of standard deviation [6]. For students who did not have a course grade on record but had a passing AP test, their AP score was used in lieu of grade points. AP scores were standardized using the same formula.

We used a linear regression analysis to investigate the relationships of the courses in the engineering curricula. From the results of this linear regression, we report the standardized β coefficients along with their p -values, N , and R^2 [6]. The β coefficients measure the standard deviations gained in the dependent variable (the target course) per standard deviation above the mean in the independent variables (the predictor courses) [6]. To allow for direct comparisons, all β values are normalized [6]. The p -values are a measure of statistical significance, and we use the standard requirement of $p < 0.05$ [6]. R^2 is the fraction of variance in the dependent variable explained by the regression model [6].

For each course in the model, we constructed a subset of

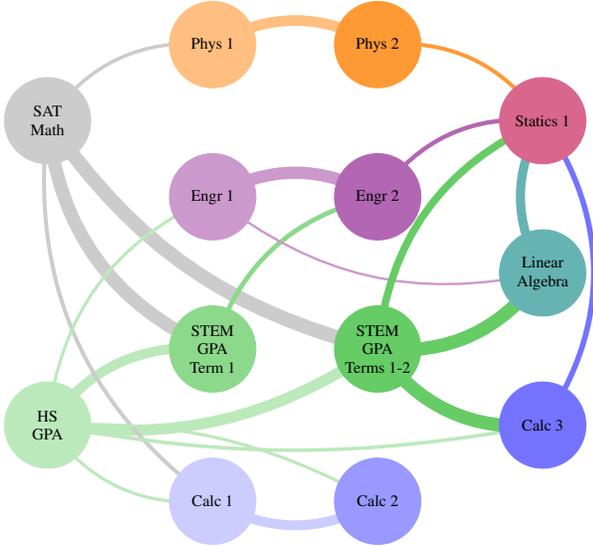


FIG. 1. A visual representation of all statistically significant predictive relationships leading to Statics 1. Chemistry is not shown since it does not predict any courses beyond the first year when STEM GPA is included. Courses are organized left to right according to the chronology of the MEMS curriculum. Vertical positions are chosen for aesthetic purposes. Colors are chosen to group courses in a sequence. Line thicknesses are scaled directly by β . Lines of the same color as a course node indicate that course is a statistically significant predictor of the recipient of the line. For visual clarity, concurrent lines are omitted except those leading to Statics 1.

the data informed by the curricula and included as independent variables, or “regressors,” all the first-year courses and advanced mathematics predictor courses that were taken prior to or concurrent with the target course as well as high school GPA, SAT Math scores, and cumulative STEM GPA at Pitt through the semester prior to the target course. We enforced the curriculum’s chronology within the data, only allowing predictor courses taken prior to or concurrent with the target course by each student (i.e., a student is dropped from an analysis if they never took the target course or if they took a predictor course after the target course). A linear regression was run with the z-scored grade points of the target course as the dependent variable and the z-scored grade points of the foundational courses, high school GPA, cumulative STEM GPA through the prior semester, and SAT Math scores as the independent variables. The independent variable with the largest $p > 0.05$ was dropped, then the linear regression re-run with the remaining independent variables [6]. This process continued iteratively until all remaining independent variables had $p < 0.05$. Since some significant predictors may have been dropped erroneously due to interference with poor predictors, the dropped regressors were then added back to the list individually to ensure they still had $p > 0.05$.

III. RESULTS AND DISCUSSION

Here we report the results for the three advanced engineering courses (Mat. Structure, Statics 1, and Statics 2) from Table I. A visual representation of the chronology and final linear regression results for foundational courses leading up to Statics 1 is shown in Fig. 1. In Fig. 1, Linear Algebra and Calculus 3, which are taken concurrently with Statics 1, produce the thickest lines beyond STEM GPA leading to Statics 1, corresponding to the largest β coefficients. Performance in both Linear Algebra and Calculus 3 is predicted almost solely by STEM GPA in the previous two terms, with Calculus 3 also weakly predicted by high school GPA and Linear Algebra weakly predicted by Engineering 1, likely due to a common emphasis on MATLAB in both courses at Pitt. Overall STEM GPA overshadows even the previous courses in the calculus sequence in predicting performance in advanced mathematics courses, which strengthens the importance of both the introductory engineering and physics sequences predicting performance in Statics 1.

In order to investigate correlations between predictor and target courses, we binned students by grades earned in a predictor course and then found the percentage of students in each bin that go on to earn each letter grade in the target course. Figure 2 shows such an analysis with the target course Statics 2 and the predictor course Linear Algebra. The observed correlation between grades earned in Linear Algebra and those earned in Statics 2 is also typical of the other foundational courses in the curriculum. The linearity of the relationship can be seen in this figure by noting the downward trend from left to right in the number of Bs, Cs, and lower grades in contrast to the upward trend for As. This linearity is clearer when we plot the mean grade points earned by Stat-

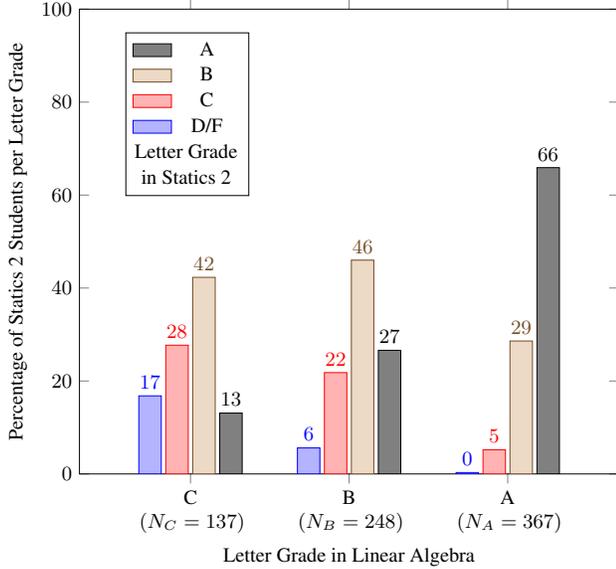


FIG. 2. Students are binned by their letter grade in Linear Algebra with the total number of students in each of these three groups shown below the letter. The percentage of each group that went on to earn an A, B, C, or D/F in the target course, Statics 2, is shown in the group of 4 bars above their Linear Algebra grade.

ics 2 students in these same bins, as shown in Fig. 3, which indicates that a linear regression analysis is appropriate. These results, especially graphs such as Fig. 2, could be particularly useful for advising.

In order to find the normalized correlation coefficients β for every statistically significant correlation within the curriculum, we performed separate iterative linear regressions with each course in Table I as the target. Table II reports the results for selected advanced engineering target courses from the MEMS Department, shown both with and without cumulative STEM GPA as a regressor. High school GPA, SAT Math, and the foundational courses that are missing from Table II were all controlled for but were not significant predictors for any of the target courses shown. A higher β indicates a stronger relationship from the regressor to the target course (recall that β is a normalized parameter). In particular, each β should be compared to the other β values in the column to estimate the regressor’s effect on the target relative to the other regressors. These results indicate “direct effects” from these predictor courses, which are the effects of the predictor course on the target while controlling for the effects of the other courses as well as GPA and SAT Math scores. The predictor courses can also have “indirect effects” if they are correlated with another predictor of the target course.

For all of the second-year engineering courses we investigated, including those in Table II, the top one or two predictors are advanced mathematics courses while the remainder of the regressors have comparable β lower than the mathematics courses. These relationships could in part be due to the advanced mathematics courses being taken concurrently with the target course, for example Statics 1 is concurrent with

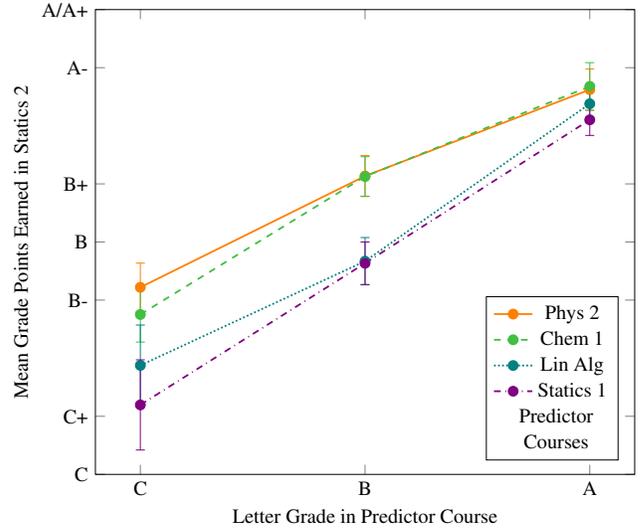


FIG. 3. Students are binned by their letter grade in the predictor course, then the mean grade points earned in the target course, Statics 2, by the students in each bin are plotted vertically along with the standard error. A selection of one course from each of the four foundational disciplines is shown. This linear trend holds for every target/regressor pair in our analysis.

TABLE II. Each column reports the β coefficients from the iterative regression for the target advanced engineering course listed at the top. A “—” indicates the regressor was included initially but not a significant predictor while “N/A” indicates the regressor was not included at any stage (e.g. was not taken prior to or concurrent with the target according to the curriculum). The predictors shown are the only statistically significant predictors ($p < 0.05$).

Regressor	Mat. Structure		Statics 1		Statics 2	
	No GPA	GPA	No GPA	GPA	No GPA	GPA
Chem 1	—	—	—	—	0.13	—
Chem 2	0.12	—	0.10	—	—	—
Engr 1	0.11	0.07	0.08	—	—	—
Engr 2	—	—	0.12	0.11	—	—
Phys 2	0.16	0.13	0.10	0.10	0.08	—
Calc 3	0.18	0.16	0.21	0.14	—	—
Lin Alg	0.26	0.25	0.24	0.23	0.17	0.13
Diff Eq	N/A	N/A	N/A	N/A	0.38	0.27
Statics 1	N/A	N/A	N/A	N/A	0.15	0.09
STEM GPA	N/A	0.18	N/A	0.19	N/A	0.35
N	654	836	823	1063	603	735
R^2	0.39	0.41	0.40	0.38	0.50	0.53

Linear Algebra and Calculus 3. However, for Statics 2 we see that Linear Algebra ($\beta = 0.17$) is a slightly better predictor than Statics 1 ($\beta = 0.15$), the previous course in the Statics sequence, despite each of those courses being taken the semester prior and the presence of the concurrent Differential Equations ($\beta = 0.38$) already accounting for the bulk of the impact of mathematics. The direct effects of the other found-

TABLE III. β coefficients from additional regressions for the target course Statics 1. The left columns include students who took Linear Algebra (LA) prior to or concurrent with Statics 1, the right columns include students who had not yet taken Linear Algebra. For each group, the regressor course with the highest β is omitted to investigate how the dependences redistribute to the remaining courses.

Regressor	Target: Statics 1			
	Had LA prior or conc.		LA not yet taken	
	Included	Omitted	Included	Omitted
Engr 2	0.11	0.13	—	—
Phys 2	0.10	0.10	0.10	0.13
Calc 3	0.14	0.19	0.29	omit
Lin Alg	0.23	omit	N/A	N/A
STEM GPA	0.19	0.29	0.32	0.49
N	1063	1063	607	607
R^2	0.38	0.35	0.40	0.36

dational courses are comparable, with each of physics, chemistry, and engineering having the second-largest direct effect on one of the three advanced target courses shown in Table II (e.g. Physics 2 is the next best predictor for Mat. Structure after the mathematics courses). When STEM GPA is included, mathematics courses remain the top predictors while the β of the remaining courses reduces unevenly, revealing a further hierarchy beyond mathematics in which chemistry is no longer a predictor. We interpret this to mean that the results without STEM GPA include a general measure of student success which is then controlled for by STEM GPA.

To investigate the hierarchy of predictor courses, we will focus further on just one of the target courses, Statics 1. Table III shows additional linear regressions with Statics 1 as the target where we investigate how the relationships change when the top predictor courses are omitted from the final state of the linear regression. In the left two columns of Table III, we look at how the β coefficients change for Statics 1 students who have taken Linear Algebra when we omit Linear Algebra as an independent variable. In the right two columns, we look at a separate population of students who have not taken Linear Algebra prior to or concurrent with Statics 1.

Comparing the first two columns of Table III shows that for the students who have taken Linear Algebra, the predictive power of Linear Algebra distributes primarily to STEM GPA followed by Calculus 3. In particular, the overall contribution from mathematics decreases from the first column to the second while the contribution of STEM GPA increases. This

suggests that additional advanced mathematics courses provide benefits above and beyond earlier mathematics courses, and that those students who have fallen behind in the curriculum are likely to continue to receive similar grades to their previous semesters. Furthermore, the predictive power of Physics 2 remains consistent across all iterations of the analysis in Table III, all of which control for STEM GPA. This indicates that the physics sequence is a strong predictor of performance in subsequent engineering courses regardless of a student’s progress through the curriculum.

IV. SUMMARY AND IMPLICATIONS

Using a linear regression analysis we find good correlation between foundational mathematics and physics courses and subsequent engineering courses in the MEMS curriculum. Consistent with Sadler and Tai’s findings for correlations between high school mathematics and college STEM performance [5], we find that the courses that best predict performance in advanced engineering courses are advanced mathematics courses, namely Calculus 3, Linear Algebra, and Differential Equations. We find that these advanced mathematics courses work synergistically — that is, the benefit of additional advanced mathematics courses is above and beyond the benefits of earlier mathematics courses. These benefits appear to be above and beyond any general measure of student performance, since the mathematics courses are consistently stronger predictors than any other course even when STEM GPA is included. Additionally, the presence of STEM GPA as a significant predictor raises questions about how well the current system promotes student improvement as they progress through the curriculum.

We also find that introductory engineering and physics courses have comparable direct effects. Moreover, in additional analyses of Statics 1, we find that reliance on physics is consistent between students who have and who have not taken Linear Algebra. We interpret this to mean that the physics courses should be considered alongside the mathematics courses as a primary foundational support for mechanical engineering. Further investigation into what caused some students to delay taking Linear Algebra could reveal potential barriers to students’ success.

We plan to extend this analysis by including demographic information such as gender and ethnicity and investigating the content and skills used in these courses.

ACKNOWLEDGMENTS

This research is supported by NSF Grant DUE-1524575.

[1] R. Felder, G. Felder, and E. Dietz, *J. Eng. Educ.* **87**, 469 (1998).
 [2] O. Eris, D. Chachra, H. Chen, S. Sheppard, L. Ludlow, C. Rosca, T. Bailey, and G. Toye, *J. Eng. Educ.* **99**, 371 (2010).
 [3] M. Besterfield-Sacre, M. Moreno, L. Shuman, and C. Atman, *J. Eng. Educ.* **90**, 477 (2010).

[4] A. Astin, *What Matters in College: Four Critical Years Revisited*, (Jossey-Bass, 1993).
 [5] P. Sadler and R. Tai, *Science* **317**, 457 (2007).
 [6] J. Cohen, *Statistical Power Analysis*, (Lawrence Erlbaum, 1988).