


## Importance of math prerequisites for performance in introductory physics

Eric W. Burkholder<sup>1</sup>, Gabriel Murillo-Gonzalez<sup>1</sup>, and Carl Wieman<sup>1</sup>  
*Department of Physics, Stanford University, Stanford, California 94305, USA*

 (Received 10 November 2020; accepted 1 February 2021; published 17 February 2021)

Previous work has looked at the relationship between high school preparation and student performance in calculus-based introductory mechanics (physics 1) courses. Here, we extend that work to look at performance in introductory calculus-based electricity and magnetism (physics 2), and we look at the significance of what college math courses have been completed in addition to high school preparation. Using multiple linear regression including these measures of prior preparation, we examine the correlation between taking various math courses in college and final exam scores in introductory physics courses at a highly selective west coast university. In physics 1, we find that prior college math coursework is not a predictor of physics 1 final exam score. In physics 2, we find that having taken a course in vector calculus is a strong predictor of physics 2 exam performance (effect size = 0.58 standard deviations,  $p < 0.001$ ), even when controlling for students' physics 1 final exam scores (effect size = 0.27 standard deviations,  $p < 0.01$ ). These effect sizes are similar in magnitude to other measures of students' incoming physics and math preparation. Qualitative analysis of student exams from physics 2 reveal that this “vector calculus gap” is due to differences in reasoning about vectors and geometry and some differences in conceptual understanding of circuits, as vector calculus itself is not required to perform well on the final exam. That is, basic reasoning related to vector calculus appears to be important, but the formalisms of vector calculus do not.

DOI: [10.1103/PhysRevPhysEducRes.17.010108](https://doi.org/10.1103/PhysRevPhysEducRes.17.010108)

### I. INTRODUCTION

Researchers in physics education research (PER) have recently started to focus on what factors explain the variation in student outcomes on various assessments such as course final exams [1–4] and concept inventories [5–9]. These investigations have largely focused on issues of equity—determining whether there are demographic performance gaps in these assessment outcomes, what factors might affect these gaps, and how they can be eliminated. A secondary outcome of this line of research has revealed that much of the variations in student outcomes can be explained by variations in students' preparation for various physics courses [10]. This has profound implications for instructional design—courses that are tailored to students with stronger physics backgrounds are effectively discriminating against students with weaker high school preparation, something that is often correlated with lower socioeconomic status of their respective communities [11].

Prior research has used measures such as SAT and ACT scores, concept inventory prescores, and high school GPA

as variables in determining students' physics outcomes—all of these variables focus on students' high school physics and math preparation. Less attention has been paid to the role of college prerequisites in explaining students' physics course performance. In this study, we build upon previous work to look at the impact of math prerequisites on performance in the introductory physics courses for engineers and scientists at Stanford University.

### II. LITERATURE REVIEW

Early work in PER focused mostly on the average outcomes of physics courses for the purposes of determining how well a particular teaching method worked. Other researchers extended such analyses to start to answer issues of equity in the physics classroom—in particular, characterizing the gap in average performance between male and female students on concept inventories and in introductory courses. This led to spirited debates about how the gender gap should be measured [12–15] and the most suitable types of analyses.

Researchers in PER have recently begun to use more sophisticated statistical methods such as multiple linear regression [10] and machine learning [16]. These methods have a number of advantages over univariate statistical tests, but the most important is that these methods allow researchers to estimate the effect sizes of multiple different variables on a single outcome. Thus, researchers can

---

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

determine the size of demographic performance gaps in course outcomes while controlling for other relevant variables like high school physics preparation. This modifies the question that researchers are able to ask from “are there differences in performance between different groups,” to “do students from different groups perform differently even if they know the same amount of physics and math at the beginning of the course?” The former question is a question of “equity of outcomes” [17], while Burkholder *et al.* characterize the latter as a question of “equity of opportunity.” [18].

The results of these more sophisticated analyses have varied. Some researchers find that there are significant gender gaps on concept inventory post-scores and final exam scores even after controlling for measures of high-school or incoming physics preparation (e.g., concept inventory prescore and SAT or ACT math score), causing researchers to question whether concept inventories may possess some inherent gender bias and whether there are equity issues in the way the introductory course sequence is taught [5–7]. One thing that has been consistent, however, is that students’ prior physics and math preparation are the most important factors that predict success in introductory courses.

A recent study, however, found that demographic performance gaps are small or zero after controlling for students’ incoming physics preparation [10]. Importantly, this study used measures of general college preparation (e.g., SAT or ACT scores) and subject-specific preparation (concept inventory prescores, AP scores, etc.). Salehi *et al.* used structural equation modeling to show that gender gaps in physics 1 course outcomes could be explained by a gender gap in concept inventory prescores, while a URM and first-generation performance gap could be explained by respective gaps in SAT or ACT math scores [10]. These results suggest that introductory physics courses may not be introducing demographic performance differences, but rather they are preserving more systemic educational differences by implicitly depending heavily on students’ incoming physics knowledge. Disadvantaged students will not have access to quality physics or calculus courses at their high schools, leaving them less prepared for the physics 1 course and, as this study shows, resulting in lower grades in those courses.

The existing literature consistently shows that academic preparation is an important variable in determining students’ performance in introductory physics courses. However, most studies have focused on measures of preparation at the K-12 level, such as SAT scores, AP scores, high school GPA, and concept inventory prescores. Fewer studies have quantitatively examined the effect of college-level mathematics coursework on outcomes in these introductory courses.

However, there are a number of studies documenting students’ difficulties with vector calculus and vector

representations in electrodynamics. For example, Schermerhorn and Thompson found that upper-level students had difficulty constructing unit vectors [19] and differential volume elements [20] for a novel spherical coordinate system. Bollen *et al.* found that students in general struggled with basic vector operations, such as vector addition [21], as well as more advanced ideas like divergence and curl [22–23]. Wallace and Chasteen find that students struggle with the idea that an integral represents a sum in Ampere’s law [24], and Pepper *et al.* find that students in upper-division electricity and magnetism (E&M) generally struggle with mathematical concepts applied to physical systems [25]. Notably, most of these studies involved upper-division E&M courses where vector calculus is thoroughly integrated into the course. Furthermore, they were qualitative studies aimed at identifying student difficulties, not quantitative studies linking these difficulties to course performance. Less is known about the role vector calculus plays in student performance in introductory E&M.

### III. RESEARCH QUESTIONS

We pose the following research questions:

1. Do students who have taken more advanced mathematics courses (vector calculus or above) prior to physics 1 and 2 perform better in the introductory physics courses after controlling for measures of high school physics and math preparation (as measured by FMCE score and SAT or ACT math score, respectively)? Similarly, do students who are behind in the recommended math sequence perform worse in these courses?
2. If there are significant gaps related to students’ math coursework, what are the differences in reasoning in solving physics exam problems between students who take advanced mathematics courses before the introductory physics sequence and those who don’t?

The first research question we will address using quantitative methods, while the second we will address using qualitative research methods.

### IV. METHODS

#### A. Course and institutional context

Our data were collected from students enrolled in the introductory physics sequence for scientists and engineers at Stanford University, a highly selective, private, research-focused institution in the western United States. Approximately 33% of the students in this course sequence are historically underrepresented in physics minorities, 22% are first-generation college students, and 48% are female.

We will refer to the calculus-based introductory mechanics course for scientists and engineers as “physics 1” and the calculus-based introductory electricity and magnetism

course for scientists and engineers as “physics 2.” Both courses are largely taught in a traditional lecture-based format, with some limited use of clicker questions during the lectures. Both courses also have recitation sections in which lecture content is reviewed and students are given the opportunity to practice problem solving using problems from *Tutorials in Introductory Physics* [26]. The textbook used for physics 1 was Young and Freedman [27], and for physics 2 was Knight [28]. Physics 1 covers kinematics, Newton’s laws, uniform circular motion, conservation of energy, conservation of momentum, and conservation of angular momentum. Physics 2 covers Coulomb’s law and electrostatics, Gauss’s law, capacitance and dielectrics, simple circuits, Ampere’s law, the Biot-Savart law, Faraday’s law, Lenz’s law, and Maxwell’s equations. Both courses have weekly problem sets which include problems from mastering physics. Both courses have two midterm exams and one final exam; the exam scores primarily determine the students’ grades.

We note that, while the teaching methods used in physics 1 and 2 were largely traditional, the gains in concept inventory scores reported in these courses are on par with what Hake called “interactive engagement” courses [29] (e.g., normalized gain in physics 1 is 0.48). Other researchers have noted problems with using normalized gain as a measure of course quality [30], and we posit that this is a poor measure of the kinds of teaching practices present in a classroom.

### B. Quantitative analysis

Data were collected from physics 1 in the winter quarter of 2018, and physics 2 in the spring quarter of 2017. There were 586 students enrolled in physics 1 and 422 students enrolled in physics 2 during these terms. Approximately 50% of students from physics 1 typically take physics 2 the following quarter. We collected students’ final exam scores from both courses, concept inventory prescores (force and motion conceptual evaluation, FMCE, for physics 1 [31] and conceptual survey of electricity and magnetism, CSEM, for physics 2 [32]), SAT or ACT math percentile scores, and data on which math courses students had taken prior to and concurrently with physics 1 and physics 2. For physics 2, we also collected those students’ physics 1 final exam scores from the winter quarter of 2017 and their FMCE prescores from that same quarter. We converted all exam and concept-inventory scores to  $z$  scores and transformed the information about mathematics coursework into two different binary variables.

Typically, students take calculus 1 (single variable differentiation), followed by taking calculus 2 (single variable integration) at the same time as physics 1 and before they take physics 2. More advanced students may start in the advanced math sequence, which covers vector calculus, differential equations, and linear algebra. For our analysis of physics 1 exam scores, we created two binary

variables. The first variable had a value of 1 if a student was behind the typical calculus sequence (enrolled in calculus 1 concurrently with physics 1) and a value of zero otherwise. The second variable had a value of 1 if students had taken vector calculus or more advanced coursework prior to physics 1 and zero otherwise. Note that a student cannot simultaneously be behind in calculus and have taken vector calculus (1, 1), but a student can have not taken vector calculus and still be in step with the math sequence (0, 0). Most students fall into the category of (i) being behind in the math sequence and not have taken vector calculus (1, 0) or (ii) not being behind in the math sequence and having taken vector calculus (0, 1).

For the analysis of physics 2 exam scores, we only had a single variable measuring prior math coursework. This variable had a value of 1 if students had taken vector calculus prior to physics 2, and zero if they had not. Calculus 2 (the corequisite for physics 1) is a prerequisite for physics 2 that all students have satisfied, while vector calculus is a corequisite for physics 2, so students who did not take vector calculus prior to physics 2 were taking it concurrently. Note that we surveyed students about their prior and concurrent mathematics coursework in both physics 1 and physics 2, so the variables measuring whether students took vector calculus prior to physics 1 or physics 2 are distinct. Approximately 8% of Stanford students take vector calculus or linear algebra in high school, but because it is not a formal AP course, they are not given college credit for this. These students are included in the group of students who took vector calculus prior to physics 1 or physics 2 as they would retake it in their first terms at Stanford. Note that we use separate datasets for physics 1 and physics 2 (from different years), so there is no overlap in mathematics preparation variables between the two populations. We also note that the vector calculus variable value of 1 contains two groups of students: (i) students who were ahead in the math sequence and took vector calculus early while taking physics 2 at the recommended time, and (ii) students who took vector calculus at the recommended time but waited to take physics 2. Both groups have taken vector calculus prior to taking physics 2.

We first conducted multiple imputation on both data sets to account for missing data. We used the mice package in R which assumes the data are missing at random (whether a value is missing is not related to that value, but may be correlated with other variables in the dataset) and uses linear and logistical regression to impute variables. We imputed 20 different datasets for both physics 1 and physics 2, conducted the appropriate linear and logistical regressions on the data, and then pooled those values to arrive at the final results which we present below. As detailed in Ref. [33], complete case analysis (deleting any student for whom we do not have complete data) likely introduces bias into regression results. Course participation is correlated with course performance, thus missing FMCE scores, etc.,

TABLE I. Errors identified in the first and fifth free response questions on the physics 2 final exam.

Problem	Conceptual	Calculus	Physical or geometric	Vector calc notation	Miscellaneous
FR1	N/A	Wrong sign after integrating. Integration mistake. Missing integral bounds.	Wrong integral bounds ( $r$ , $R$ , Infinity). Missing direction vector. Using wrong radial variable.	Missing vector hats in problem setup. Dot product mistake.	Wrong E-field, potential carried over. Wrong constant or number in volume/area formula. Wrong power in volume/area formula. Left blank.
FR5	The inductor stores energy. The current changes after a long time. Missing circuit elements.	N/A	N/A	N/A	Guessed final answer. Attempted to solve differential equation.

are inherently linked to the outcome variables and biased. Multiple imputation is a way to mitigate this bias. For more detail on missing data and multiple imputation, see Ref. [33].

For the physics 1 data, we conducted a multivariable linear regression of final exam scores on combinations of four other variables: FMCE prescore, SAT or ACT math score, prior vector calculus experience, and being behind in the math sequence. We report the regression coefficients and total R-squared values for several different combinations of these variables in the results section.

For physics 2 data, we first conducted a multivariable linear regression of physics 2 final exam score on gender, FMCE prescore from physics 1, CSEM prescore, SAT or ACT math score, physics 1 final exam score, and whether a student had taken vector calculus prior to taking physics 2. We then did separate regression analyses on the scores on the multiple-choice question portion of the final exam and on the scores on the free-response question portion of the exam. For the multiple-choice questions, we used a random-intercept logistic regression model to measure the effects of each variable on the probability of getting a particular question right. The random intercept model allows us to account for baseline variations in difficulty between multiple-choice questions. We conducted a similar analysis on the free-response questions using a random-intercept linear regression. For more detail on random-intercept models, see Refs. [26,34].

### C. Qualitative analysis

Our quantitative analysis indicated that the largest vector calculus gaps were on free response questions 1 and 5, and that for the other questions the vector calculus gaps were small (effect size  $< 0.1$  standard deviations). To understand what aspects of student reasoning were responsible for these gaps, we performed a qualitative analysis of physics 2 final exam solutions for these two questions. The second author selected the 66 students who did not take vector

calculus prior to physics 2, as well as 66 students with similar physics 1 final exam scores who had taken vector calculus for analysis. This was done by identifying the final exam scores for students who had not taken vector calculus, and finding a student who had taken vector calculus with a score that as closely as possible matched that score. For example, a student who had not taken vector calculus but scored a 73% on the exam might be matched with a student who had taken vector calculus and scored a 75%. This approximately controls for prior physics preparation so that we can make a fair evaluation that is equivalent to our quantitative model. The second author first made a rubric for grading the free-response questions. The first and second authors independently coded 10 physics 2 exam solutions to evaluate the validity and reproducibility of the rubric. In this independent coding, the authors refined the rubric and came up with a list of codes to characterize the types of errors that students were making on these two questions. A list of the error codes may be found in Table I. We grouped errors into five categories: conceptual errors (e.g., not understanding how a circuit element works), algorithmic calculus errors, errors in physical or geometric reasoning (e.g., spatial reasoning), mistakes in vector calculus notation, and miscellaneous other errors (e.g., not remembering a formula correctly).

The first two authors then coded another 10 responses independently using the error codes and a new rubric. Cohen's kappa for the first free-response question (FR1, Fig. 1.) was 0.81, and for the fifth free-response question (FR5, Fig. 2) it was 0.75. The second author then coded the solutions to these two questions for remainder of the  $66 \times 2$  exams.

Once all of the exam solutions were coded, we conducted additional quantitative analysis of the results. For each subpart of FR1 and FR5, we used a  $t$  test to determine if there were differences between students who had previously taken vector calculus and those who had not. If we found a significant difference, we then coded the

**FR1** A uniformly charged insulating sphere of radius  $R$  and positive volume charge density  $\rho$  is illustrated below.

(a) Using Gauss's law, derive an expression for the magnitude and direction of the electric field for  $r > R$ . Express your answer in terms of  $r, \rho, R$ , and the any constants as required. Clearly indicate all relevant steps in your derivation.

1. **+1** Student used Gauss's Law.
2. **+0.5** Correct Final answer is in terms of  $r, \rho, R$ , and any required constants.
3. **+0.5** Correct Final answer is expressed in magnitude and direction.

(b) Again using Gauss's Law, derive an expression for the magnitude and direction of the electric field in the region  $r < R$ . Express your answer in terms of  $r, \rho, R$  and any constants as required. Clearly indicate all relevant steps in your derivation.

1. **+1** Student used Gauss's Law.
2. **+0.5** Correct Final answer is in terms of  $r, \rho, R$ , and any required constants.
3. **+0.5** Correct Final answer is expressed in magnitude and direction.

(c) Assume the potential approaches zero as  $r$  approaches infinity. Using the expression determined in part (a), find an expression for the potential as a function of  $r$  in the region  $r > R$ .

1. **+1** Student set up their solution using  $\Delta V = \int \mathbf{E} \cdot d\mathbf{S}$
2. **+1** Student arrived at the correct algebraic expression for the electric potential using the expression from 21(a).

(d) Again, assume the potential approaches zero as  $r$  approaches infinity. Using the expressions obtained in parts (b) and (c), obtain an expression for the potential in region  $0 < r < R$ .

1. **+1** Student set up their solution using  $\Delta V = \int \mathbf{E} \cdot d\mathbf{S}$
2. **+1** Student arrived at the correct algebraic expression for the electric potential using the expressions from 21(b) and 21(c)

(e) Sketch the magnitude of the electric field **and** the electric potential as a function of  $r$  in the region  $0 < r < 3R$ .

1. **+0.5** Electric field function for  $r < R$  is a function of  $r$  and has the appropriate x-intercept.
2. **+0.5** Electric field function for  $r > R$  is a function of  $1/r^2$  and has the appropriate x-intercept.
3. **+0.5** Electric potential function for  $r < R$  is a function of  $-r^2$  and has the appropriate x-intercept.
4. **+0.5** Electric potential function for  $r > R$  is a function of  $1/r$  and has the appropriate x-intercept.
5. **-1** Axes are not appropriately labeled.

FIG. 1. Rubric for free response question 1.

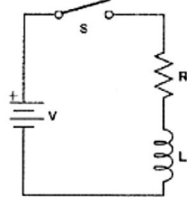
responses for that subpart for evidence of each error identified during initial grading. We totaled the number of students who had made a certain type of error and whether they had previously taken vector calculus or not. We then used Fisher's exact test (when counts were small) or a chi-squared test to determine if the frequency of the error was more prevalent in the vector calculus group or the nonvector calculus group.

FR1 required students to use Gauss's law to compute the electric field and electric potential inside and outside a uniformly charged sphere, as well as graph the potential and electric field. The problem as well as the final rubric

used to grade it (by the authors of this study, not the TAs who graded the problems for the quantitative analysis) may be found in Fig. 1. FR5 required students to qualitatively analyze an LR circuit.

We further identified 8 multiple-choice questions which required calculus, vectors, or geometric reasoning to determine whether there were vector calculus gaps on these individual questions. We analyzed each of the options for each multiple-choice question and identified the correct choice, as well as the conceptual error a student would have to make to select each of the incorrect responses. We then tallied which option students chose to look for differences

**FR5** An LR circuit is shown below. A switch  $S$  is held in the open condition for a long time, then the switch is closed. Express your answers to the questions below in terms of  $V$ ,  $R$ ,  $L$  and constants as required.



- (a) Immediately after the switch is closed, what is the voltage across the resistor? Explain.
- +0.5 For stating that the voltage is zero.
  - +0.5 For giving an appropriate explanation (i.e. inductor will oppose the change in current immediately after the switch is closed).
- (b) A long time after the switch is closed, what is the current through the inductor? Explain.
- +0.5 For stating that the current is  $I = v / R$ .
  - +0.5 For explaining why the current is  $I = v / R$  (i.e. the current does not change after a 'long time' has passed, so the inductor offers no resistance).
- (c) What is the power supplied by the battery to the circuit a long time after the switch is closed? Explain.
- +0.5 For getting the correct answer with the Power Formula in terms of  $V$  and  $R$
  - +0.5 For explaining why the inductor is not considered after a long time has passed in the circuit.
- (d) Use Kirchoff's laws to obtain an differential equation that describes the time evolution of the current in the circuit. You do not need to solve this equation.
- +1 Student uses the appropriate expressions for the voltage of the battery, resistor, and inductor in terms of  $V$ ,  $R$ ,  $L$ ,  $I$ , and  $t$ .
  - +2 Student uses Kirchoff's Voltage Law.
  - +1 Student arrives at the correct differential expression using the Voltage Law and the given terms.

FIG. 2. Rubric for free response question 5.

in answer choice between students with and without prior vector calculus. Details of the analysis may be found in the Appendix.

## V. RESULTS

### A. Quantitative results

The results of our multiple regression analysis of physics 1 final exam grades are in Table II. The regression

coefficients indicate the effect size for an individual variable on the outcome in units of standard deviations. For example, the regression coefficient of the FMCE prescore tells you the number of standard deviations increase in final exam scores per standard deviation increase in FMCE prescores. The regression coefficient of vector calculus prior tells you the size of the difference in units of standard deviation in final exam scores between students who do and do not take vector calculus prior to

TABLE II. Regression models predicting physics 1 final exam score. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ . Numbers in parentheses are the standard errors of the coefficients. All coefficients are in units of standard deviations.

Variable	Model 1	Model 2	Model 3a	Model 3b	Model 4	Model 5a	Model 5b
FMCE prescore	0.49*** (0.038)				0.35*** (0.041)	0.35*** (0.041)	0.35*** (0.041)
Math SAT or ACT		0.47*** (0.043)			0.31*** (0.046)	0.31** (0.046)	0.31** (0.046)
Vector calculus			-0.061 (0.093)			-0.082 (0.079)	
Behind in calculus				-0.42 (0.28)			-0.17 (0.23)
$R$ squared	0.24	0.23	0.000 87	0.0039	0.32	0.32	0.32

TABLE III. Regression models predicting physics 2 final exam score. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ , † $p < 0.10$ . Numbers in parentheses are the standard errors of the coefficients. All coefficients are in units of standard deviations.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Gender ( $F = 1$ )	-0.086 (0.12)								
FMCE prescore		0.15* (0.071)		0.12† (0.068)				-0.033 (0.062)	
CSEM prescore		0.36*** (0.060)		0.33*** (0.058)				0.15 ** (0.052)	0.13*** (0.032)
SAT or ACT percentile			0.30*** (0.061)	0.21*** (0.055)				0.086 † (0.050)	0.085* (0.037)
Physics 1 final exam					0.66*** (0.045)		0.64*** (0.046)	0.55*** (0.065)	0.54*** (0.045)
Vector calculus						0.58*** (0.12)	0.27** (0.10)	0.27 ** (0.10)	0.27** (0.069)
R squared	0.0019	0.21	0.091	0.25	0.44	0.061	0.45	0.48	0.48

physics 1 or 2. The R-squared values indicate the total fraction of variation in final exam scores explained by all of the variables included in the model.

Each standard deviation increase in FMCE prescore or SAT or ACT math score predicts a ~0.5 standard deviation increase in physics 1 final exam score (models 1 and 2). There is no statistically significant correlation between taking an advanced math course and physics 1 exam score (model 3a). There is a negative but insignificant correlation between not meeting math pre-requisites and physics 1 exam score (models 3b and 5b); we expect this effect is nonzero in general, but that it is insignificant due to lack of statistical power here. Model 4, which includes the FMCE prescore and math SAT or ACT score, is the simplest, best-fitting model to the data. Models 5a and 5b show that the effect of math course remains insignificant after controlling for FMCE prescore and SAT or ACT math score. Though our conclusions regarding the effects of “Behind in Calculus” are limited by statistical power, model 5b shows that the effect is small compared to more general measures of preparation for physics 1. Variance inflation factors are below 1.3 for all variables in all multivariate models, indicating that multicollinearity is not a concern [35]. In total, FMCE prescores and SAT or ACT math scores explain 32% of the variation in physics 1 exam scores. This is precisely the same analysis as was conducted in Ref. [10], but with a different dataset (from a different year), and this result is in agreement with the result of that study. In sum, these results suggest that high school preparation, but not college mathematics coursework, is important for performance in physics 1.

The results from our regression analyses of physics 2 final exam grades on various other measures of physics and college preparation may be found in Table III. We find that there is no statistically significant gender gap in final exam scores (model 1), even when not controlling for incoming preparation. This is different from Ref. [10], albeit for a different course, wherein a gender gap in physics 1 final exam scores was explained by gender differences in FMCE pre-scores (high school physics preparation). However, this is a different population, as only 58% of females and 56% of males that take physics 1 also take physics 2 the following term. Also, much less of the physics 2 material is covered in high school physics. We note that students who take physics 2 immediately following physics 1 are less likely to have taken vector calculus (73% vs 83%,  $p = 0.017$ ) than students that delay taking physics 2. It is unlikely that the students who wait are taking other courses besides vector calculus (e.g., circuits) related to physics 2 content as physics 2 is a prerequisite for all courses in electrical engineering.

We see that physics specific measures of high school preparation explain 21% of the variance in physics 2 final exam scores (model 2). Both FMCE and CSEM are predictive of performance, but there is a stronger

TABLE IV. Mixed-effects regression models predicting multiple-choice and free response scores on the physics 2 final exam. Numbers in parentheses are the standard errors of the coefficients. All coefficients are in units of standard deviations. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

	Multiple choice	Free response
CSEM prescore	0.13** (0.041)	0.090*** (0.027)
SAT or ACT percentile	0.12** (0.040)	0.023 (0.034)
Physics 1 final exam	0.51*** (0.056)	0.35*** (0.033)
Vector calculus	0.10 (0.085)	0.27*** (0.056)

relationship between CSEM prescores and final exam scores than between FMCE prescores and final exam scores, which makes sense as the CSEM is much more closely aligned with the course content. General college preparation (SAT or ACT math scores) explains 9.1% of the variance in final exam scores (model 3). Altogether, these three measures of high school preparation explain 25% of the variance in physics 2 final exam scores (Model 4), slightly smaller than the proportion of variance explained in physics 1 final exam scores by similar measures.

Physics 1 final exam grade is the strongest single predictor of physics 2 final exam performance, explaining 44% of the variance in final exam scores (model 5). In model 6 we also see a large gap in physics 2 final exam scores between students who take vector calculus prior to physics 2 versus concurrently with physics 2 ( $b = 0.58$ ). This gap remains statistically significant and moderate in size even when we control for physics 1 final exam scores (model 7).

The most parsimonious model—that which explains the greatest proportion of variance without being overfit—as judged by the R-squared value and likelihood ratio tests between models is model 9. This model says that CSEM prescores, SAT or ACT math scores, physics 1 final exam scores, and vector calculus preparation are all significant predictors of physics 2 final exam performance, together explaining 48% of the variation in physics 2 final exam scores. Altogether, these results suggest that both high school preparation and prior college mathematics coursework are important for success in physics 2.

To provide more detail into how these measures of incoming preparation, particularly the effects of vector calculus, on physics 2 final exam scores, we performed

an item-level analysis using linear mixed effects models (Table IV). First, we constructed a model of the multiple-choice question responses. We found that, accounting for the variation in difficulty of the multiple-choice questions as above, CSEM prescores, SAT or ACT math score, and physics 1 final exam score were still clearly predictive of performance. However, we found no statistically significant gap in multiple-choice outcomes between students who take vector calculus prior to or concurrently with physics 2.

The factors that predict performance on the free-response questions are different than for the multiple-choice questions (Table IV). CSEM prescore and physics 1 final exam score are predictive of performance, but less so than on the multiple-choice questions. Indeed, we find that SAT or ACT math score is not predictive of performance on the free-response questions, controlling for other factors. Unlike the multiple-choice questions, there is a significant vector calculus gap on the free response questions. Further analysis reveals that the largest gaps are on free response questions 1 and 5 ( $b = 0.39, 0.44$ , respectively, compared with the average of 0.27 for the exam as a whole, as indicated in Table III). We do not report R-squared values in Table IV as pooled values are not available for mixed effects models in the mice package.

## B. Qualitative analysis

In light of the results from our quantitative modeling, we conducted a qualitative analysis of student responses to free response questions 1 and 5.

Based on our rubric, on free response question 1, we found a significant vector calculus gap on parts a, b, and c of the question ( $p = 0.023$ ,  $p = 0.019$ , and  $p = 0.0014$ , respectively.  $P$  values are not corrected to account for multiple comparisons). Most students got parts d, and e incorrect so we observed no difference. The errors coded are given in Tables V–VII. Parts a and b involved computing the electric field inside and outside the sphere using Gauss’s law. We found that students who did not take vector calculus prior to physics 2 were more likely to make physical or geometric errors—forgetting to include the direction of the electric field or using the wrong Gaussian surface to make their calculations ( $p = 0.0059$ ). Students who did not take vector calculus prior to physics 2 were also more likely to make errors in their vector calculus

TABLE V. Error counts from qualitative analysis of FR1a. Number in brackets is the number of students for whom we found a mistake, out of 66 total students in each category (No vector calc and prior vector calc).

Error category	No vector calc	Prior vector calc
Physical or geometric	<ul style="list-style-type: none"> <li>• Missing direction (15)</li> <li>• Using wrong radial variable, or <math>r</math> vs <math>R</math> (7)</li> </ul>	<ul style="list-style-type: none"> <li>• Missing direction (10)</li> </ul>
Vector calc notation	<ul style="list-style-type: none"> <li>• Missing vector hat in problem set up (16)</li> </ul>	<ul style="list-style-type: none"> <li>• Missing vector hat in problem set up (7)</li> </ul>
Misc. errors	<ul style="list-style-type: none"> <li>• Wrong constant (or number) (12)</li> <li>• Wrong variable (8)</li> </ul>	<ul style="list-style-type: none"> <li>• Wrong constant (or number) (8)</li> <li>• Wrong variable (4)</li> </ul>



TABLE VI. Error counts from qualitative analysis of FR1b. Number in brackets is the number of students for whom we found a mistake, out of 66 total students in each category (No vector calc and prior vector calc).

Error category	No vector calc	Prior vector calc
Physical or geometric	<ul style="list-style-type: none"> <li>• Missing direction (20)</li> <li>• Using wrong radial variable (8)</li> <li>• Incorrect ratios (11)</li> </ul>	<ul style="list-style-type: none"> <li>• Missing direction (12)</li> <li>• Incorrect ratios (3)</li> </ul>
Vector calc notation	<ul style="list-style-type: none"> <li>• Missing vector hat in problem set up (13)</li> </ul>	<ul style="list-style-type: none"> <li>• Missing vector hat in problem set up (7)</li> </ul>
Misc. errors	<ul style="list-style-type: none"> <li>• Wrong constant (11)</li> </ul>	<ul style="list-style-type: none"> <li>• Wrong constant (8)</li> </ul>

TABLE VII. Error counts from qualitative analysis of FR1c. Number in brackets is the number of students for whom we found a mistake, out of 66 total students in each category (No vector calc and prior vector calc).

Error category	No vector calc	Prior vector calc
Calculus errors	<ul style="list-style-type: none"> <li>• Wrong sign after integrating (3)</li> <li>• Wrong terms after integrating (5)</li> <li>• Missing int bounds (5)</li> </ul>	<ul style="list-style-type: none"> <li>• Wrong sign after integrating (1)</li> <li>• Wrong terms after integrating (7)</li> <li>• Missing int bounds (3)</li> </ul>
Physical or geometric	<ul style="list-style-type: none"> <li>• Wrong int bounds <math>R</math> to <math>r</math> (lower to upper bound) (12)</li> <li>• Wrong int bounds <math>R</math> to <math>\infty</math> (6)</li> <li>• Wrong int bounds <math>r</math> to <math>R</math> (3)</li> <li>• Wrong int bounds <math>r</math> to <math>\infty</math> (4)</li> <li>• Wrong int bounds 0 to <math>r</math> (4)</li> </ul>	<ul style="list-style-type: none"> <li>• Wrong int bounds <math>R</math> to <math>r</math> (lower to upper bound) (12)</li> <li>• Wrong int bounds <math>R</math> to <math>\infty</math> (5)</li> <li>• Wrong int bounds <math>r</math> to <math>R</math> (0)</li> <li>• Wrong int bounds <math>r</math> to <math>\infty</math> (1)</li> <li>• Wrong int bounds 0 to <math>r</math> (0)</li> </ul>
Vector calc notation	<ul style="list-style-type: none"> <li>• Missing vector hats in problem setup (18)</li> </ul>	<ul style="list-style-type: none"> <li>• Missing vector hats in problem setup (9)</li> <li>• Dot product between two numbers (2)</li> </ul>
Misc. errors	<ul style="list-style-type: none"> <li>• Wrong E-field from FR1a (23)</li> <li>• Left blank (3)</li> </ul>	<ul style="list-style-type: none"> <li>• Wrong E-field from FR1a (13)</li> </ul>

notation ( $p = 0.0091$ ), but this did not affect their scores on the problem. There were no differences in miscellaneous errors on part a ( $p = 0.26$ ). Part c required students to compute the electric potential inside of the sphere using the E-field derived in part a. We found that students in both groups were equally likely to make calculus mistakes (sign errors, integration mistakes, missing integral bounds), but that students who did not take vector calculus prior to physics 2 were far more likely to use the wrong integral bounds ( $p = 0.017$ ), reflecting a misunderstanding of the geometry.

On free response question 5, we found a significant vector calculus gap on part b using our rubric, which required students to specify the current through the inductor a long time after the switch had been closed. We found that students who had not taken vector calculus prior to physics 2 were more likely to say that the inductor stored energy or that the current would still be changing after a long time (12/66 vs 3/66,  $p = 0.017$ ).

When we conducted our analysis of the multiple-choice questions that could potentially probe geometric reasoning, calculus, or vectors, we found no differences between students who did and did not take vector calculus prior to physics 2. However, most students got nearly all of the multiple-choice questions correct, and for the only two questions that a significant fraction of the students got

incorrect, the errors were almost entirely making a simple sign error. This explains why we did not find any difference in the quantitative analysis of the multiple-choice questions. For more detail on the multiple-choice analysis, see the Appendix.

## VI. DISCUSSION

We conducted a mixed-methods analysis of physics 1 and 2 final exam performance for students enrolled in physics for scientists and engineers at Stanford University. We find that high school physics preparation is a strong predictor of final exam performance in physics 1 and physics 2, similar to Ref. [10]. In addition, we find that physics 1 final exam grades are the single strongest predictor of physics 2 final exam performance. Even after controlling for these measures of physics 2 preparation, we find that there is a vector calculus gap in physics 2 final exam scores—students who take vector calculus prior to physics 2 perform better. We found no effect of advanced mathematics preparation on performance in physics 1, and no effect of being behind in the math sequence on performance in physics 1.

The lack of correlation between calculus coursework and physics 1 exam performance is not surprising. Physics 1 does not require the use of advanced mathematical

techniques *per se*—it only requires a rather limited and basic understanding of calculus which nearly all of these students have before taking any calculus in college. Facility with algebra and algebraic reasoning and trigonometry is likely more important, hence the dependence on math SAT and ACT scores. However, one might expect math coursework to be a proxy for general academic preparation (i.e., a student who was able to take multivariable calculus in high school will have a stronger general math background). This does not seem to be the case in this population; further investigations are needed to see if this is simply due to an unusual characteristic of the student population we study here. We find a negative effect of being behind in math (not meeting the minimum co-requisite of calculus 2), which is reduced by 75% once we control for high school physics and math preparation (as measured by SAT or ACT math scores and FMCE scores). This effect is not statistically significant, likely due to the small number of students who fall into this category ( $N = 13$ ).

The significant correlation between vector calculus preparation and physics 2 final exam performance is also unsurprising, but the details are less clear. Many physical laws in electricity and magnetism rely on the use of vector calculus, and indeed one can pose many problems that require the use of vector calculus to arrive at an answer (e.g., spherical surface integrals). However, none of the problems on the final exam explicitly required students to use vector calculus to solve them. At most, they only required single variable integration and differentiation. We note that the size of the vector calculus gap might be slightly larger at Stanford than at other institutions because approximately 8% of the students who take vector calculus prior to physics 2 take it twice—once in high school, and once again when arriving at Stanford. Because it is a relatively small number of students, however, we expect any impact on our results from this population to be quite small.

Our analysis of students' solutions reveals that students who did not take vector calculus made more mistakes with the direction of vector fields and geometric reasoning, suggesting that this effect is not simply due to general mathematics preparation. This is in line with studies showing that students struggle with differential length vectors and volume elements in electrodynamics [19–20], and suggests that these difficulties persist into upper-division courses. In vector calculus, students get significant practice reasoning about three-dimensional geometries, which might plausibly give them an advantage on physics 2 exam problems that involve geometry but no explicit use of vector calculus. We saw no differences in basic competency with differentiation and integration between students who take vector calculus prior to or concurrently with physics 2.

A surprising result was the correlation with taking vector calculus on the performance on FR5, which was a circuit

problem, hence involving no thinking about three-dimensional space. We found that students who did not take vector calculus prior to physics 2 were more likely to be mistaken about how an inductor functioned. We see three possible explanations for this. The first is that *LR* circuits in this course are taught simultaneously with differential equations. For students with weaker math backgrounds, this significantly increases the cognitive load associated with learning about *LR* circuits and could interfere with learning. A second possible explanation is that *LR* circuits are taught very quickly at the end of the course and students with weaker math backgrounds were already further behind and did not absorb the new material as well. Finally, it is possible that some of the students who had already taken vector calculus had also taken differential equations as these were part of the same course sequence (we only had information about which course sequence students had previously been enrolled in for physics 2).

We found no differences between students who did and did not take vector calculus prior to physics 2 on the multiple-choice questions, but this was not surprising as these questions were largely testing memory of basic concepts or phenomena and involved very little use of math. Also, the questions were not very discriminating.

One limitation in this study is that our vector calculus group in physics 2 contains two groups of students that we cannot distinguish: (i) students who were ahead in the math sequence and took vector calculus early while taking physics 2 at the recommended time, and (ii) students who took vector calculus at the recommended time but waited to take physics 2. It seems plausible that students in group (i) would be even further ahead of the nonvector calculus students than students in group (ii). Thus, we may be underestimating the impact of vector calculus on physics 2 performance.

These results are limited because we have only analyzed student performance and preparation in a single electricity and magnetism course at a single, highly selective institution. We encourage other researchers to attempt similar analyses at their own institutions to see if they observe similar vector calculus gaps in course performance. We could imagine that different courses may place more or less emphasis on three-dimensional geometric reasoning and spend more or less time explicitly covering the mathematical ideas used, and as a result see different gaps. Because Stanford is a highly selective institution, we might also think there would be more variation in SAT and ACT math scores at other schools, and thus that this factor might be more predictive of course performance than it is here.

## VII. CONCLUSIONS

These results show that for this student population which calculus courses a student has taken has little to no correlation with physics 1 performance, but vector calculus preparation has a significant correlation with physics

2 performance after controlling for baseline physics preparation. This suggests that a student who starts their college career by taking calculus 1 concurrently with physics 1 may not experience trouble in physics 1, but they might be significantly disadvantaged in physics 2 because they do not have any vector calculus experience. In our analysis, even though vector calculus was not important for doing well on the physics 2 final exam, differences in familiarity working with vectors and reasoning about 3D geometries still lead to differences in exam performance. We cannot determine whether vector calculus is important for mastering the material in physics 2, but we can say that other mathematical reasoning skills that are practiced in vector calculus seem to be important. Notably, in this population, the vast majority of students are actually ahead in the recommended math sequence, which may bias the instructors to favor these more prepared students, and as a result leave the course less accessible for the students who come in with less preparation and follow the recommended sequence.

While these results may seem to imply that having taken vector calculus should be a prerequisite for physics 2, this would impose a major barrier to the completion of a STEM major for many students. Students who come from lower socioeconomic school districts are less likely to have calculus preparation in high school, and for them such a prerequisite would prevent them from taking physics 2 until their second year, which in many cases would make completion of many STEM majors quite difficult.

This is a challenging problem for physics departments to address and we propose a possible solution: provide a companion course to physics 2 that provides additional assistance for students who have not completed vector calculus. This course could focus on geometric reasoning and familiarity with vectors as well as concepts related to integration and differentiation. There is some evidence suggesting that co-curricular supports are indeed helpful for students' performance [36], but it remains an open question as to how to best design these experiences to optimize students' learning. We also hypothesize that shifting to highly interactive learning environments (SCALE-UP or extensive use of Peer Instruction) can help reduce the dependence of course performance on prior preparation, thus some models of instruction may be able to eliminate this vector calculus gap. How successful these methods are in reducing the gap remains an open question for further investigation, however.

#### APPENDIX: ANALYSIS OF MULTIPLE-CHOICE QUESTIONS

A qualitative analysis was performed on 20 multiple-choice questions from the physics 2 exam. Each question was evaluated on one main aspect: Does this problem require advanced math (geometry, integration, and/or vector calculus)? We can organize the mc questions into a category of "significant problems," these are the problems

that answer "yes" to our question. 8 out of 20 problems fell into this category, these are problems 1, 6, 8, 10, 11, 13, 14, and 18. Problems 7, 9, 12, 17, and 19 were circuit problems. Problem 3 only used algebra to solve, and students had to remember how a spherical conducting shell carried its charge. Problems 2, 4, 5, 15, 16, and 20 did not require students to use advanced math. Problems 15, 16, and 20 do require students to interpret a physical phenomenon, but it is not complex enough to require the student to use advanced math. Out of this trio (problems 15, 16, and 20), the only calculus that students use is interpreting  $d\phi/dt$  (change in flux over time).

We then characterized the concepts probed by and option breakdowns of the significant problems. It is important to note that when we look at concepts, we are referring to mathematical concepts and not physics concepts. When we evaluate a question's multiple choices, we will write "Correct" when the option is the correct choice, "?" when there is no logical explanation for why a student would pick that option, and a detailed explanation when an option is wrong but a reasonable choice for a student (e.g., if a student is not as comfortable with the cross product, then we can see why they would pick option A.).

**Problem 1:** (2 points) A uniformly charged rod with a positive charge density  $\gamma$  lies along the  $x$  axis as illustrated below. What is the magnitude of the  $y$  component of the electric field at the location  $y = y_0$  on the  $y$  axis due to the short segment  $dx$  at  $x = x_0$ ?

Concepts: Integration, geometry

Options:

- A)  $\frac{1}{4\pi\epsilon_0} * \frac{\lambda dx}{\sqrt{x_0^2 + y_0^2}}$
- B)  $\frac{1}{4\pi\epsilon_0} * \frac{\lambda dx}{x_0^2 + y_0^2}$
- C)  $\frac{1}{4\pi\epsilon_0} * \frac{\lambda dx}{x_0^2 + y_0^2} * \frac{x_0}{\sqrt{x_0^2 + y_0^2}}$
- D)  $\frac{1}{4\pi\epsilon_0} * \frac{\lambda dx}{x_0^2 + y_0^2} * \frac{y_0}{\sqrt{x_0^2 + y_0^2}}$
- E) None of the above

Breakdown:

- A) Forgot to square the inverse distance
- B) Only found  $E$
- C) Only found  $E_x$
- D) Correct
- E) ?

**Problem 6:** (2 points) An infinite plane of charge is located in the  $x$ - $y$  plane and produces a uniform electric field of  $1 \frac{V}{m}$  in the  $+z$  direction above the plane. What is the difference in the electric potential  $\Delta V = V_2 - V_1$  between points  $P_2$  at  $(1m, 0m, 1m)$  and  $P_1$  at  $(0m, 1m, 3m)$ ?

Concepts: Integration

Options:

- A)  $2V$
- B)  $-2V$
- C)  $\sqrt{6}V$
- D)  $-\sqrt{6}V$
- E) Not enough information

Breakdown:

- A) Correct
- B) Includes minus sign; this means that either the limits are wrong, or they forgot to include a minus sign in the original expression for potential
- C)  $\sqrt{6}$  is the distance between the two points, so if you used the length between p2 and p1 for your integral bounds you should get either C or D as your answer
- D) Same as C, but negative
- E) ?

**Problem 8:** (2 points) Which of the following cylindrical wire has the largest resistance? All wires are made of the same material.

Concepts: Geometry

Options:

- A) A wire of length  $L$  and diameter  $2d$ .
- B) A wire of length  $L$  and diameter  $d/2$ .
- C) A wire of length  $L/2$  and diameter  $d$ .
- D) A wire of length  $L$  and diameter  $d$ .
- E) A wire of length  $2L$  and diameter  $2d$ .

Breakdown:

- A) ?
- B. Correct. You can still answer this even if you incorrectly treat the diameter as a radius.
- C) ?
- D) ?
- E) ?

**Problem 10:** (2 points) A half-ring (semicircle) of uniformly distributed charge  $Q$  has radius  $R$ . What is the electric potential at its center? Assume that the potential from the half-ring approaches zero at distances very far from the half-ring.

Concepts: Integration, geometry

Options:

- H)  $\frac{Q}{4\pi\epsilon_0 R^2}$
- I)  $\frac{Q}{8\pi\epsilon_0 R^2}$
- J)  $\frac{Q}{4\pi\epsilon_0 R}$
- K)  $\frac{Q}{8\pi\epsilon_0 R}$
- L)  $-\frac{Q}{4\pi\epsilon_0 R^2}$

Breakdown:

- A. Used  $\frac{1}{r^2}$  instead of  $\frac{1}{r}$  in the original expression for potential
- B) ?
- C. Correct
- D) ?
- E) ?

**Problem 11:** (2 points) Consider two closely spaced and oppositely charged parallel metal plates. The plates are square with sides of length  $L$  and carry charge  $+Q$  and  $-Q$  on their facing surfaces. What is the magnitude of the electric field in the region between the plates?

Concepts: Integration, geometry

Options:

- A)  $E = \frac{Q}{\epsilon_0 L^2}$
- B)  $E = \frac{2Q}{\epsilon_0 L^2}$
- C)  $E = \frac{Q}{2\epsilon_0 L^2}$
- D)  $E = 0$
- E)  $E = \frac{Q}{4\pi\epsilon_0 L^2}$

Breakdown:

- A. Correct
- B. True if students interpreted the plates as only having charge on one side
- C.  $E = 0$  is only true for parallel plates with equal charge.
- D) ?
- E) ?

**Problem 13:** (2 points) Two very long parallel wires are a distance  $d$  apart and carry equal currents in opposite directions. The locations, if any, where the net magnetic field due to these currents is equal to zero are

Concepts: Geometry, Vector Calculus

Options:

- A) midway between the wires
- B) a distance  $d/2$  to the left of the left wire and also a distance  $d/2$  to the right of the right wire.
- C) a distance  $d$  to the left of the left wire and also a distance  $d$  to the right of the right wire.
- D) a distance  $d/\sqrt{2}$  to the left of the left wire and also a distance  $d/\sqrt{2}$  to the right of the right wire.
- E) The net field is not zero anywhere.

Breakdown:

- A. True if both currents were going in the same direction
- B. True only with this set up:  $\frac{\mu_0 I}{2\pi(\frac{d}{2})} + \frac{\mu_0 I}{2\pi(R)} = 0$
- C. True only with this set up:  $\frac{\mu_0 I}{2\pi(D)} + \frac{\mu_0 I}{2\pi(R)} = 0$
- D) ?
- E. Correct

**Problem 14:** (2 points) Consider a solenoid of length  $L$ ,  $N$  windings and radius  $b$  ( $L$  is much longer than  $b$ ). If the length of the solenoid became half as long ( $L/2$ ), the number of windings doubled ( $2N$ ), and all other quantities remained the same (the radius  $b$  and the current  $I$ ), the magnetic field inside the solenoid

Concepts: Geometry

Options:

- A) remains the same
- B) becomes twice as strong
- C) becomes four times as strong
- D) becomes one half as strong

Breakdown:

- A. If the student used the simplified version of finding the B field for a solenoid ( $B = \mu n I$ ) and confused  $n$  with the number of turns instead of  $n = \text{Number of turns/Length}$ .
- B) ?
- C. Correct
- D) ?

**Problem 18:** (2 points) An electron moves with an initial velocity  $v$  as shown below in the plane containing a current carrying wire. The current flows to the right, as shown below. In which direction is the magnetic force exerted on the electron?

Concepts: Vector Calculus, Geometry

Options:

- A) Into the page
- B) Out of the page
- C) Upward
- D) Downward
- E) The magnetic force is zero since the velocity is parallel to the current

Breakdown:

- A. ?
- B. ?
- C. Correct
- D. True if particle were positively charged
- E. ?

We counted for each question how many students were selecting each option. We present our results in Table VIII.

The vast majority of students answered questions 1, 8, 10, and 14 correctly. On problems 6, 13, and 18 the most popular incorrect answer simply involved a sign mistake, so this does not seem to measure students' understanding of advanced mathematical concepts. On problem 11 the most

TABLE VIII. MC Results: Bold indicates the correct response. If the Group/VC = 0, students have not taken vector calculus prior to physics 2, whereas if Group/VC = 1 students have taken vector calculus prior to physics 2.

Problem No.	Group/VC =	A	B	C	D	E
1	<b>0</b>	4	4	6	<b>50</b>	2
	<b>1</b>	4	9	9	<b>43</b>	1
6	<b>0</b>	<b>24</b>	23	7	9	3
	<b>1</b>	<b>27</b>	20	8	10	1
8	<b>0</b>	1	<b>65</b>	0	0	0
	<b>1</b>	0	<b>65</b>	1	0	0
10	<b>0</b>	3	1	<b>58</b>	4	0
	<b>1</b>	5	3	<b>57</b>	1	0
11	<b>0</b>	<b>44</b>	11	7	2	2
	<b>1</b>	<b>48</b>	11	3	3	1
13	<b>0</b>	16	4	1	1	<b>44</b>
	<b>1</b>	11	3	4	1	<b>47</b>
14	<b>0</b>	13	5	<b>47</b>	1	N/A
	<b>1</b>	7	0	<b>59</b>	0	N/A
18	<b>0</b>	1	4	<b>33</b>	25	3
	<b>1</b>	2	6	<b>32</b>	24	2

popular incorrect answer involved a factor of two difference, again not measuring any understanding of advanced mathematical ideas.

- [1] L. E. Kost, S. J. Pollock, and N. D. Finkelstein, Characterizing the gender gap in introductory physics, *Phys. Rev. ST Phys. Educ. Res.* **5**, 010101 (2009).
- [2] L. E. Kost-Smith, S. J. Pollock, and N. D. Finkelstein, Gender disparities in second-semester college physics: The incremental effects of a smog of bias, *Phys. Rev. ST Phys. Educ. Res.* **6**, 020112 (2010).
- [3] Z. Hazari, R. H. Tai, and P. M. Sadler, Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors, *Sci. Educ.* **91**, 847 (2007).
- [4] R. H. Tai and P. M. Sadler, Gender differences in introductory undergraduate physics performance: University physics versus college physics in the USA, *Int. J. Sci. Educ.* **23**, 1017 (2001).
- [5] A. Madsen, S. B. McKagan, and E. C. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020121 (2013).
- [6] R. Henderson, J. Stewart, and A. Traxler, Partitioning the gender gap in physics conceptual inventories: Force Concept Inventory, Force and Motion Conceptual Evaluation, and conceptual survey of electricity and magnetism, *Phys. Rev. Phys. Educ. Res.* **15**, 010131 (2019).
- [7] R. Henderson, G. Stewart, J. Stewart, L. Michaluk, and A. Traxler, Exploring the gender gap in the conceptual survey of electricity and magnetism, *Phys. Rev. Phys. Educ. Res.* **13**, 020114 (2017).
- [8] V. P. Coletta, J. A. Phillips, and J. Steinert, FCI normalized gain, scientific reasoning ability, thinking in physics, and gender effects, *AIP Conf. Proc.* **1413**, 23 (2012).
- [9] M. Lorenzo, C. H. Crouch, and E. Mazur, Reducing the gender gap in the physics classroom, *Am. J. Phys.* **74**, 118 (2006).
- [10] S. Salehi, E. W. Burkholder, G. P. Lepage, S. J. Pollock, and C. E. Wieman, Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics, *Phys. Rev. Phys. Educ. Res.* **15**, 020114 (2019).
- [11] *Talking about Leaving Revisited: Persistence, Relocation and Loss in Undergraduate STEM Education*, edited by E. Seymour and A-B. Hunter (Springer, New York, 2020).
- [12] S. D. Willoughby and A. Metz, Exploring gender differences with different gain calculations in astronomy and biology, *Am. J. Phys.* **77**, 651 (2009).
- [13] J. Day, J. B. Stang, N. G. Holmes, D. Kumar, and D. A. Bonn, Gender gaps and gendered action in a first-year physics laboratory, *Phys. Rev. Phys. Educ. Res.* **12**, 020104 (2016).

- [14] J. M. Nissen, R. M. Talbot, A. N. Thompson, and B. Van Dusen, Comparison of normalized gain and Cohen's  $d$  for analyzing gains on concept inventories, *Phys. Rev. Phys. Educ. Res.* **14**, 010115 (2018).
- [15] V. P. Coletta and J. J. Steinert, Why normalized gain should continue to be used in analyzing preinstruction and post-instruction scores on concept inventories, *Phys. Rev. Phys. Educ. Res.* **16**, 010108 (2020).
- [16] C. Zabriskie, J. Yang, S. DeVore, and J. Stewart, Using machine learning to predict physics course outcomes, *Phys. Rev. Phys. Educ. Res.* **15**, 020120 (2019).
- [17] B. Van Dusen and J. Nissen, Equity in college physics student learning: A critical quantitative intersectionality investigation, *J. Res. Sci. Teach.* **57**, 33 (2020).
- [18] E. Burkholder, C. Walsh, and N. G. Holmes, Examination of quantitative methods for analyzing data from concept inventories, *Phys. Rev. Phys. Educ. Res.* **16**, 010141 (2020).
- [19] B. P. Schermerhorn and J. R. Thompson, Physics students' construction of differential length vectors in an unconventional spherical coordinate system, *Phys. Rev. Phys. Educ. Res.* **15**, 010111 (2019).
- [20] B. P. Schermerhorn and J. R. Thompson, Physics students' construction and checking of differential volume elements in an unconventional spherical coordinate system, *Phys. Rev. Phys. Educ. Res.* **15**, 010112 (2019).
- [21] L. Bollen, P. van Kampen, C. Baily, M. Kelly, and M. De Cock, Student difficulties regarding symbolic and graphical representations of vector fields, *Phys. Rev. Phys. Educ. Res.* **13**, 020109 (2017).
- [22] L. Bollen, P. van Kampen, C. Baily, and M. De Cock, Qualitative investigation into students' use of divergence and curl in electromagnetism, *Phys. Rev. Phys. Educ. Res.* **12**, 020134 (2016).
- [23] L. Bollen, P. van Kampen, and M. De Cock, Students' difficulties with vector calculus in electrodynamics, *Phys. Rev. ST Phys. Educ. Res.* **11**, 020129 (2015).
- [24] C. S. Wallace and S. V. Chasteen, Upper-division students' difficulties with Ampere's law, *Phys. Rev. ST Phys. Educ. Res.* **6**, 020115 (2010).
- [25] R. E. Pepper, S. V. Chasteen, S. J. Pollock, and K. K. Perkins, Observations on student difficulties with mathematics in upper-division electricity and magnetism, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010111 (2012).
- [26] L. C. McDermott and P. S. Shaffer, *Tutorials in Introductory Physics* (Prentice Hall, Upper Saddle River, NJ, 2002).
- [27] H. D. Young and R. A. Freedman, *University Physics with Modern Physics*, 14th ed. (Pearson Boston, 2015).
- [28] R. D. Knight, *Physics for Scientist and Engineers: A Strategic Approach with Modern Physics* (Pearson Boston, 2017).
- [29] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *Am. J. Phys.* **66**, 64 (1998).
- [30] J. M. Nissen, R. M. Talbot, A. N. Thompson, and B. Van Dusen, Comparison of normalized gain and Cohen's  $d$  for analyzing gains on concept inventories, *Phys. Rev. Phys. Educ. Res.* **14**, 010115 (2018).
- [31] R. Thornton and D. Sokoloff, Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula, *Am. J. Phys.* **66**, 338 (1998).
- [32] D. Maloney, T. O'Kuma, C. Hieggelke, and A. Van Heuvelen, Surveying students' conceptual knowledge of electricity and magnetism, *Am. J. Phys.* **69**, S12 (2001).
- [33] J. Nissen, R. Donatello, and B. Van Dusen, Missing data and bias in physics education research: A case for using multiple imputation, *Phys. Rev. Phys. Educ. Res.* **15**, 020106 (2019).
- [34] E. Theobald, Students are rarely independent: When, why, and how to use random effects in discipline-based education research, *CBE Life Sci. Educ.* **17**, rm2 (2018).
- [35] M. H. Kutner, C. J. Nachtsheim, and J. Neter, *Applied Linear Regression Models*, 4th ed. (McGraw-Hill Irwin, New York, 2004).
- [36] C. Stanich, M. A. Pelch, E. J. Theobald, and S. Freeman, A new approach to supplementary instruction narrows achievement and affect gaps for underrepresented minorities, first-generation students, and women, *Chem. Educ. Res. Pract.* **19**, 846 (2018).