PHYSICAL REVIEW PHYSICS EDUCATION RESEARCH **22,** 020103 (2026)

Editors' Suggestion

Click, watch, learn: The impact of student self-study materials on physics electricity and magnetism course outcomes

James K. Hirons,[1] Jonathan D. Perry,[2] Kaylyn Sitka,[1] Dawson T. Nodurft,[1] Scott Crawford,[3] William Bassichis,[1] and Tatiana L. Erukhimova[1,*]

[1]*Department of Physics & Astronomy, Texas A&M University, College Station, Texas, 77843, USA*
[2]*Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA*
[3]*Department of Statistics, Texas A&M University, College Station, Texas, 77843, USA*

(Received 15 August 2025; accepted 12 June 2026; published 13 July 2026)

Performance in introductory courses, particularly physics, is often crucial for student success in science, technology, engineering, and mathematics majors and can impact an individual's tendency to persist in their chosen field. To enhance students' individual learning experiences, faculty at many universities have worked to develop open-access, self-study materials to help build conceptual understanding and problem-solving skills. Faculty at Texas A&M University have contributed to these efforts, creating more than 200 online video resources and a broad bank of prior exams available to students. This work explores and measures the impact that these resources can have on student course outcomes in an introductory, calculus-based electricity and magnetism course. Data were collected from three fall semesters, 2021–2023, including classroom performance, a conceptual assessment, and relevant university-level data to contextualize student background and preclass abilities. We present results from regression analyses examining the relative importance of multiple resources offered concurrently to students. We also discuss how the use of supplemental material affects course performance, alongside other potential predictors such as demographics and prior preparation. Similar to prior studies, we found that relevant prior preparation in mathematics was the strongest predictor of student performance, with prior physics knowledge being a weaker but still statistically significant predictor. Students' utilization of supplemental prior exams was the second-highest predictor of student performance. First-generation students were observed to have a slightly lower average performance on exams. However, interaction terms in the regression models suggested that first-generation students using supplemental prior exams were able to close this gap. Through anonymous surveys, students reported favorable impressions of the materials, with over 80% of students sharing that they had a noticeable contribution to their learning outside of the classroom and 98% stating they would recommend them to their peers.

DOI: 10.1103/p4jl-cjjd

I. INTRODUCTION

Introductory physics courses play a critical role in building foundational knowledge for students pursuing science, technology, engineering, and mathematics (STEM) majors. Calculus-based physics, in particular, is a required component for all physics and engineering students during their first 2 years of college. As part of students' early academic experiences, these courses are often pivotal to their success and progression within their chosen fields [1,2]. A recent 5-year longitudinal study on attrition and persistence in undergraduate physics programs found that the majority of students who lose interest in completing a physics major make that decision during their first or second year [3], underscoring the importance of student experiences in introductory physics courses. Research shows that success in introductory physics courses is strongly linked to students' self-efficacy in math and physics, prior academic preparation, and access to effective learning resources [4–12]. While strong mathematical foundations are consistently associated with better performance, incoming students often display a wide range of preparedness levels [12–14].

Compounding these challenges, introductory courses are frequently taught in large-enrollment settings, where classes have heterogeneous levels of student preparation, ranging from those with substantial preparation or high scores on Advanced Placement exams to those who may have never had a prior physics course. Due to the size of these courses, there are often limited opportunities for individualized attention and tailored support. Prior

*Contact author: etanya@tamu.edu

literature shows that first-generation students can experience additional academic challenges, compounding the issues of large classes, related to academic self-regulation [15,16]. In recent decades, a substantial body of knowledge has evolved, offering instructors myriad strategies for effective teaching [17–20]. While these strategies have made significant strides in improving student learning outcomes across many introductory courses, there remains room for improvement beyond changes to classroom pedagogy. It is important to recall that a lot of student learning still takes place outside of the scheduled class time, with students spending sometimes considerable effort preparing for lectures, completing assignments, and studying for exams, often without direct access to guidance from the instructor or instructional team [20]. In today's digital landscape, students have become increasingly reliant on online resources to support their learning [21].

Online platforms such as FlipItPhysics and Interactive Video Vignettes provide exemplars of efforts that have demonstrated measurable benefits for student learning and performance by offering structured, interactive content that reinforces core STEM concepts [22–24]. FlipItPhysics, for example, uses prelecture videos to replace traditional lectures, allowing classroom time to focus on problem-solving and demonstrations. Broadly, supplemental learning materials—including structured videos, problem-solving guides, and archived exams—have emerged as valuable tools for bridging gaps in student preparation [7,24–28]. Prior research has shown that supplemental resources can aid in students' development of conceptual understandings in physics [7,25,27]. In addition, interactive problem-solving tools have been shown to improve participation and improve students' ability to navigate complex problems [29,30]. However, the impact of these materials can vary based on individual interaction patterns and study strategies.

The increased availability of digital course materials during the pandemic proved valuable to students. When surveyed, many expressed a desire to retain access to recorded lectures and online content even in postpandemic, face-to-face classrooms [31,32]. To support student success, address gaps in preparation, and accommodate diverse learning needs, instructors at many institutions have developed open-access resources that foster self-directed learning outside the classroom [22,27–29,33]. As the field of physics education evolves, multimedia resources have become an increasingly prominent instructional tool, reshaping the ways students engage with complex material [25,34]. This is especially critical, as success in introductory science and mathematics courses can determine whether students persist through mid-level and advanced coursework [1].

In this study, we focused on the role of learning resources by examining how multiple supplemental learning resources influenced student success in a large, calculus-based introductory electricity and magnetism course at Texas A&M University. More than 10 years ago, physics faculty there began creating a series of online supplemental materials to support student learning. We hypothesized that the usage of our department's supplemental materials can have a significant, measurable improvement on student course outcomes. Our specific research aims were to investigate three key questions: (i) what impact did supplemental resources have on student outcomes in the introductory electricity and magnetism course, (ii) how did the impacts differ between first-generation and continuing-generation students, and (iii) what were students' perceptions of the impact of these resources. This work expands on previous studies, which have investigated the effect of individual resources, such as online videos, on learning by broadly investigating how multiple resources, each designed to meet different self-study needs, affect students' course outcomes. In the following sections, we describe the course context, supplemental resources, and analysis methods; present results from several regression models and surveys; and discuss the implications of our results.

II. RESEARCH METHODS

A. Course description

The data for this study were collected from students who took a second-semester calculus-based electricity and magnetism class during the fall semesters of 2021–2023 at Texas A&M University. These were large-enrollment courses that had around seven sections enrolling ∼140 per section, with an additional honors section of around 50 students each semester. These courses were taught by a large team of instructors, consisting of five or six faculty and upward of 12 graduate teaching assistants (TAs). All students used the same textbook, *Don't Panic: A Guide to Introductory Physics For Students of Science and Engineering*, Volume 2, by William Bassichis, a coauthor of this paper. In addition to 3 h of lecture per week, students attended an 80-min recitation session each week of no more than 28 students. The recitation portion, where students practiced problem-solving both individually and in groups, was taught by TAs.

All instructors (faculty and graduate students) met weekly to gather feedback from the previous week and discuss plans for the following week, ensuring instructional uniformity for students taking the course. Instructors utilized in-class lecture quizzes, either on paper or through the i-Clicker system, as well as weekly quizzes administered through Gradescope. At the beginning and end of the semester, students took the Brief Electricity and Magnetism Assessment (BEMA) [35], completion of which was a small bonus toward their overall quiz grade for the course, equivalent to one weekly quiz. The course summative assessments included three common midterm exams and a comprehensive final exam. All exam questions were free-response, multistep problems, with students expected to provide detailed solutions that indicated their thought process: the law(s) they used to approach the problem and the steps that led to the final answer.

A focus of the pedagogy of the course was the development of expertlike problem-solving abilities. This has often been considered an essential component of physics education, fostering the development of expert-level problem-solving skills [20,36]. It begins with a clear definition of problem solving, not as the memorization of formulas or procedures but as the application of fundamental principles to novel and complex situations. A key distinction between novice and expert problem solvers lies in their approach: while novices often rely on surface features and rote learning, experts engage with the underlying concepts and structure of problems. Instruction in the course, therefore, emphasizes strategic thinking and conceptual understanding. For example, rather than memorizing that capacitors in series share the same charge, students are guided to understand why this is the case, using physical laws such as charge conservation and $\oint \vec{E} \cdot d\vec{r} = 0$ in steady state. Exposure to challenging problems, such as analyzing circuits with initially charged capacitors or determining electric potential differences in wires with varying geometries, encourages deeper engagement and adaptability. Even when problems appear similar, subtle differences, such as the shape of a wire or the distribution of resistivity, can significantly alter the required approach, underscoring the need for flexible, principle-based reasoning.

The three midterm exams were common exams administered to all students taking the class at the same time. All instructors and TAs graded student papers immediately following the exam and returned them during the next lecture, ensuring that feedback was delivered while the content remained fresh in students' minds, a practice shown by prior research to support students' learning [20]. To ensure uniformity, each problem on the exam was graded by the same team, consisting of a lead faculty member and TAs. The final comprehensive exams were written by individual instructors, and these exams were scheduled on different days. However, all instructors submitted their final exams to a course coordinator to check for uniformity between instructors. Importantly, instruction was not oriented toward exam preparation alone; all exam problems were original and designed to assess student understanding.

Students were encouraged to use the open-access supplemental materials [37] created over the years to assist learners with varying backgrounds and levels of preparation in mastering the course learning outcomes outside the classroom. These self-study materials were posted on the departmental website and were created with support from internal grants. Students were not given any course credit for their usage of supplemental materials. The open-access materials included four distinct resources tailored to different types of learners.

The resources consisted of the following:

1. Chapter outline videos, developed by authors W. B. and T. L. E., were designed to offer a structured overview and summary of key topics in each chapter of the required textbook, *Don't Panic*. These videos were intended for students to watch before reading the chapter and attending lectures.
2. Conceptual and example videos, developed by W. B., T. L. E., and J. D. P., were created to reinforce key topics after they had been introduced in class and to emphasize their application to common and important problems. These short videos included explanations, derivations, and demonstrations to help illustrate the concepts.
3. Problem-solving videos, also referred to as recitation videos, developed by T. L. E., provided detailed solutions to multistep problems that students might encounter in their homework or major summative assessments. Their in-depth approach was designed to help students develop and refine their problem-solving skills and techniques.
4. Collection of midterm and final exams from previous years, developed by W. B. and T. L. E., spanning over 15 years. The midterm exams provided only the answers, without solutions to the problems, while the final exams included detailed solutions.

B. Data description

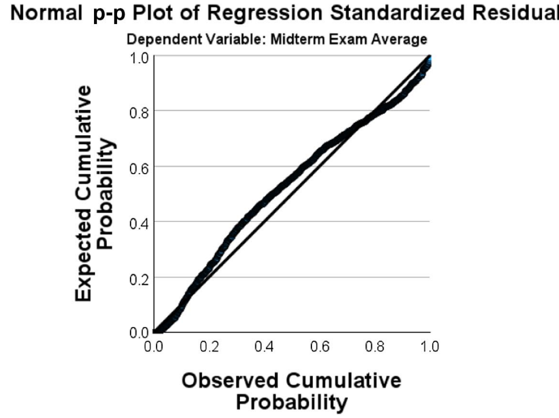
To investigate the impact of supplemental materials on student learning, we collected course-level, departmental-level, and university-level data for this study. The course-level data included students' midterm exam grades (out of a possible 100), final exam performance, final grades, and BEMA performance measured during the first and last days of recitation. Prior preparation in physics was measured using the BEMA prescore, while exam performance and final grades provided numeric and categorical representations of student success. Anonymous surveys were also conducted after each of the three midterm exams and were aimed at gauging self-reported student impressions of the impact and utility of the supplemental materials on their course performance.

The departmental-level data included student usage of the supplemental resources, which was tracked by monitoring students' clicks on each individual embedded hyperlink on our supplemental material website. These counts provided an estimate of each student's usage of individual videos or past exams and enabled us to create four additional explanatory variables: one for each type of video and one for the past exams.

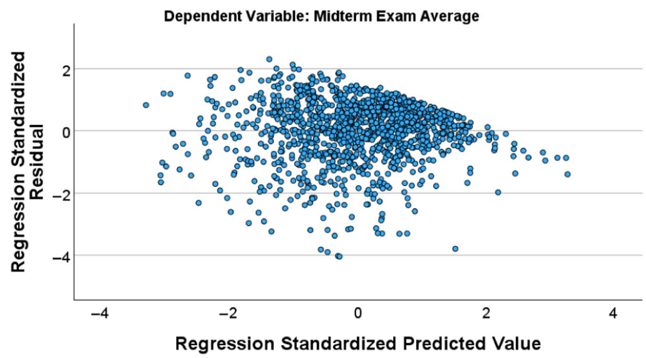
The university-level data included student demographics and background preparation. The demographic data included information related to student gender,

TABLE I. Demographics for students participating in study, $N = 2124$.

Demographic	Categories				
Ethnicity	White 51.22%	Hispanic 22.18%	Asian 16.71%	Black 2.97%	Other 6.92%
First generation	Yes 14.12%	No 83.24%	No record 2.64%		
Gender	Male 77.40%		Female 22.60%		



(a)



(b)

FIG. 1. Plots for check of assumptions for regression models, including (a) p-p plot of residuals and (b) a scatterplot of residuals.

ethnicity, and first-generation student status, as shown in Table I. Prior preparation in mathematics was assessed using the quantitative portion of the Scholastic Aptitude Test (SAT-math), the student's final letter grade in the first course of the calculus sequence taken at Texas A&M, and their performance on the required university-administered Math Placement Exam (MPE). Additionally, students' prior preparation in language was assessed using the qualitative portion of the SAT-reading. Letter grades from the students' calculus courses were treated on a standard four-point scale, with $A \equiv 4$, $B \equiv 3$, etc. For uniformity in the analysis, these numeric grade values, along with the other prior preparation variables, were normalized as z scores [38]. All three data sets were matched for each student by the university assessment team, anonymized, and provided to the researchers.

C. Statistical methods

Multiple linear regression analysis was conducted using the statistical software package SPSS. Prior to interpreting the results, key assumptions underlying the regression model were systematically evaluated: normality of residuals, homoscedasticity, and absence of multicollinearity. To assess the normality assumption, a p-p plot comparing the cumulative distribution functions of the observed and expected residuals was generated. The results pertaining to normality and homoscedasticity are presented in Fig. 1. The results fall within the acceptable threshold for normality, as shown in Fig. 1(a). Homoscedasticity was examined via a scatterplot of standardized predicted values against standardized residuals, as shown in Fig. 1(b). This assumption was met, though minor clustering was observed, attributable to the ceiling effect of exam scores, which limit average values to a maximum of 100. Assumptions were also checked for each independent, nonbinary variable included in regression models. These plots, shown in Supplemental Material [39], indicate that assumptions were met for each parameter. To evaluate multicollinearity, variance inflation factors (VIFs) were computed for all explanatory variables across the regression models. All VIFs were ≤ 1.7, which is well below the conventional threshold of 4 [40], indicating that multicollinearity did not impact estimated regression coefficients.

Variables expressed as counts are often skewed due to the inherent lower bound of zero; for instance, a student cannot engage with a video fewer than zero times. This limitation is reflected in the residual patterns observed for the supplemental resource variables, as shown in Fig. 2(a). Consequently, the explanatory variables corresponding to supplemental resources demonstrated non-normality when analyzed in linear space. To mitigate this issue, a logarithmic transformation was applied to these variables. Post-transformation residual plots revealed that the assumption of normality was largely satisfied, aside from observable clustering at the lower bound, as shown in Fig. 2(b).

In interpreting results of multiple regression models, several key metrics help in evaluating both the overall model performance and the relative influence of individual predictors. R-squared values, $0 \leq R^2 \leq 1$, quantify the proportion of total variance explained by the model. Beta coefficients represent the relative contribution of each independent variable to the observed effect. A threshold of statistical significance of $p < 0.05$ was adopted, a regularly selected value for such studies [38].

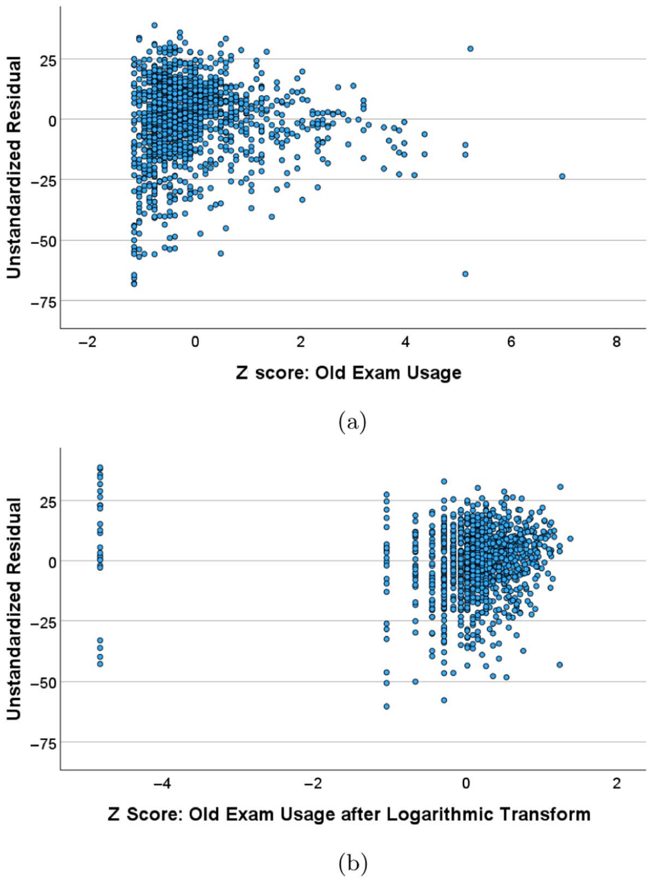


FIG. 2. Scatterplots of residuals for usage of supplemental prior years exams (a) no modifications and (b) postlogarithmic transform.

Initial regression models incorporated all available explanatory variables. To refine the model, we then employed backward elimination [41], which reduced the number of explanatory variables until only statistically significant factors remain. These explanatory factors included demographics (student gender, ethnicity, and first-generation status), usage metrics for each of the four supplemental materials, and indicators of prior academic preparation: SAT quantitative score, BEMA pretest score, a score on the university-administered MPE, and the final letter grade earned in Calculus I.

A total of 2556 students were enrolled over the three semesters, 2124 of whom opted into this study. However, due to missing data (e.g., BEMA pre, calculus grade, and first-generation status), we had complete data for 1368 students. As final exams and final course grades were assigned by individual instructors, the most uniform measure of student performance within the course was the common midterm exams. Therefore, the response variable in our multiple regression models was the average score between all three midterm exams. Distributions of average exam scores for the three midterm exams are shown in Figs. 3 and 4 for the complete dataset and the reduced dataset, respectively. We did not include the BEMA post-test as a response variable due to evidence suggesting students did not authentically engage with the questions, as well as low participation rates.

To examine impacts between multiple explanatory variables, we employed a limited number of interaction terms. These interaction terms facilitated predictions for specific student subgroups and enabled testing of the hypothesis that outcomes may differ by, for instance, generational status [10,14,42]. Interaction terms were created by multiplying the value of each relevant variable by a binary indicator of first-generation status. For example, the interaction term for student engagement with supplemental prior exams captured usage counts exclusively for first-generation students, with values set to zero for continuing-generation students.

When removing data with missing entries, the dataset was reduced by approximately 35%. Removal of data followed listwise deletion; that is, any row which was missing at least one piece of information was removed from the data. These included entries such as their final calculus grade or first-generation status. To enhance the predictive power of this analysis, we performed imputation. Imputation is a systematic method of modeling missing data [43]. Due to the sporadic nature of missingness within the dataset, the criteria of being missing completely at random were met [43]. To address this type of missing data, we employed multiple imputation by

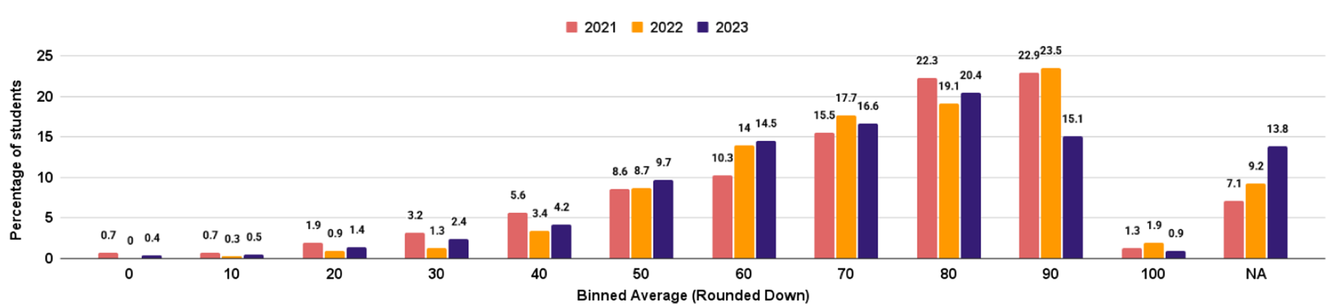


FIG. 3. Distribution of average score across all three midterm exams sorted by semester for all students participating in the study. Bars labeled "NA" indicate students who missed one or more of the common midterm exams.

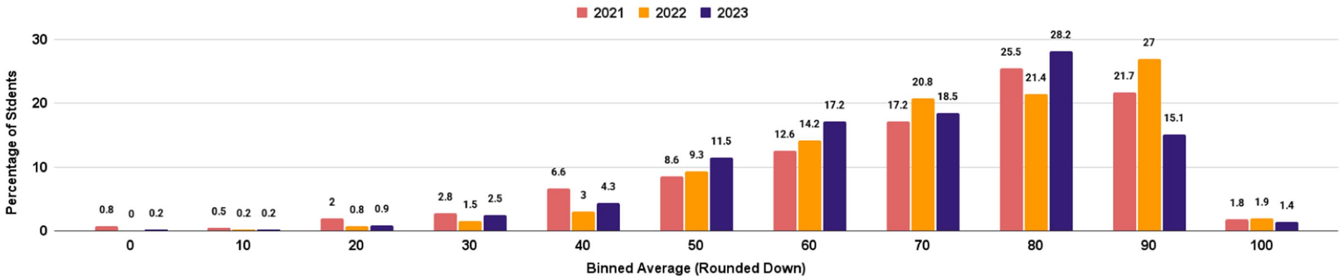


FIG. 4. Distribution of average score across all three midterm exams sorted by semester for the reduced dataset.

chained equations, generating 20 imputed datasets. Standardized beta coefficients were estimated by fitting linear models to each imputed dataset and pooling results using Rubin's rules [43]. Backward elimination was conducted on pooled results to identify statistically significant predictors. Model fit was assessed by pooling R-squared values across imputations. Average pooled R-squared values were computed for both the full and reduced models.

III. RESULTS

Preliminary regression models incorporated student ethnicity as an explanatory variable, revealing a statistically significant negative correlation between exam performance and students identifying as Hispanic or Asian. No significant effects were identified for the remaining ethnic groups. This pattern aligned with previous research indicating that observed performance disparities between minority and majority students may be attributable to other factors, particularly precourse academic preparation [10,44]. Subsequent analysis suggested that the statistical significance associated with ethnicity was predominantly driven by first-generation student status within these broader categories. In response to this finding, we omitted ethnicity from further analysis.

This section discusses three models. The first model was a multiple linear regression that included interaction terms linking first-generation student status to each of the four supplemental resources, with all variables represented in linear space. The second model applied logarithmic transformations to variables exhibiting non-normality, as previously noted. The third model combined interaction terms and logarithmic transformations to enable a more in-depth analysis of factors contributing to student performance. Table II shows the demographics for the backward-eliminated dataset for models 1 and 3, described in the following subsections. The demographics for the second model differ only minimally.

TABLE II. Demographics for students with listwise valid data, $N = 1368$.

Demographic	Categories				
Ethnicity	White 54.09%	Hispanic 21.86%	Asian 15.72%	Black 2.92%	Other 5.41%
First generation	Yes 12.21%		No 87.79%		
Gender	Male 77.49%		Female 22.51%		

A. First model: Interaction terms

The first model explored the interaction between students' first-generation status, their use of supplemental materials, and their academic performance. Interaction terms met all assumptions for regression modeling. Descriptive statistics and results for this model are presented in Table III. Our sample consisted of 1368 students, 12.2% of whom had first-generation status.

The model has an R-squared value of 0.302, indicating that it took into account 30.2% of the variability in student performance. The strongest predictor of student performance was their Calculus 1 grade ($\beta = 0.327$). The secondstrongest predictor was the use of supplemental prior exams ($\beta = 0.221$). Other significant prior preparation variables included MPE scores ($\beta = 0.175$) and BEMA pretest scores ($\beta = 0.157$). A small negative correlation was observed between first-generation status and course performance ($\beta = -0.066$). However, a positive correlation was found for the interaction term between first-generation status and usage of supplemental prior exams ($\beta = 0.050$). Among the supplemental videos, only the chapter outline videos showed a statistically significant effect, with a negative correlation ($\beta = -0.084$). The contribution to the variance of each of the four most significant factors from this model is shown in Table IV.

To address the missingness within the data, as described in Sec. II C, we completed multiple imputation with $m = 100$. This increased the R^2 value of the model to 0.362. Results from the imputed model are shown in parentheses in Table III. Similar to the original model, students' prior performance in calculus was the strongest predictor of their performance ($\beta = 0.336$). The second strongest predictor of student performance remained students' utilization of previous years' exams ($\beta = 0.232$).

TABLE III. Descriptive statistics and results for a multiple regression model with midterm exam average as the outcome variable, including an interaction term between first-generation status and usage of supplemental materials. Results from imputed models are presented in parentheses.

Descriptive statistics for interaction-term model, $N = 1368$		
Variable	Mean	STD
Exam average	75.676 (73.769)	17.472 (18.901)
Math placement	0.073 (−0.041)	0.868 (1.031)
Calculus 1 grade	0.072 (−0.014)	0.953 (1.014)
BEMA Pre	0.0004 (0.0003)	0.979 (1.001)
SAT reading	−(−0.007)	−(1.003)
Outline videos	$0.005(-2.472 \times 10^{-16})$	1.020 (1)
Prior exams	$0.046\ (5.795 \times 10^{-16})$	1.008 (1)
First generation (binary)	0.122 (0.145)	n/a
First generation × prior exams	0.020 (0.006)	0.445 (0.429)

Multiple regression model w/ interaction term		
$R^2 = 0.302$ (0.362), outcome var: ME avg		
Independent variable	Beta	Significance
Calculus 1 grade	0.327 (0.336)	<0.001 (<0.001)
Prior exams	0.221 (0.232)	<0.001 (<0.001)
Math placement	0.175 (0.169)	<0.001 (<0.001)
BEMA Pre	0.157 (0.137)	<0.001 (<0.001)
SAT reading	−(0.112)	−(< 0.001)
First generation × prior exams	0.050 (0.047)	0.047 (0.005)
First generation (binary)	−0.066 (−0.049)	0.004 (0.017)
Outline videos	−0.084 (−0.068)	<0.001 (<0.001)

The two other measures of student prior preparation, MPE score ($\beta = 0.169$) and BEMA pretest score ($\beta = 0.137$), remained statistically significant predictors of student performance. As in the original model, first-generation status was negatively linked to exam performance ($\beta = -0.049$); however, the utilization of previous years' exams by first-generation students yielded a positive correlation ($\beta = 0.047$) to their performance. The only supplemental video resource that showed statistical significance in the first model after multiple imputation was the chapter outline videos ($\beta = -0.068$). Finally, one notable difference between the original and imputed models was that the latter showed statistical significance ($\beta = 0.112$) for students' SAT-reading scores.

TABLE IV. Contribution to R^2 from significant predictors in multiple regression models.

Variable	First model	Second model	Third model
Calculus 1 grade	0.109	0.114	0.109
Math placement	0.108	0.109	0.108
Prior exams	0.056	0.066	0.067
BEMA Pre	0.022	0.021	0.022

B. Second model: Logarithmic transformations

Due to the observed nonlinearity within the residual plots for all supplemental materials, we then implemented a logarithmic transformation to better understand the relationship between resource utilization and course performance; these variables are marked with an asterisk in Tables VI and VII. As shown in Table V, the sample size for this unimputed model was $N = 1402$. The slight increase in N was due to the removal of first-generation status, as this identifier was missing from 34 students in the database. The first-generation status variable showed no statistical significance within this model and was therefore removed from the reduced model during backward elimination. Table V highlights student demographics in the reduced model with first-generation status removed.

Within this multiple regression model, seen in Table VI, the model had an R-squared value of 0.314, which implies that it accounted for 31.4% of the variability. Contributions to the variability of each variable are shown again in Table IV. This model showed that a student's previous success in calculus was the highest predictor of student performance in introductory electricity and magnetism ($\beta = 0.325$). Also, as previously seen, student usage of supplemental prior exams was the second highest predictor of student performance ($\beta = 0.262$). Two other measures of prior preparation, MPE ($\beta = 0.176$) and BEMA ($\beta = 0.145$), also showed positive correlation with students' performance on exams. Student usage of chapter outline videos showed a negative correlation ($\beta = -0.106$) with their performance, while the supplemental review and recitation-style problem-solving videos showed no significance, similar to what was observed in the first model.

Multiple imputation increased the R^2 of the model to 0.355. Results from the imputed model are shown in parentheses in Table VI. Similar to the original model, students' prior performance in calculus was the strongest predictor of midterm exam performance ($\beta = 0.333$). The second strongest predictor of student performance remained students' utilization of previous years' exams ($\beta = 0.216$). The two other measures of students' prior preparation, MPE score ($\beta = 0.170$) and BEMA pretest score ($\beta = 0.131$), remained statistically significant

TABLE V. Demographics for students with listwise valid data in second model, $N = 1402$.

Demographic	Categories				
Ethnicity	White 53.21%	Hispanic 21.61%	Asian 15.55%	Black 2.92%	Other 6.71%
First generation	Yes 11.91%	No 85.66%	No record 2.43%		
Gender	Male 77.25%		Female 22.75%		

TABLE VI. Descriptive statistics and results for a multiple regression model with midterm exam average as the outcome variable, including logarithmic transformation of supplemental material explanatory variables. Results from imputed models are presented in parentheses.

Descriptive statistics for LOG transform model, $N = 1402$		
Variable	Mean	STD
Exam average	75.809 (73.769)	17.454 (18.901)
Math placement	0.078 (−0.041)	0.869 (1.031)
Calculus 1 grade	0.080 (−0.014)	0.953 (1.014)
BEMA Pre	0.0004 (0.0003)	0.976 (1.001)
SAT reading	−(−0.007)	−(1.003)
Outline videos*	$-0.009\ (-3.404 \times 10^{-17})$	1.001 (1)
Conceptual videos*	$-(-1.915 \times 10^{-16})$	−(1)
Recitation videos*	$-(-6.857 \times 10^{-16})$	−(1)
Prior exams*	$0.092\ (-1.168 \times 10^{-16})$	0.825 (1)

Multiple regression model including logarithmic transform		
$R^2 = 0.314$ (0.355), Outcome var: ME avg		
Independent variable	β	Significance
Calculus 1 grade	0.325 (0.333)	<0.001 (<0.001)
Prior exams*	0.262 (0.216)	<0.001 (<0.001)
Math placement	0.176 (0.170)	<0.001 (<0.001)
BEMA Pre	0.145 (0.131)	<0.001 (<0.001)
SAT reading	−(0.103)	−(< 0.001)
Outline videos*	−0.106(−0.090)	<0.001 (<0.001)
Conceptual videos*	−(−0.045)	−(0.043)
Recitation videos*	−(0.045)	−(0.046)
First generation (binary)	−(−0.040)	−(0.049)

predictors of student performance. Similar to the first model, the imputed model showed statistical significance ($\beta = 0.103$) for students' SAT-reading scores. Notably, first-generation status was negatively linked to exam performance ($\beta = -0.040$), whereas first-generation status was not statistically significant in the original model. In terms of the supplemental resources, all three video resources showed statistical significance after multiple imputation. Chapter outline videos were still negatively correlated with exam performance ($\beta = -0.090$), and conceptual videos also showed a slight negative correlation ($\beta = -0.045$). However, the recitation-style problem-solving videos were positively correlated with student performance on midterm exams ($\beta = 0.045$).

C. Third model: Hybrid methods

In the first model, we found that both first-generation status and the use of supplemental materials by first-generation students were statistically significant predictors of students' average midterm scores. However, in the second model, first-generation status was not statistically significant. To obtain a more complete picture of the key predictors of student success in this course, we built a hybrid model that combined both approaches, specifically interaction terms and logarithmic transformations for the supplemental material variables. Descriptive statistics and results are shown in Table VII. For this model, we had a sample size of 1368, 12.2% of which were first-generation students.

TABLE VII. Descriptive statistics and results for a multiple regression model with midterm exam average as the outcome variable, including logarithmic transformation of supplemental material explanatory variables and first-generation interaction term. Results from imputed models are presented inparentheses.

Descriptive statistics for hybrid model, $N = 1368$		
Variable	Mean	STD
Exam average	75.676 (73.769)	17.472 (18.901)
Math placement	0.073 (−0.041)	0.868 (1.031)
Calculus 1 grade	0.072 (−0.014)	0.953 (1.014)
BEMA Pre	0.0004 (0.0003)	0.979 (1.001)
SAT reading	−(−0.007)	−(1.003)
Outline videos*	$-0.012(-3.404 \times 10^{-17})$	1.0003 (1)
Conceptual videos*	$-(-1.915 \times 10^{-16})$	−(1)
Recitation videos*	$-(-6.857 \times 10^{-16})$	−(1)
Prior exams*	$0.087\ (-1.168 \times 10^{-16})$	0.832 (1)
First generation (binary)	0.122 (0.145)	n/a
First generation × prior exams*	0.232 (0.006)	0.923 (0.930)

Hybrid multiple regression model		
$R^2 = 0.313$ (0.355), Outcome var: ME avg		
Independent variable	β	Significance
Calculus 1 grade	0.314 (0.333)	<0.001 (<0.001)
Prior exams*	0.236 (0.216)	<0.001 (<0.001)
Math placement	0.173 (0.170)	<0.001 (<0.001)
BEMA Pre	0.147 (0.131)	<0.001 (<0.001)
SAT reading	−(0.103)	−(< 0.001)
First generation × prior exams*	0.075 (−)	0.028 (−)
First generation (binary)	−0.092(−0.040)	0.005 (0.049)
Outline videos*	−0.109(−0.90)	<0.001 (<0.001)
Conceptual videos*	−(−0.045)	−(0.043)
Recitation videos*	−(0.045)	−(0.046)

With an $R^2 = 0.313$, the predictive power of this hybrid model is comparable to prior models. As before, each variable's contribution to the variance is presented in Table IV. Notably, a student's calculus grade remains the strongest predictor of exam performance ($\beta = 0.314$), whereas the usage of the supplemental prior exams followed in second ($\beta = 0.236$). Other variables representing prior preparation, MPE ($\beta = 0.173$) and BEMA ($\beta = 0.147$), were positively correlated with performance in our course. First-generation status was negatively correlated with average exam performance ($\beta = -0.092$), indicating that these students scored lower on average compared to their continuing-generation peers. At the same time, first-generation students' usage of prior years' exams showed a positive correlation ($\beta = 0.075$), suggesting that the observed differences could be counteracted by the utilization of this resource. Student usage of chapter outline videos showed a negative correlation ($\beta = -0.109$) with student performance, while the supplemental review and recitation-style problem-solving videos showed no significance. This echoes the results from our prior models.

After multiple imputation, none of the interactions between first-generation status and the usage of supplemental resources showed statistical significance. Due to this, the imputed versions of models 2 and 3 are identical, as seen in Tables VI and VII and discussed in Sec. II B.

For a more fine-grained analysis, the third model was then also employed to analyze specific topics from each of the three midterms, both using listwise deletion (without multiple imputation) and multiple imputation. The data for student usage of supplemental materials were divided into three intervals: prior to the first midterm, between the first and second midterms, and between the second and third midterms. Broadly, results for each of these additional models exhibited results comparable to Table VII; that is, students' calculus grades were the highest predictor of their performance on individual exam items, with math placement and BEMA prescore showing lower standardized coefficients. Also, the use of prior years' exams was the second-highest predictor of performance for each individual topic. Additionally, for a small number of specific topics (Coulomb's Law, Electric Potential, Electric Flux, and TimeDependent Circuits), recitation videos were statistically significant with β ranging from 0.054 to 0.094. In conjunction, first-generation status was a negative predictor for the Coulomb's Law topic under listwise deletion and was observed under imputation for the topics of Electric Field, Capacitors, and Ohm's Law, with β around 0.05 to 0.08, which mirrors the effects seen in models presented in prior sections for overall student performance. Interactions between first-generation student status and use of recitation videos showed positive correlation for topics of Electric Fields with β values of 0.043 (unimputed) and 0.080 (imputed) and for dc circuits with a β of 0.104 for the unimputed model. Interactions with chapter outline video use were inconsistent, ranging from weakly negative to weakly positive ($\beta = 0.05$ to 0.06).

D. Student perceptions through surveys

Anonymous surveys were conducted to gauge students' impressions about the utility and impact of our supplemental self-study resources. The data compiled from all three semesters are discussed below. When asked about their physics experience from high school, Fig. 5, fewer than a quarter of students felt prepared for university-level physics. Meanwhile, roughly 2/3 of students said they did not feel prepared for this course, looking back on their high school experience.

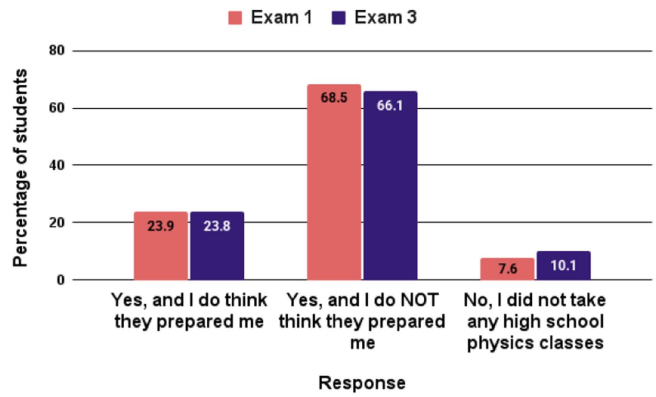


FIG. 5. Student responses to survey question 1: "Did you take any physics classes in high school? If so, do you feel like they prepared you for this class?" Numbers represent the percentages of students who selected each survey item.

The next two questions focused on students' use of online resources to support their learning outside of class time, as shown in Figs. 6 and 7. Almost all students, 99%, indicated that at least a little of their outside learning came from online resources. Over 80% of students reported that most of, if not all, of their outside learning of the material could be credited to our department's supplemental resources. From both of these survey questions, it can be seen that very few students rely purely on lectures and the textbook for their learning.

Students were also asked which of the four types of supplemental resources they utilized, as shown in Fig. 8. Roughly, half of the students used the recitation-style problem-solving videos, and over 30% and 20% of them

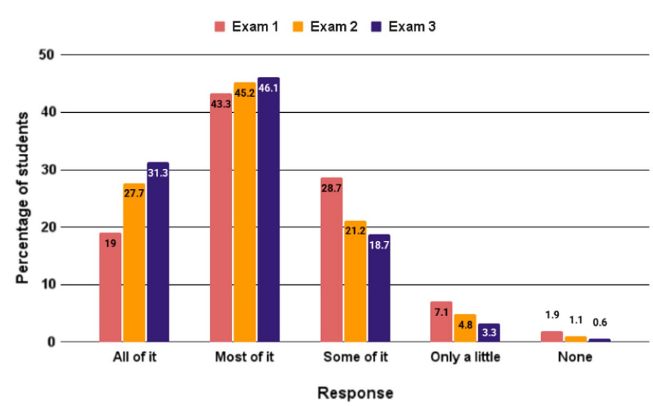


FIG. 6. Student responses to survey question 2: "How much of your learning for this course, outside of class time, do online supplemental resources contribute to?" Numbers on each bar represent the percentage of students who selected that survey item.

used the supplemental concept videos and chapter outline videos, respectively. A much higher percentage of students, over 82%, made use of the repository of previous years' exams over the semester. Roughly half of the students also reported that they made intentional changes to their study habits after using these resources, as shown in Fig. 9.

When asked to what degree students felt the supplemental materials impacted their performance on midterm exams, over 88% of students reported that these materials had a positive impact on their exam performance, and less than 5% of students responded that they had a negative impact, as shown in Fig. 10. When asked about whether or not they would recommend using these materials to their classmates and peers, as shown in Fig. 11, and over 98% of students indicated that they would recommend utilizing the supplemental materials to other students.

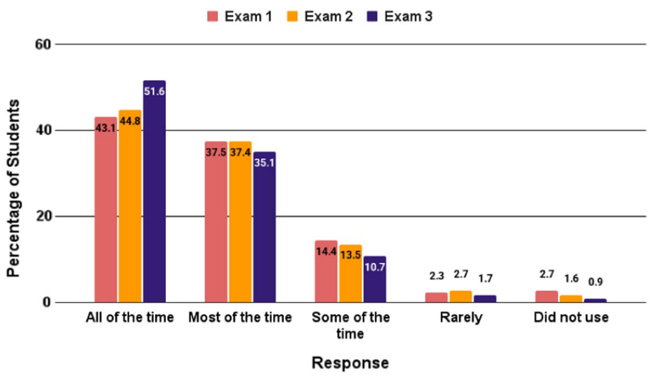


FIG. 7. Student responses to survey question 3: "Did the supplemental resources provided by the department noticeably contribute to your learning of the material outside of class time?" Numbers represent the percentage of students who selected that survey item.

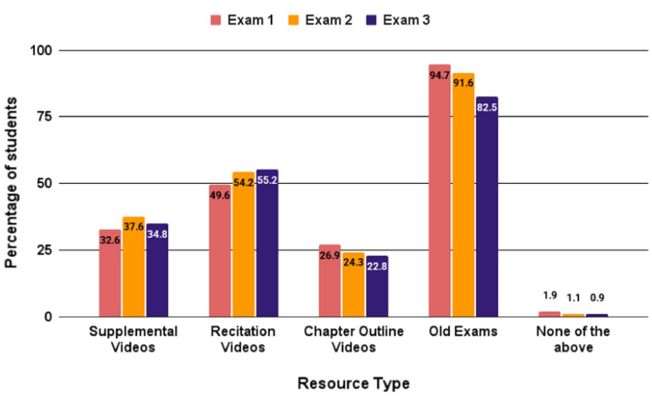


FIG. 8. Student responses to survey question 4: "Which of the following supplemental resources provided by the department did you use? (Select all that apply)" Numbers represent the percentage of students who selected that survey item.

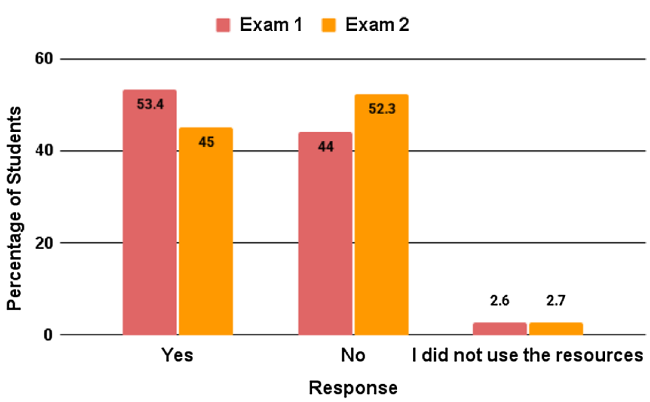


FIG. 9. Student responses to survey question 5: "Did you make any significant or intentional changes to your study methods for physics after using the supplemental resources provided by the department?" Numbers represent the percentage of students who selected that survey item.

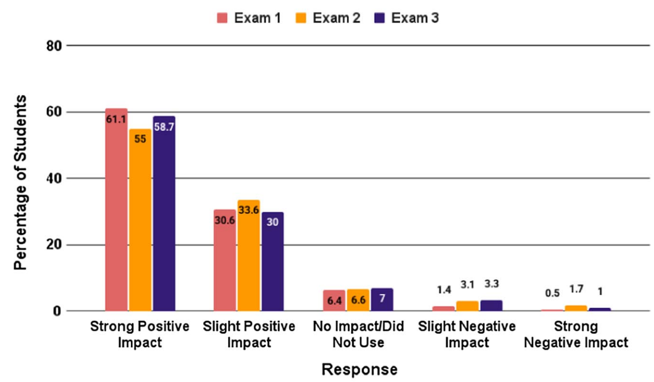


FIG. 10. Student responses to survey question 6: "Did the supplemental resources provided by the department have a noticeable impact on your exam performance?" Numbers represent the percentage of students who selected that survey item.

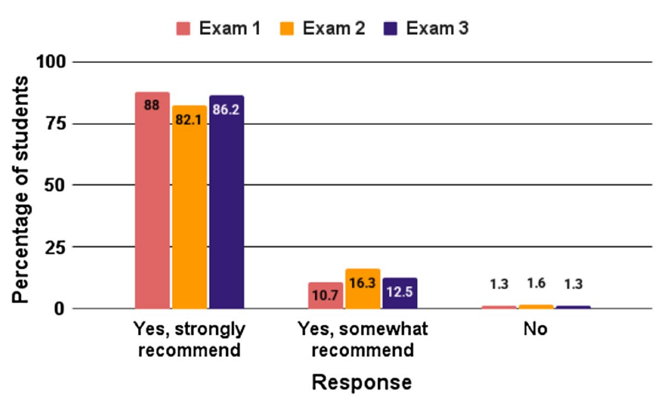


FIG. 11. Student responses to survey question 7: "Would you recommend the supplemental resources provided by the department to other students?" Numbers represent the percentage of students who selected that survey item.

IV. DISCUSSION

Our regression models indicated that students' prior preparation in math and physics was highly predictive of average midterm exam performance. This is similar to prior studies where math preparation was the highest predictor of student learning [9,10,12,14,45–47] and where background physics knowledge was a less significant predictor of student performance [12]. In particular, we observed that students' recent mathematics preparation, based on final letter grade in Calculus I, was the single strongest significant predictor of their performance in a calculus-based introductory electricity and magnetism class at Texas A&M University. Prior physics knowledge and abilities, indicated by conceptual inventory pretest scores, were also a positive predictor of student performance, though they carried less weight than math background, which is consistent with prior literature [9,14]. First-generation student status was observed to be a negative predictor of overall student performance. However, when SAT Reading remained significant after backwards elimination, the beta coefficient for first-generation status became smaller. This shift in significance, from first-generation status to a prior preparation variable (SAT Reading), aligns with previous research from Bui [48], indicating that academic preparation accounts for much of the observed performance gap among first-generation students. Their work shows that first-generation students tend to enter college with lower levels of academic preparation, including lower SAT/ACT scores, fewer advanced high school courses, and less exposure to college-level academic expectations. SAT Reading scores can capture a broad set of academic literacies relevant to physics problem solving, such as interpreting symbolic language, navigating descriptive text, and following multistep reasoning, and therefore could mediate the relationship between first-generation status and exam performance.

Beyond prior preparation, one of the most notable findings is the strong positive correlation between student performance and the use of supplemental prior exams. This resource consistently emerged as the second strongest significant predictor of success, suggesting that students who engage with these materials can greatly benefit, regardless of their initial mathematics or physics background. This is an important result, as more than 83% of students indicated they used this resource according to anonymous surveys. As prior exams offered only answers, not solutions, it is possible that students used this resource to test themselves prior to exams. Student use of repeated testing as a preparation strategy has been shown to positively impact performance, particularly if the exam or assessment does not immediately follow intensive studying (e.g., cramming) [20,49–51]. Additionally, attempting prior versions of exams under testlike conditions can be effective in helping students assess their level of preparation by providing an unbiased, objective measure of their abilities and identifying areas for further study before taking exams [20].

We see some suggestive signals about the positive correlations between students' engagement with video resources and their course performance, which complements prior studies on the impact of video resources [7,24,25]. This is evidenced by the positive relationships exhibited in the imputed versions of the second and third models for the recitation videos. However, there are some methodological differences that can account for results related to the video resources not being consistent across the models. In the cited studies, students were assigned to engage with multimedia materials prior to class, and in at least one study, classroom instruction time was reduced commensurately with the length of the weekly videos. As a result, the videos were a more integral and required component of student learning, rather than optional supplemental materials. Additionally, our results differ slightly from those of Perry *et al.* [28], who observed that the supplemental concept videos included in their study had a positive impact, as measured by correlation coefficients, on student exam performance, particularly for the most conceptually challenging material (e.g., Faraday's law) during the semester. In this work, no consistent relationship between student performance and conceptual videos was observed, though we do observe positive impacts from a subset of recitation videos on specific topics throughout the course. Importantly, none of the above-cited studies tracked student engagement with prior exams, which our analysis found to be the second-strongest predictor of student performance across all regression models. Linde *et al.* [52] highlight survey findings on educators' beliefs that supplemental resources foster the development of problem-solving skills and self-regulated learning. Through interacting with these materials, students can build the skills necessary to succeed not only within the course but throughout their academic and professional careers [29].

Notably, the interaction terms between first-generation status and usage of prior exams in the first models (with and without imputation) and the third model (without imputation) suggest that the effect of using these supplemental materials may be more pronounced for first-generation students compared to their continuing-generation peers. As seen in both the original and imputed first and third models, first-generation status had a slightly negative correlation with students' exam performance. Interestingly, the models suggest that first-generation students who make use of prior exams can effectively eliminate the gap in their exam performance related to their first-generation status alone. It is also possible that utilizing these supplemental resources may aid these students in developing crucial self-study skills. Prior research has shown that while accounting for general high school preparation can explain some differences in first-generation students' course performance, the observed deficits cannot be attributed to mathematics and physics preparation alone [10]. This may be due to the fact that first-generation students often report greater challenges in managing their learning and possessing adequate study skills [15,16]. The implementation and further utilization of these supplemental materials could allow for all students, but especially first-generation students, to develop these crucial skills that are necessary for not just their academic future but for their journey toward becoming lifelong learners.

An unexpected result was the consistent negative beta coefficient associated with student usage of supplemental chapter outline videos across all models, which were aimed at guiding students through the reading of the textbook. Student usage of conceptual videos, designed to explain basic concepts and give brief examples of phenomena in each unit, also showed negative correlations with exam performance in the imputed versions of the second and third models. One possible interpretation is that students relying more heavily on these videos were already struggling academically and seeking additional support in any form available. Alternatively, it is possible that some students were attempting to use these videos as their primary learning resource rather than as a supplement to reading, as intended, and in these cases, resource effectiveness can be reduced significantly [53]. In such cases, their reliance on these materials may indicate a substitution effect, where they engage less with other learning strategies, such as attending lectures, participating in problem-solving exercises, or seeking peer collaboration. Prior research supports this interpretation, indicating that exclusively relying on online videos is less effective than attending regular lectures alone or using videos to complement classroom learning [54]. Live instruction provides opportunities for real-time interaction, clarification, and deeper engagement with course content, elements that may be missing when students depend solely on passive video consumption.

The final video resource, recitation-style problem-solving videos, showed only weak positive correlation with student performance, appearing in the imputed second and third models and for a small number of specific topics. This was unexpected as these videos take an in-depth, step-by-step approach to guiding students through how to set up and develop a complete solution to problems introduced in the class and similar to those on assessments. Although the signals around the video resources studied are mixed, student surveys showed that over 50% of students used the recitation-style problem-solving videos and approximately 30% utilized the supplemental concept videos. These resources also received high levels of approval, with nearly all students indicating they would recommend them to their peers, suggesting a perceived positive impact on the learning process.

Student perceptions gathered via surveys suggest that a large portion of students felt unprepared for their university-level physics courses after their completion of the required high school curriculum. Over 60% of students reported learning from online resources in general; however, we note that a majority of students ($\geq 80\%$) reported that our open-access resources had a noticeable contribution to their learning outside of class. These results align with a central finding by Hill *et al.* [27], which highlights that students who engaged with online learning materials alongside lectures tended to achieve greater conceptual gains compared to those who relied solely on traditional course instruction. A similar proportion of students also claimed that usage of the supplemental resources had a tangible impact on their exam performance. A very encouraging result from our student surveys can be seen in student recommendations; an overwhelming amount of students ($\geq 98\%$) recommend the utilization of the department's open-access materials to their classmates and peers. This aligns with findings from anonymous surveys conducted by Perry *et al.* [28], in which students reported overwhelmingly favorable views regarding the impact of the supplemental concept videos.

In large-enrollment introductory physics courses, where opportunities for personalized instruction are limited, many students struggle to accurately assess their preparedness for examinations. A lack of confidence can exacerbate ineffective study habits and hinder meaningful engagement with course material. To address this challenge, practicing prior exam problems, particularly those designed with a structured format that emphasizes the identification of relevant physical laws and the logical progression of multistep solutions, can be a valuable strategy. This practice complements other formative assessments such as homework assignments and quizzes, reinforcing conceptual understanding and problem-solving strategies essential for success in physics examinations. This approach not only helps students develop their "physics muscles" through active problem-solving but also fosters confidence and provides realistic feedback on their readiness. Students' high ratings of the materials also show their positive attitude toward expert-developed, open-access, and self-paced learning materials, which can be used on demand during their self-directed study time. Creating tailored supplemental, multimedia (exams, videos, etc.) materials that align with the specific content and structure of the course further enhances student learning and supports overall academic success in these demanding instructional settings.

Given the many complex and often unknown variables that influence student success across a course, these results highlight the importance of considering how students engage with different learning resources. Ultimately, our models offer important insight into the predictors of student success in introductory electricity and magnetism, helping to clarify the role of supplemental materials in shaping learning outcomes.

V. LIMITATIONS

Although the regression models in this study indicate positive effects of student use of open-access supplemental resources, several limitations should be considered when interpreting the results. Usage data were based on the number of clicks on individual hyperlinks or visits to specific resource pages. However, the dataset did not capture time spent on each resource or actual engagement levels. Given the constraints of the website infrastructure, page visits were determined to be the most feasible proxy for resource utilization. Additionally, as previously noted, some students were excluded from the final regression models due to missing data, most commonly because calculus grades were only collected for students who completed the course in residence or because some students did not complete the BEMA pretest during the first recitation period.

The aforementioned limitations behind the regression models shown were then mitigated by the usage of multiple imputation. Although the missing data in this study were consistent with Missing Completely at Random, the use of multiple imputation still introduces model-dependent assumptions [43]. Imputed values reflect the structure of the imputation model rather than observed behavior; imputation may smooth the variability in the effects of supplemental resources on students' exam performance.

Variance inflation factors and p values each carry important limitations for interpreting statistical models. VIFs reflect only linear dependence and are sensitive to low-variance or structurally correlated binary predictors, which can lead to elevated values that stem from coding constraints or group imbalance rather than substantive multicollinearity [55]. Likewise, p values combine information about effect size and sample size, depend on model specification, and do not convey the magnitude or practical significance of an effect; they summarize the extremeness of the observed data under a null model rather than the probability that a hypothesis is true [56–58]. As with any statistical analysis, interpreting the results of this study in context is crucial for future applications.

VI. CONCLUSION

We analyzed factors influencing student learning outcomes within an introductory calculus-based electricity and magnetism course, including demographics, prior preparation in mathematics and physics, and usage of open-access supplemental materials. We aggregated data from more than 2000 students over three semesters to construct multiple linear regression models. Three models were run with and without imputation. Consistent with prior studies, our results showed mathematical background, measured by students' overall letter grade in their first-semester calculus course, was the strongest predictor of success, $\beta \sim 0.314 - 0.336$. Importantly, across all three models, the second strongest predictor was student usage of supplemental prior exams $\beta \sim 0.216 - 0.262$, which were used by a majority of students ($> 83\%$). These relative beta values imply that students' use of prior exams can significantly reduce the performance gap predicted by their level of math preparation. This could be due to them enhancing their self-regulated learning skills and continuously developing their problem-solving abilities, which better prepares them to tackle new and unfamiliar problems on the course exams. We observed that first-generation status was negatively correlated with student performance on exams. However, when combined with their usage of supplemental prior exams, first-generation students showed a positive correlation with their performance, demonstrating that student use of this resource helps to bridge observed gaps.

Data from our anonymous surveys indicated that two-thirds of students felt unprepared for this course after their high school physics experience. About half of the students watched problem-solving videos, and approximately one-third of the students watched supplemental review videos. Responses further revealed that over 80% of students who used the supplemental materials believed that they made a noticeable contribution to their learning and more than 88% felt they positively impacted their performance. Remarkably, approximately 98% of surveyed students recommended these materials to their classmates and peers.

These findings underscore the importance of accessible supplemental resources in supporting student populations with diverse backgrounds and levels of preparation, particularly those who face added challenges such as independently managing their learning. Resources like those described in this paper can help students develop the problem-solving and critical thinking skills essential for long-term success and contribute to retention in STEM programs. As universities refine instructional strategies, integrating structured self-study tools, such as archived exams and instructional supplemental videos, may play an important role in promoting equitable outcomes in physics education. Future research should explore how variations in resource design and usage patterns affect long-term learning, ensuring that supplemental materials continue to evolve as effective and inclusive tools. Moving forward, we plan to conduct a similar analysis for our introductory calculus-based mechanics course.

ACKNOWLEDGMENTS

We are grateful to A. Lewis Ford and Ryan Carmichael for their assistance with collecting and synthesizing the data for this study. This work was partially supported by Texas A&M University and the Texas Higher Education Coordinating Board Digital Design for Student Success (D2S2) program.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available because they contain sensitive personal information. The data are available from the authors upon reasonable request.

[1] C. Nord, S. Roey, R. Perkins, M. Lyons, N. Lemanski, J. Brown, and J. Schuknecht, The nation's report card: America's high school graduates. U.S. Department of Education, National Center for Education Statistics, Washington, DC, Tech. Rep.NCES 2011462, 2011.

[2] C. Good, A. Rattan, and C. S. Dweck, Why do women opt out? Sense of belonging and women's representation in mathematics, J. Pers. Soc. Psychol. **102**, 700 (2012).

[3] A. M. Porter, R. Y. Chu, and R. Ivie, Attrition and persistence in undergraduate physics programs: a five-year longitudinal study (2018–2023), Tech. Rep. (AIP, College Park, Maryland, 2024.

[4] R. W. Lent, S. D. Brown, and K. C. Larkin, Relation of self-efficacy expectations to academic achievement and persistence, J. Counsel. Psychol. **31**, 356 (1984).

[5] R. W. Lent, S. D. Brown, and K. C. Larkin, Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking, J. Counsel. Psychol. **34**, 293 (1987).

[6] C. Lindstrom and M. D. Sharma, Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter?, Int. J. Innov. Sci. Math. Educ. **19**, 1 (2011).

[7] H. R. Sadaghiani, Using multimedia learning modules in a hybrid-online course in electricity and magnetism, Phys. Rev. STPhys. Educ. Res. **7**, 010102 (2011).

[8] C. J. Ballen, C. Wieman, S. Salehi, J. B. Searle, and K. R. Zamudio, Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning, CBE—Life Sci. Educ. **16**, ar56 (2017).

[9] S. Salehi, E. Burkholder, G. P. Lepage, S. Pollock, and C. 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).

[10] J. Stewart, G. L. Cochran, R. Henderson, C. Zabriskie, S. DeVore, P. Miller, G. Stewart, and L. Michaluk, Mediational effect of prior preparation on performance differences of students underrepresented in physics, Phys. Rev. Phys. Educ. Res. **17**, 010107 (2021).

[11] P. N. Iwuanyanwu, Facilitating problem solving in a university undergraduate physics classroom: The case of students' self-efficacy, Interdiscip. J. Environ. Sci. Educ. **18**, e2270 (2022).

[12] D. E. Meltzer and D. H. King, Pre-instruction diagnostic tests can predict grade probabilities in introductory physics, Am. J. Phys. **93**, 528 (2025).

[13] P. M. Sadler and R. H. Tai, Success in introductory college physics: The role of high school preparation, Sci. Educ. **85**, 111 (2001).

[14] L. E. Kost, S. J. Pollock, and N. D. Finkelstein, Characterizing the gender gap in introductory physicsPhys. Rev. Spec. Top.—Phys. Educ. Res. **5**, 010101 (2009).

[15] M. J. Stebleton and K. M. Soria, Breaking down barriers: Academic obstacles of first-generation students at research universities, Learn. Assistance Rev. **17**, 7 (2012).

[16] L. Falcon, Breaking down barriers: First-generation college students and college success, Innovation Showcase **10** (2015).

[17] D. L. Schwartz and J. D. Bransford, A time for telling, Cognit. Instr. **16**, 475 (1998).

[18] N. R. Council, *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*, edited by S. R. Singer, N. R. Neilsen, and H. A. Schweingruber (The National Academies Press, Washington, DC, 2012).

[19] N. R. Council, *Adapting to a Changing World: Challenges and Opportunities in Undergraduate Physics Education* (The National Academies Press, Washington, DC, 2013).

[20] J. Mestre and J. Docktor, *Science Of Learning Physics, The: Cognitive Strategies For Improving Instruction* (World Scientific Publishing Company, Singapore, 2020).

[21] C. Dolch, O. Zawacki-Richter, M. Bond, and V. I. Marín, Higher education students' media usage: A longitudinal analysis, Asian J. Distance Educ. **16**, 31 (2021).

[22] T. Stelzer, G. Gladding, J. P. Mestre, and D. T. Brookes, Comparing the efficacy of multimedia modules with traditional textbooks for learning introductory physics content, Am. J. Phys. **77**, 184 (2009).

[23] P. W. Laws, M. C. Willis, D. P. Jackson, K. Koenig, and R. Teese, Using research-based interactive video vignettes to enhance out-of-class learning in introductory physics, Phys. Teach. **53**, 114 (2015).

[24] J. C. Moore, Efficacy of multimedia learning modules as preparation for lecture-based tutorials in electromagnetism, Educ. Sci. **8**, 23 (2018).

[25] Z. Chen, T. Stelzer, and G. Gladding, Using multimedia modules to better prepare students for introductory physics lecture, Phys. Rev. ST.—Phys. Educ. Res. **6**, 010108 (2010).

[26] W. Fakcharoenphol, E. Potter, and T. Stelzer, What students learn when studying physics practice exam problems, Phys. Rev. ST Phys. Educ. Res. **7**, 010107 (2011).

[27] M. Hill, M. D. Sharma, and H. Johnston, How online learning modules can improve the representational fluency and conceptual understanding of university physics students, Eur. J. Phys. **36**, 045019 (2015).

[28] J. D. Perry, T. L. Erukhimova, and W. H. Bassichis, New video resource for calculus-based introductory physics, design and assessment. I. Electricity and magnetism, Am. J. Phys. **87**, 335 (2019).

[29] C. Singh and D. Haileselassie, Developing problem-solving skills of students taking introductory physics via web-based tutorials, J. Coll. Sci. Teach. **39**, 42 (2010).

[30] C. Choirudin, M. Lubis, and M. A. Masuwd, Enhancing high school students' mathematical problem-solving skills through interactive media: A classroom action research approach, J. Teach. Learn. Math. **2**, 104 (2025).

[31] B. R. Wilcox and M. Vignal, Understanding the student experience with emergency remote teaching, PERC Proc. **2020**, 581 (2020).

[32] M. Dew, L. Ford, D. T. Nodurft, T. Erukhimova, and J. Perry, Student responses to changes in introductory physics learning due to the COVID-19 pandemic, Phys. Teach. **59**, 162 (2021).

[33] M. Hill, M. Sharma, and Y. Xu, Pre-lecture online learning modules in university physics: Student participation, perceptions, and performance, Int. J. Innovation Sci. Math. Educ. **25**, 14 (2017).

[34] M. J. Cahill, K. M. Hynes, R. Trousil, L. A. Brooks, M. A. McDaniel, M. Repice, J. Zhao, and R. F. Frey, Multiyear, multi-instructor evaluation of a large-class interactive-engagement curriculum, Phys. Rev. St Phys. Educ. Res. **10**, 020101 (2014).

[35] L. Ding, R. Chabay, B. Sherwood, and R. Beichner, Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, Phys. Rev. ST Phys. Educ. Res. **2**, 010105 (2006).

[36] J. L. Docktor, J. Dornfeld, E. Frodermann, K. Heller, L. Hsu, K. A. Jackson, A. Mason, Q. X. Ryan, and J. Yang, Assessing student written problem solutions: A problem-solving rubric with application to introductory physics, Phys. Rev. Phys. Educ. Res. **12**, 010130 (2016).

[37] Texas A&M intro physics supplemental materials, classroom.physics.tamu.edu (2019).

[38] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Erlbaum Associates, New York City, NY, 1977).

[39] See Supplemental Material at http://link.aps.org/supplemental/10.1103/p4jl-cjjd for plots related to assumption testing for regression models.

[40] I. Pardoe, L. Simon, and D. Young, Regression methods: 12.4—detecting multicollinearity using variance inflation factors (2024).

[41] J. J. Faraway, *Linear Models with R* (Chapman & Hall/CRC, New York City, NY, 2009).

[42] M. Lunt, Introduction to statistical modelling 2: Categorical variables and interactions in linear regression, Rheumatology **54**, 1141 (2015).

[43] S. v. Buuren, *Flexible Imputation of Missing Data* (CRC Press, Boca Raton, FL, 2012).

[44] E. Brewe, V. Sawtelle, L. H. Kramer, G. E. O'Brien, I. Rodriguez, and P. Pamelá, Toward equity through participation in Modeling Instruction in introductory university physics, Phys. Rev. St Phys. Educ. Res. **6**, 010106 (2010).

[45] D. D. Long, G. W. McLaughlin, and A. M. Bloom, The influence of physics laboratories on student performance in a lecture course, Am. J. Phys. **54**, 122 (1986).

[46] G. E. Hart and P. D. Cottle, Academic backgrounds and achievement in college physics, Phys. Teach. **31**, 470 (1993).

[47] 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 (2010).

[48] K. V. T. Bui, First-generation college students at a four-year university: Background characteristics, reasons for pursuing higher education, and first-year experiences, Coll. Stud. J. **36**, 3 (2002).

[49] M. Wheeler, M. Ewers, and J. Buonanno, Different rates of forgetting following study versus test trials, Memory **11**, 571 (2003).

[50] H. L. Roediger III and J. D. Karpicke, Test-enhanced learning: Taking memory tests improves long-term retention, Psychol. Sci. **17**, 249 (2006).

[51] J. D. Karpicke and H. L. Roediger III, Repeated retrieval during learning is the key to long-term retention, J. Mem. Lang. **57**, 151 (2007).

[52] I. Linde, L. Hačatrjana, and L. Daniela, Teachers' feedback on the use of supplementary materials for developing students' self-regulated learning and problem solving skills, in *82nd International Scientific Conference of the University of Latvia* (2024), pp. 20–33

[53] S. DeVore, E. Marshman, and C. Singh, Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics, Phys. Rev. St Phys. Educ. Res. **13**, 010127 (2017).

[54] M. Meehan and J. McCallig, Effects on learning of time spent by university students attending lectures and/or watching online videos, J. Comput. Assist. Learn. **35**, 283 (2018).

[55] R. Salmerón-Gómez, C. B. García-García, and J. García-Pérez, A redefined variance inflation factor: Overcoming the limitations of the variance inflation factor: A redefined variance inflation factor: R. Salmerón-Gómez *et al.*, Comput. Econ. **65**, 337 (2025).

[56] R. S. Nickerson, Null hypothesis significance testing: A review of an old and continuing controversy, Psychol. Methods **5**, 241 (2000).

[57] A. Gelman and H. Stern, The difference between "significant" and "not significant" is not itself statistically significant, Am. Stat. **60**, 328 (2006).

[58] R. L. Wasserstein and N. A. Lazar, The ASA statement on p-values: Context, process, and purpose, Am. Stat. **70**, 129 (2016).