The Princeton Review SAT® Prep Effectiveness Study (2024)

Study Type: ESSA Evidence Level II

Prepared for:
Princeton Review

Prepared by LearnPlatform by Instructure:
Meetal Shah, Ph.D., Senior Researcher
Alexandra Lee, Ph.D., Researcher
Christina Davis, M.P.A., Researcher

September 2024

EXECUTIVE SUMMARY

The Princeton Review® contracted with LearnPlatform by Instructure, a third-party edtech research company, to examine the impact of The Princeton Review SAT® Prep on high school student outcomes. LearnPlatform by Instructure designed this study to satisfy Level II requirements (Moderate Evidence) of the Every Student Succeeds Act (ESSA).

Study Sample, Measures, and Methods

This quasi-experimental study occurred during the spring 2024 semester and included a matched analysis sample of 1,252 high school students (626 treatment, 626 comparison) from across 15 schools in one school district in Texas. The sample was predominantly Hispanic (61%), followed by white (22%), African American (7%), Asian (6%), and multi-racial (4%). In terms of socioeconomic status (SES), this sample was classified as 26% economically disadvantaged. Three percent of the sample has English language learner (ELL) designation, and 58% of the sample identified as female.

Researchers used administrative and standardized assessment data to examine the impact of The Princeton Review SAT Prep on student outcomes. Descriptive statistics were used to examine participant characteristics and understand program implementation. Regression analyses were used to study impact on outcomes. For impact analysis, researchers created matched samples and conducted baseline equivalence testing. All analyses met What Works Clearinghouse (WWC) baseline equivalence standards. Researchers also included student-level covariates to control for potential selection bias.

Program Implementation and Student Findings

Student usage. In spring 2024, participating students (*n* = 626) each attended an average of four synchronous SAT Prep virtual sessions with The Princeton Review.

Student outcomes. Researchers examined the association between different attendance levels in The Princeton Review SAT Prep program sessions and student outcomes. There was a positive statistically significant association between different attendance levels for The Princeton Review SAT Prep sessions and students' spring 2024 SAT scores. That is, students who attended more than 4 sessions (in total) had higher GPAs than students who attended fewer than 4 sessions. Next, researchers examined the impact of The Princeton Review SAT Prep on student outcomes by comparing The Princeton Review SAT Prep users to a matched (statistically similar) sample of non-users. On average, students who used The Princeton Review SAT® Prep had statistically significant higher spring 2024 SAT scores than students who did not use the program (controlling for baseline achievement and significant covariates).

Conclusion

Given positive outcome findings, this study meets ESSA evidence requirements for Level II (Moderate Evidence). Specifically, this quasi-experimental study was properly designed and implemented,

documented baseline equivalence, included statistical controls, had more than 350 students across multiple schools, and had multiple positive statistically significant findings.

ESSA Study Key Takeaways

Students attended multiple SAT® Prep sessions with The Princeton Review.

Participating students each attended an average of 4 SAT Prep sessions with The Princeton Review.

The Princeton Review SAT Prep had a positive impact on students' spring 2024 SAT® scores.

Students who attended more SAT Prep sessions with The Princeton Review had higher spring 2024 SAT scores. This result was statistically significant ($r = .12$; $p = .003$).

Students who attended SAT Prep sessions with The Princeton Review had higher spring 2024 SAT scores than non-users. This result was statistically significant (Hedges' $g = 0.16$; $p < .001$).

Table of Contents

Introduction

The Princeton Review recognizes that students often need specific, targeted guidance to understand and perform well on complex standardized tests. The Princeton Review provides test preparation for a wide range of exams, the including SAT® and ACT®. The company also provides career and education services and skill-building to ensure that students access tools that equip them to succeed in achieving academic, post-secondary, and professional goals. The resources and services provided by The Princeton Review are designed to help all students in a learning community achieve positive learning and assessment outcomes (see logic model in Appendix A).

As part of its ongoing efforts to demonstrate effectiveness, The Princeton Review contracted with LearnPlatform by Instructure, a third-party edtech research company, to examine the impact of The Princeton Review SAT® Prep on student outcomes. LearnPlatform by Instructure designed this study to satisfy Level II requirements (Moderate Evidence) of the Every Student Succeeds Act (ESSA).

The study had the following research questions:

Implementation Research Questions

1. To what extent did students engage with The Princeton Review SAT Prep?
   a. How many SAT Prep sessions with The Princeton Review did students complete?

Effectiveness Research Questions

2. What was the association between different attendance levels in The Princeton Review SAT Prep program and students' spring 2024 SAT scores?
3. What was the impact of The Princeton Review SAT Prep on spring 2024 SAT scores for students who used the program compared to those who did not?

Methods

This section of the report briefly describes the study design, setting and participants, measures, analysis methods, and baseline equivalence practices.

Study Design

This study used a quasi-experimental design[1] with propensity score matching to align with ESSA Level II evidence standards. The treatment group included students who used The Princeton Review SAT Prep during spring 2024. The comparison group included students who did not access The Princeton Review SAT Prep but who were demographically similar, with comparable fall 2024 achievement.

Setting and Participants

This study occurred during the spring 2024 semester and included a matched analysis sample of 1,252 high school students (626 treatment, 626 comparison) from across 15 schools in one school district in Texas. Any student who was provided access to The Princeton Review SAT® Prep program (n = 943) was considered a user; 260 of these students were missing baseline and/or outcome data and therefore were dropped from the sample. Of the remaining 643 students, 17 students were dropped during propensity score matching, resulting in a sample of 626 students who used the program.

The sample was predominantly Hispanic (61%), followed by white (22%), African American (7%), Asian (6%), and multi-racial (4%); 58% of the sample identified as female. In terms of socioeconomic status (SES), this sample was classified as 26% economically disadvantaged. Three percent of the sample has English language learner (ELL) designation.

Measures

This study includes the following measures to provide insights into the impact of The Princeton Review SAT Prep on student outcomes.

The Princeton Review SAT Prep Usage Metrics. Researchers utilized spring 2024 student-level usage data (i.e., total sessions attended). Usage data informed whether and the extent to which students used The Princeton Review SAT Prep during the spring semester.

Student Outcomes. Researchers used administrative data, specifically fall 2023 PSAT™/SAT scores as a measure of students' baseline achievement and spring 2024 SAT scores to evaluate student outcomes.

---

[1] A quasi-experimental design compares the outcomes of The Princeton Review SAT® Prep users to non-users.

Data Analysis

Researchers conducted descriptive statistics to describe participant characteristics and support implementation analyses. Researchers then conducted regressions to examine The Princeton Review SAT Prep outcomes and impacts. All analyses included fall 2023 PSAT/SAT scores as a baseline measure and student-level covariates. Researchers also used nearest neighbor propensity score matching to account for potential selection bias for the comparative model. In addition, researchers calculated standardized effect sizes (i.e., Hedges' g) to determine the magnitude of difference in The Princeton Review SAT Prep student outcomes compared to statistically similar non-users.

Baseline Equivalence

To ensure the validity to the study's findings and adhere to ESSA Level II standards, researchers conducted baseline equivalence tests on matched treatment and comparison student samples. These findings are discussed in the Student Findings section and referenced in Appendix B.

Program Implementation

This section presents descriptive findings for The Princeton Review SAT Prep implementation. Researchers analyzed program usage metrics to determine the extent to which students used The Princeton Review SAT Prep during spring 2024.

Students at the district were provided access to The Princeton Review SAT Prep program suite of online resources for one full year from the time of their enrollment in the program in February 2024. In addition to the platform resources (such as practice tests and drills) that students complete asynchronously, they also had access to live instruction for their courses, which ran from late February to early April 2024. Depending on students' class schedule, they met synchronously for test preparation sessions once or twice a week for anywhere between 1.5 and 3 hours.

To what extent did students engage with The Princeton Review SAT Prep program? How many SAT Prep sessions with The Princeton Review did students complete?

Among students who were provided access to The Princeton Review SAT Prep program, students attended an average of four Princeton Review SAT® Prep sessions (SD = 3.5 sessions).

Fifty-seven percent of students attended 4 or more SAT® Prep sessions with The Princeton Review.

Number of students

Low Use 0−3	Moderate Use 4−7	High Use 8−10
272	184	170

Number of Princeton Review SAT® Prep sessions attended
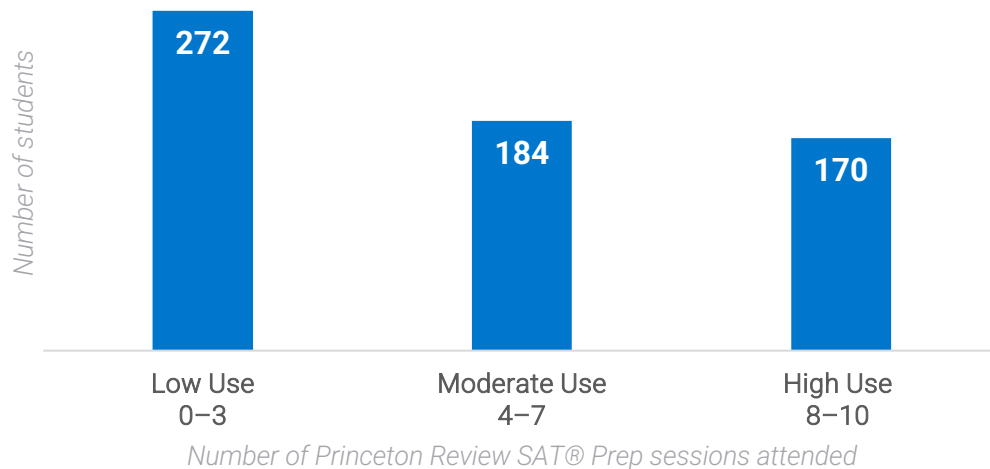
Student Findings

To answer effectiveness research questions, researchers conducted two sets of regressions: (1) for The Princeton Review SAT Prep users only (i.e., correlative analysis), and (2) regressions using matched samples of users to non-users (i.e., comparative analysis). The following section details the findings for both sets of regressions: (a) the association between different attendance levels in The Princeton Review SAT Prep program and student outcomes, and (b) the impact of The Princeton Review SAT Prep program on student outcomes. Researchers report statistically significant findings at the $p < .05$ level. To allow for better interpretability of results, marginal means charts are presented below. The orange vertical lines at the top of each bar represent a 95% confidence interval. (See Appendix B for more details about the model and the corresponding Hedges' g effect sizes.)

What was the association between different attendance levels in The Princeton Review SAT Prep program and students' spring 2024 SAT scores?

Researchers conducted regressions for The Princeton Review SAT Prep students, exploring the association between different attendance levels in The Princeton Review SAT Prep program and spring 2024 SAT scores. Models included fall 2023 PSAT/SAT test scores, students' socioeconomic status, and gender as covariates.

Finding 1. Students who attended 4–7 sessions (moderate use) and 8–10 sessions (high use) had significantly higher spring 2024 SAT scores than students who attended 0–3 sessions (low use; Figures 1). These results were statistically significant.

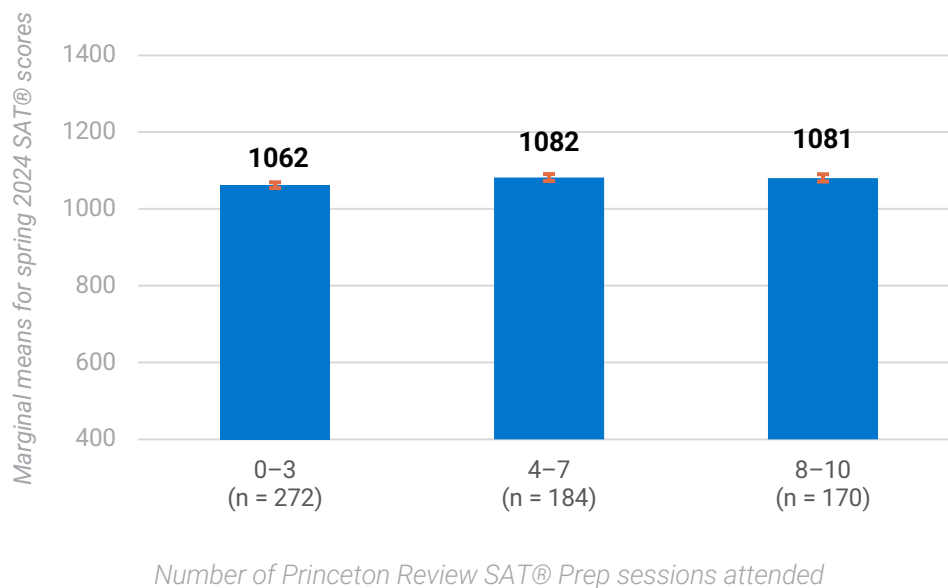

Number of Princeton Review SAT® Prep sessions attended

Figure 1. Association between different attendance levels in the Princeton Review SAT Prep program and users' spring 2024 SAT scores (n = 626). Mean differences between low (0–3 sessions), moderate (4–7 sessions), and high (8–10 sessions) usage groups were statistically significant (g = 0.13, p = .001; g = 0.13, p = .001).

What was the impact of The Princeton Review SAT Prep on spring 2024 SAT scores for students who used the program compared to those who did not?

Researchers created a matched sample of The Princeton Review SAT Prep users and non-users based on students' fall PSAT/SAT scores, school, race, socioeconomic status, and gender. The matched sample demonstrated baseline equivalence (Hedges' g = 0.01). Next, researchers used the matched sample to run a regression model, examining the impact of participating in The Princeton Review SAT Prep on students' spring 2024 SAT scores, controlling for fall PSAT/SAT scores and other statistically significant demographic variables (i.e., socioeconomic status and gender). The Princeton Review SAT Prep users had higher spring 2024 SAT scores compared to non-users, and this difference was statistically significant (g = 0.16; p < .001); Figure 2).

Finding 2. The Princeton Review SAT Prep users had higher spring 2024 SAT scores compared to non-users. This difference was statistically significant.
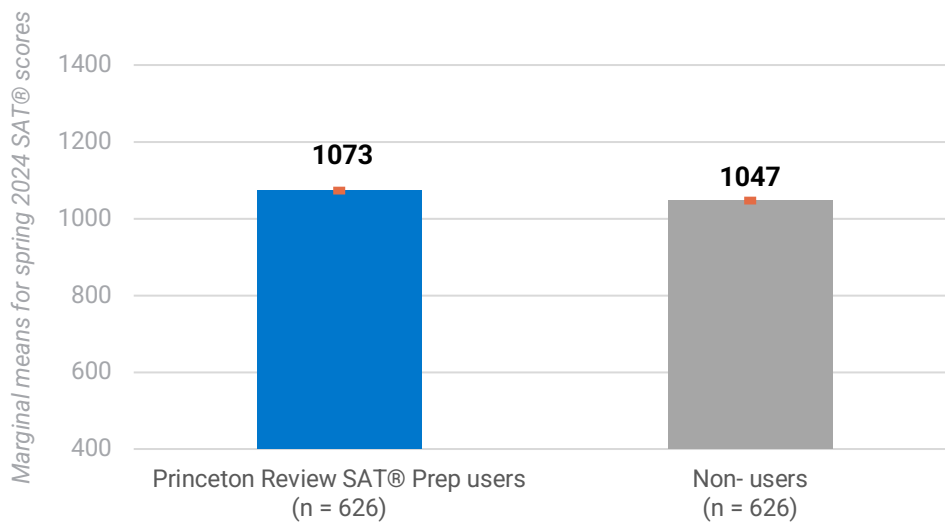
Figure 2. Adjusted mean spring 2024 SAT scores for The Princeton Review SAT Prep users and non-users. Mean difference was statistically significant (g = 0.16, p < .001; n = 1,252).

Conclusions

Given multiple positive outcome findings, this study provides results to satisfy ESSA evidence requirements for Level II (Moderate Evidence). Specifically, this quasi-experimental study met the following criteria for Level II:

✔ Proper design and implementation

✔ Baseline equivalence for treatment and comparison groups

✔ Statistical controls through covariates

✔ At least 350 students in the analysis sample

✔ Representative, multi-site study

✔ At least one statistically significant, positive finding

Appendix A. The Princeton Review SAT Prep Logic Model

The Princeton Review
Logic Model

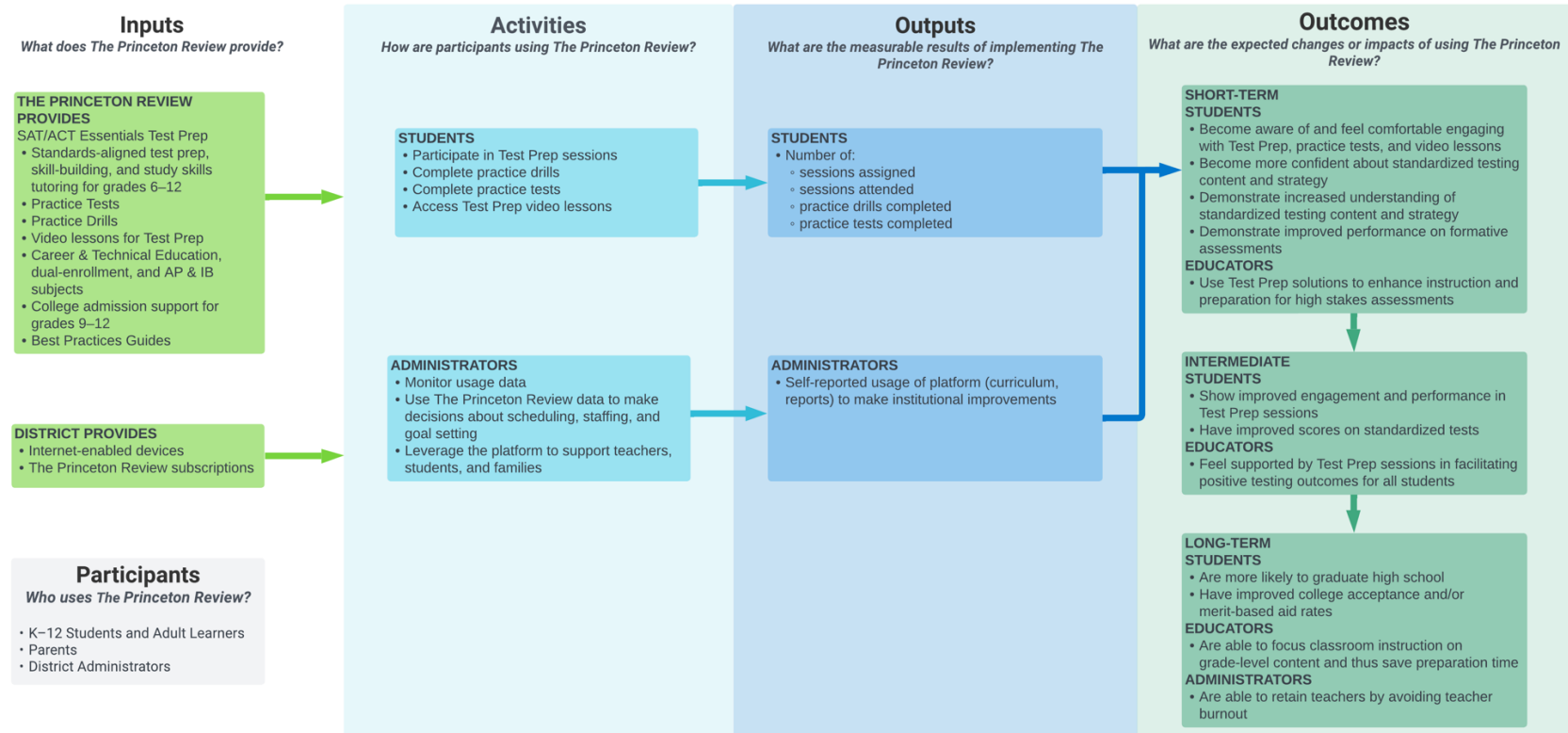**Problem Statement**: Students often need specific, targeted guidance to understand and perform well on complex standardized tests. The Princeton Review provides test preparation for a wide range of exams, the including SAT® and ACT®. The company also provides career and education services and skill-building to ensure that students access tools that equip them to succeed in achieving academic, post-secondary, and professional goals. The resources and services provided by The Princeton Review are designed to help all students in a learning community achieve positive learning and assessment outcomes.

Inputs
What does The Princeton Review provide?

THE PRINCETON REVIEW PROVIDES
SAT/ACT Essentials Test Prep
• Standards-aligned test prep, skill-building, and study skills tutoring for grades 6–12
• Practice Tests
• Practice Drills
• Video lessons for Test Prep
• Career & Technical Education, dual-enrollment, and AP & IB subjects
• College admission support for grades 9–12
• Best Practices Guides

DISTRICT PROVIDES
• Internet-enabled devices
• The Princeton Review subscriptions

Participants
Who uses The Princeton Review?

• K–12 Students and Adult Learners
• Parents
• District Administrators

Activities
How are participants using The Princeton Review?

STUDENTS
• Participate in Test Prep sessions
• Complete practice drills
• Complete practice tests
• Access Test Prep video lessons

ADMINISTRATORS
• Monitor usage data
• Use The Princeton Review data to make decisions about scheduling, staffing, and goal setting
• Leverage the platform to support teachers, students, and families

Outputs
What are the measurable results of implementing The Princeton Review?

STUDENTS
• Number of:
  ◦ sessions assigned
  ◦ sessions attended
  ◦ practice drills completed
  ◦ practice tests completed

ADMINISTRATORS
• Self-reported usage of platform (curriculum, reports) to make institutional improvements

Outcomes
What are the expected changes or impacts of using The Princeton Review?

SHORT-TERM
STUDENTS
• Become aware of and feel comfortable engaging with Test Prep, practice tests, and video lessons
• Become more confident about standardized testing content and strategy
• Demonstrate increased understanding of standardized testing content and strategy
• Demonstrate improved performance on formative assessments
EDUCATORS
• Use Test Prep solutions to enhance instruction and preparation for high stakes assessments

INTERMEDIATE
STUDENTS
• Show improved engagement and performance in Test Prep sessions
• Have improved scores on standardized tests
EDUCATORS
• Feel supported by Test Prep sessions in facilitating positive testing outcomes for all students

LONG-TERM
STUDENTS
• Are more likely to graduate high school
• Have improved college acceptance and/or merit-based aid rates
EDUCATORS
• Are able to focus classroom instruction on grade-level content and thus save preparation time
ADMINISTRATORS
• Are able to retain teachers by avoiding teacher burnout

LEARN PLATFORM
BY INSTRUCTURE

Appendix B. Additional Information on Study Design and Methods

Analytic Sample. A total of 643 Princeton Review SAT Prep students had complete demographic, achievement, and usage data. Researchers had a sample of 3,019 comparison students with complete demographic and achievement data. Using these this dataset of 643 The Princeton Review SAT Prep students and 3,019 non–The Princeton Review SAT Prep students, researchers used propensity score matching (i.e., nearest neighbor matching without replacement) to create a matched study sample of 622 The Princeton Review SAT Prep students and 622 non–The Princeton Review SAT Prep students with similar fall PSAT/SAT test scores, school, race, socioeconomic status, and gender. Consequently, the matched study sample included 1,252 students. Additional demographic information on participating students is included in Table B1.

Table B1. Student demographics by group for matched sample

Characteristic	The Princeton Review SAT Prep students (*n* = 626) Percent	*n*	Non–The Princeton Review SAT Prep students (*n* = 626) Percent	*n*	Total sample (*n* = 1,252) Percent	*n*
Race/Ethnicity						
Hispanic	58%	360	64%	399	61%	759
Asian	6%	40	6%	38	6%	78
Two or More	4%	26	4%	27	4%	53
White	22%	137	22%	139	22%	276
Black	10%	60	4%	22	7%	82
Socioeconomic Status (economically disadvantaged)						
Yes	26%	165	25%	156	26%	321
No	74%	461	75%	470	74%	931
Gender						
Female	59%	368	58%	364	58%	732
Male	41%	258	42%	262	42%	520
English Language Learner						
Yes	3%	19	4%	22	3%	41
No	97%	607	96%	604	97%	1,211

After generating the matched sample, researchers conducted regression analyses using the propensity score matched sample to confirm there were no statistically significant differences between student groups. The matched sample met WWC baseline equivalence standards with Hedges' *g* effect size 0.01 (*p* = .812; Table B2).

Table B2. Baseline equivalence analysis of fall 2023 PSAT/SAT scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	p-value
Fall 2023 PSAT/SAT test scores (Hedges' g effect size = 0.01)	1.80	7.56	0.24	.812
School	-2.41	0.89	-2.72	.007
Race	13.22	2.08	6.35	<.001
ELL	-109.78	21.73	-5.05	<.001
SES	-36.43	8.95	-4.07	<.001
Gender	26.50	7.69	3.45	.001

Table B3. Descriptive statistics for the usage categories for The Princeton Review SAT Prep sessions

Usage categories: total sessions attended		n	Mean	SD
Low	0−3 sessions	272	1.1	1.1
Moderate	4−7 sessions	184	5.3	1.1
High	8−10 sessions	170	9.1	0.8

Overall Relationship Between Total Sessions Attended and Students' Spring 2024 SAT® Scores

Table B4. Students' spring 2024 SAT scores by The Princeton Review SAT Prep sessions attended

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	p-value
Moderate Use vs. Low Use (Hedges' g = 0.13)	20.58	5.92	3.48	<.001
High Use vs. Low Use (Hedges' g = 0.13)	19.46	6.06	3.21	<.001
High Use vs. Moderate Use (Hedges' g = -0.01)	-1.12	6.59	-0.17	.865
Fall 2023 P/SAT scores	1.04	0.02	56.77	<.001
Socioeconomic Status	-13.92	5.70	-2.44	.015
Gender	10.29	5.06	2.04	.042

Difference Between Students who used The Princeton Review SAT Prep and Students Who Did Not Use the Program

Researchers conducted regression analysis to examine whether there were any differences between students who used The Princeton Review SAT Prep during spring 2024 and a statistically similar sample of students who did not use the program. In addition to the outcome of interest (i.e., spring 2024 SAT scores), the model included the student group, fall 2023 PSAT/SAT scores, and student demographic characteristics (Table B5).

Table B5. Differences between spring 2024 SAT scores by student group

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Students who used The Princeton Review SAT Prep vs. Students who did not use the program (Hedges' *g* = 0.16)	26.10	3.63	7.19	<.001
Fall 2023 PSAT/SAT scores	1.03	0.01	77.93	<.001
Socioeconomic status	11.26	4.22	-2.67	.008
Gender	10.21	3.70	2.76	.006