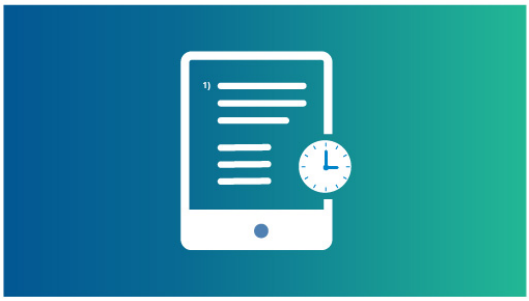


LSAT Takers and Khan Academy Preparation

Kimberly Dustman
Law School Admission Council

Gregory Camilli
Law School Admission Council

Ann Gallagher
Law School Admission Council

April 2021

LSAT Takers and Khan Academy Preparation

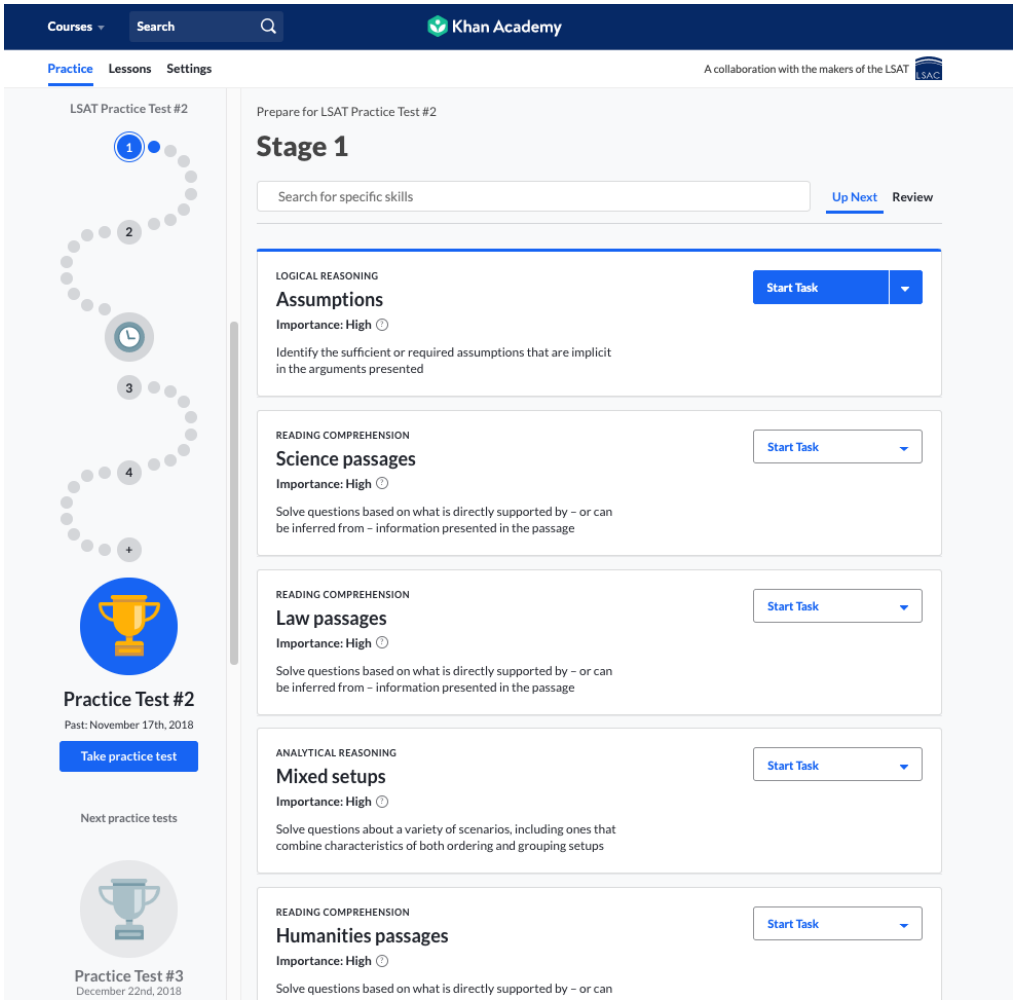
Abstract

Free online Law School Admission Test (LSAT) preparation resources from the Law School Admission Council and Khan Academy have been widely utilized by LSAT and LSAT-Flex test takers. In September 2020, nearly 70,000 individuals engaged with Khan Academy's Official LSAT® Prep platform. The purpose of this study was to examine the potential effects of engagement on actual LSAT performance. Our analyses showed that a higher level of engagement (measured in terms of practice time and number of practice exams taken) was associated with higher performance on the LSAT. These results held not only for the overall population but also across multiple demographic subgroups. The results also showed that the performance of test takers with lower initial practice exam scores was associated with slightly higher LSAT score gains per practice minute, indicating that these students benefitted at least as much as students who scored higher initially. Because this was a quasi-experimental controlled study, the possibility of alternative influences on LSAT performance cannot be ruled out. However, we believe that engagement with the Khan Academy platform is currently the best explanation for the LSAT score increases observed in this study.

Introduction

Free online Law School Admission Test (LSAT) preparation resources from the Law School Admission Council (LSAC) and Khan Academy have been widely utilized by LSAT and LSAT-Flex test takers. In September 2020, nearly 70,000 individuals spent time on Khan Academy's Official LSAT® Prep platform. Thousands of official practice questions from real tests are available, and learners can create a personalized practice plan tailored to their strengths and weaknesses. Interactive lessons, problem sets, timed practice exams, and videos engage learners in making real progress toward their goals. The Khan platform (Figure 1) is designed to build confidence in the learning process for prospective test takers while teaching the fundamental reasoning skills that are central to both legal education and academic work.

FIGURE 1. Khan Academy's Official LSAT Prep platform



One purpose of this preliminary study was to examine usage of the Khan Academy test-prep platform and to test the hypothesis that such usage has a positive effect on LSAT performance. However, we caution that the research design is not a true experiment in which individuals are randomly assigned to treatment and control conditions (Shadish et al., 2002)—though a number of potentially confounding variables were controlled through regression analysis. While we believe Khan usage is the best explanation for our findings, the possibility of alternative explanations cannot be ruled out (Grosz et al., 2020). A second purpose of this study is to determine whether the association between level of Khan usage and LSAT performance differed across racial, ethnic, and gender subgroups.

Literature Review

Research in the field of test preparation shows that while students generally engage in some form of test-prep study, they often have unequal access to costly test-prep services. Powers (1998) surveyed a random sample of students registered to take the SAT to determine whether the availability and use of a variety of SAT-prep programs provided access and learning opportunities to students without adequate family resources or income. Results of the survey indicated that some methods of test prep were more frequently cited than others, including taking the PSAT (Preliminary SAT) and utilizing test-prep materials provided by the College Board, whereas other, more resource-intensive methods were used less often and were likely beyond the means of average test takers. Examples of such methods include private tutoring, services provided by commercial test-prep companies, and programs sponsored by a school or other organization outside of school. Students in the study self-reported an average of approximately 11 hours preparing for the SAT. Students in the sample who attended coaching or tutoring sessions outside of school paid an average of $400 for these services. According to a recent LSAC report, Sweeney et al. (2019) found that the most popular self-indicated methods of preparation for the LSAT in 2017–2018 (prior to the availability of Khan Academy's Official LSAT Prep) were self-study, official LSAC prep materials, and commercially available non-LSAC books or software.

Researchers have cautioned that self-selection can distort the perceived effectiveness of a program: Students who engage in test preparation may differ systematically from students who do not in terms of characteristics such as motivation or degree of financial support. Study designs must control for these differences before attributing score gains to test preparation (Powers, 1993). Powers and Camara (1999) found that coached SAT takers differed from noncoached SAT takers in terms of socioeconomic status, race, and ethnicity. Similarly, Buchmann et al. (2010) found that family financial capital influenced the likelihood that test takers will engage in test prep.

While massive open online courses (MOOCs) and independent learning content sites are becoming increasingly popular, there has been little research studying their effectiveness in terms of student learning outcomes. Despite a high dropout rate for some MOOCs (Hew & Cheung, 2014), e-learning reaches a wide variety of learners and is quickly being incorporated into the educational landscape (Abdulaziz, 2018). Gardner and Brooks (2018) explored how learners' online course data could be used to create predictive models of student success, but they also noted methodological gaps in much

of this research, such as the lack of experimental design. Other research investigates psychological considerations in how MOOCs are designed and used to aid student instruction (Terras & Ramsay, 2015) as well as how to best maintain motivation and engagement (Sun et al., 2019). Given the prevalence of online courses and e-learning materials, research is needed to understand their impact on student learners.

Sample

Within 2 days after every test administration, LSAC emails students who have taken the most recent administration of the LSAT, inviting them to complete the Post-LSAT Questionnaire (PLQ). One PLQ question asks LSAT takers to indicate whether they used Khan Academy to prepare for the test. Those who indicate in the affirmative are given the option to have their data used for research purposes. For any given test administration, about 65–75% of Khan users consent to sharing their data with LSAC researchers; in this report, they are identified as the Consenting Khan Users subgroup.

Data for Consenting Khan Users from June 2018 (Khan Academy's Official LSAT Prep launch date) through July 2020 were matched with Khan Academy usage information, resulting in a sample of 12,471 records. Of these test takers, 61% took the LSAT once, 27% took the LSAT twice, 8% took the LSAT three times, and 3% took it four or more times.

To determine the representativeness of the sample, we compared the Consenting Khan Users subgroup (n = 12,471) to (a) the All Self-Reported Khan Users subgroup (i.e., all 2018–2019 test takers who indicated that they had used Khan to prepare for the LSAT; n = 38,572 test takers) and (b) the All Test Takers group (i.e., all 2018–2019 test takers; n = 100,183).

Table 1 displays distributions of test takers across the various demographic subgroups; some differences across samples are evident (see also Table A-1 in the Appendix). Those in the Consenting Khan Users subgroup were more likely to identify as Black, white, and female compared to those in the All Self-Reported Khan Users subgroup and the All Test Takers group. They were also more likely to be Pell Grant recipients, and more likely than those in the All Test Takers group to be first-generation college graduates.[1] Those in the Consenting Khan Users subgroup were also less likely to be first-generation college graduates compared to those in the All Self-Reported

[1] "First generation" was defined as individuals whose parents did not receive a bachelor's degree.

Khan Users subgroup. Because of these differences, we used proportional poststratification weighting[2] so that the Consenting Khan Users subgroup better reflected the All Self-Reported Khan Users subgroup in terms of gender, race, and ethnicity.

TABLE 1. Demographic breakdown of Khan sample, Khan users, and all test takers

Group	Consenting Khan Users (June 2018-July 2020) [a]	All Khan Users (June 2018-May 2019) [b]	All Test Takers (June 2018-May 2019)[c]
% American Indian	0.29	0.39	0.36
% Asian	7	10	11
% Black	15	13	11
% Hispanic	8	8	9
% White	54	53	51
% Two or more races/ethnicities	10	9	9
% Female	64	59	57
% Male	36	41	43
% Pell Grant recipient	37	27	25
% First generation	35	45	30
Average age	26	25	25
Average LSAT score	151	151	151
Sample size	12,471	38,572	100,183

[a] Khan users who consented to sharing their data with LSAC researchers for LSAT exams taken between June 2018 and July 2020.

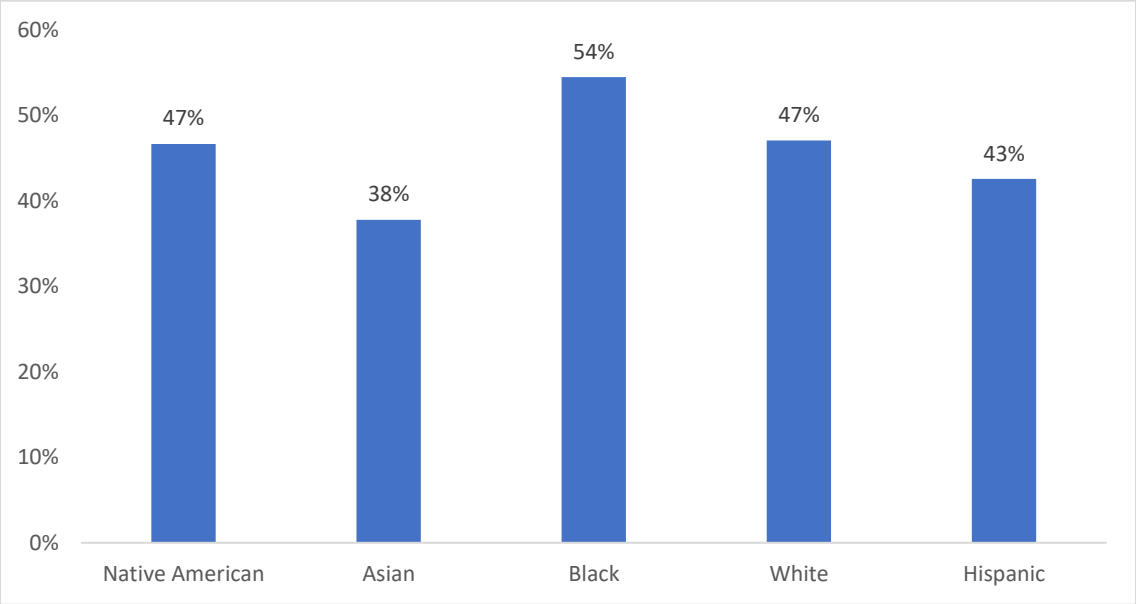
[b] LSAT test takers who indicated they used Khan Academy when asked at time of taking LSAT for LSAT exams taken during the 2018-2019 testing year.

[c] All individuals who took the LSAT during the 2018-2019 testing year.

Across subgroups based on race and ethnicity, nearly half of test takers indicated that they used Khan Academy to prepare for the LSAT. As shown in Figure 1, Black test takers were the most likely demographic subgroup to have used Khan Academy as part of their LSAT preparation.

[2] See adjusting for nonresponse by weighting: http://www.restore.ac.uk/PEAS/nonresponsetxt.php#cation.

FIGURE 1. Percentage of test takers who self-reported using Khan Academy by race and ethnicity (n = 5,177) [3,4]



Descriptive Statistics

Analyses of Khan usage indicated that approximately 8% of the sample set a practice schedule; of those, the average practice goal was 8.6 hours per week. Usage was primarily measured through the following variables calculated over the time intervals prior to each time an individual took the LSAT: practice minutes, video minutes, and number of unique practice exams taken.[5] For example, for Khan users who took the LSAT for the first time, usage was calculated over the interval from the launch of Khan Academy's Official LSAT Prep platform (June 2018) to the date of their LSAT exam. For students who took the LSAT a second time, Khan usage was restricted to the time interval between their first and second exam dates. The average number of weeks spent engaging with the platform (defined as having positive minutes active in

[3] Native American includes individuals of Native Alaskan descent. See LSAC report *Understanding and Interpreting Law School Enrollment Data: A Focus on Race and Ethnicity* for more information on race and ethnicity categories: http://www.lsac.org/data-research/research/understanding-and-interpreting-law-school-enrollment-data-focus-race-and.

[4] Results are reported for the February 2020 administration, as this was the last time Khan usage data was collected on test day.

[5] Practice minutes are defined as time spent on Khan Academy's Official LSAT Prep practice tasks and includes both time spent on discrete skill practice tasks and time spent on practice exams/diagnostics. Video minutes are defined as time spent watching Khan Academy's LSAT instructional videos. At the time of this analysis, ten unique full-length LSAT practice exams were available through the Khan Academy platform.

the week) was 8 weeks. The average total time that Khan users spent on the platform prior to taking an LSAT was 24.5 hours: 18.3 hours were spent on practice tasks and 1.8 hours on instructional videos (Table 2).

TABLE 2: Engagement statistics by total sample and across UGPA subgroups

Engagement Statistic	Total Sample	UGPA <3.32	UGPA 3.32–3.75	UGPA >3.75
Goal practice hours/week				
Average	8.6	8.8	8.6	7.7
Mdn	6.0	6.0	5.0	5.0
SD	8.3	8.1	8.4	7.6
Number of weeks using Khan				
Average	8.1	7.4	7.9	8.6
Mdn	6.0	6.0	6.0	7.0
SD	6.7	6.5	6.8	6.9
Total hours spent on Khan				
Average	24.5	21.6	23.8	26.5
Mdn	14.7	10.4	13.6	18.4
SD	28.2	28.7	27.7	27.4
Video hours				
Average	1.8	2.0	1.7	1.6
Mdn	0.7	0.7	0.7	0.7
SD	2.8	3.3	2.8	2.4
Practice hours				
Average	18.3	16.0	17.6	20.1
Mdn	10.8	7.8	9.8	13.6
SD	21.0	21.3	20.4	20.9
Number of unique practice exams taken				
Average	4.5	4.1	4.4	4.7
Mdn	4.0	3.0	4.0	4.0
SD	2.9	2.8	2.9	3.0
Practice LSAT score gain				
Average	2.6	2.0	2.8	2.8
Mdn	3.0	3.0	3.0	3.0
SD	7.8	7.9	7.3	7.7
Sample size	12,471	3,399	3,439	4,373

About 51% of Khan users did not complete an LSAT practice exam. Of those who did, a little under half (47%) completed 1–3 practice exams, 44% completed 4–9 practice exams, and 9% completed all 10 available practice exams. The average

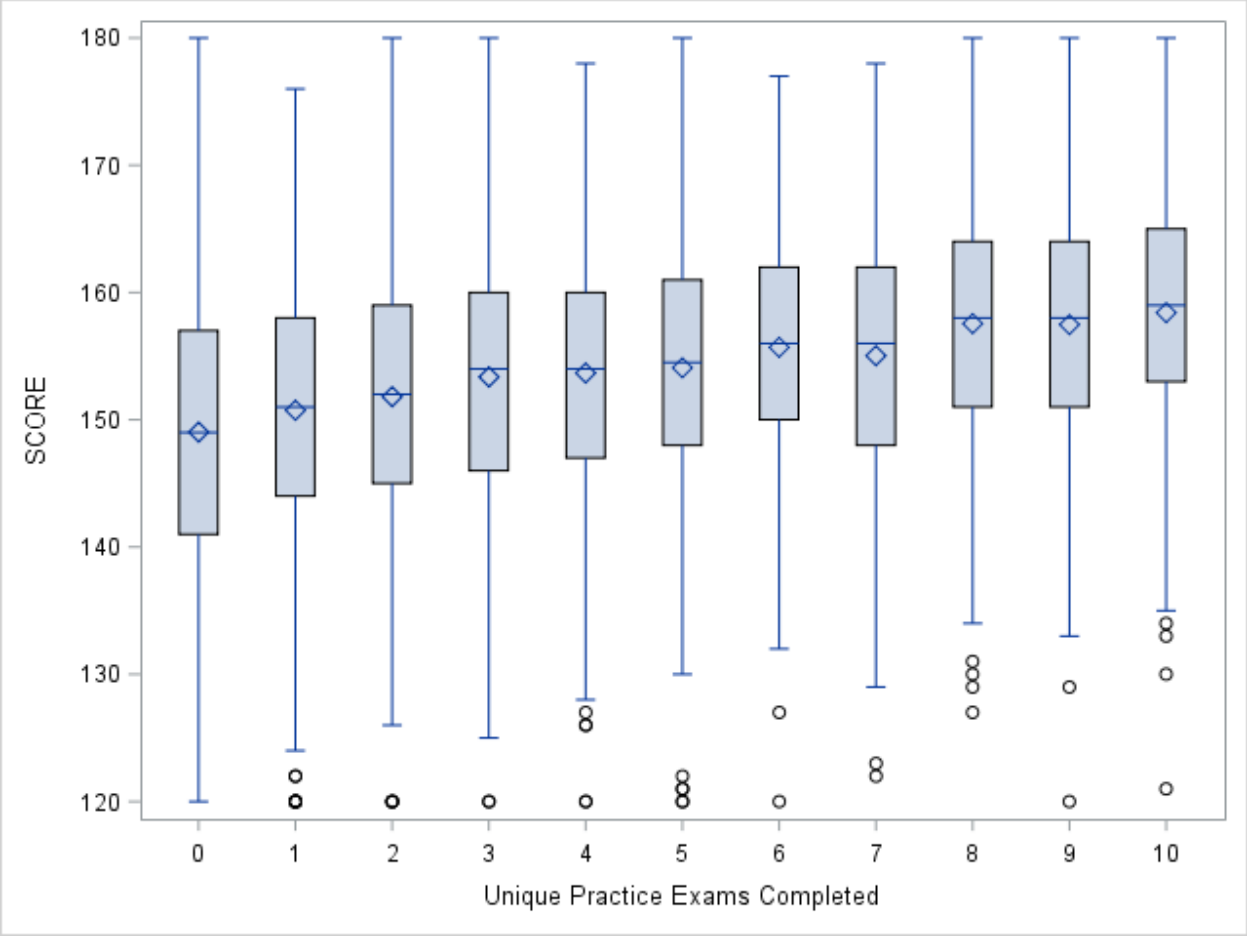
number of unique practice exams completed was 4.5 exams, and the average gain between the first practice LSAT score and the most recent practice LSAT score was about 3 points. While some variation in usage statistics was observed across demographic subgroups, differences were generally small and not statistically significant. Descriptive statistics are also reported across undergraduate GPA (UGPA) subgroups. Those with a higher UGPA tended to use Khan more than those with a lower UGPA. Prior research has shown that a higher GPA may be reflective of stronger study habits, increased motivation, and study enjoyment (Alshawwa et al., 2015). Accordingly, persistence may be a mediating factor influencing academic success.

Two primary usage variables were used to measure engagement in our analyses: (a) minutes spent on practice activities[6] and (b) number of unique practice exams taken, binned into six categories (0, 1–2, 3–4, 5–6, 7–8, and 9–10; see Table 7). Other usage variables were examined but were found not to correlate with actual LSAT scores. Video minutes, for example, were not correlated with LSAT scores, but they were positively associated with age ($r = 0.16$, $p < .0001$), meaning that older test takers were somewhat more likely to spend time viewing videos.

Because the relationship between the number of practice exams and LSAT scores was not assumed to be linear, our subsequent regression analysis examined the potential boost for practice exams binned into five categories (1–2, 3–4, 5–6, 7–8, and 9–10) relative to having taken no practice exams. Taking more practice exams generally resulted in increased LSAT scores relative to number of practice exams taken (unbinned; see box plots in Figure 2).

[6] Because the raw distribution of minutes spent on practice activities was right-skewed with high-end outliers, we used a log transformation to normalize the distribution.

FIGURE 2. LSAT score by number of unique practice exams completed



As shown in Table 3, Pearson correlations of UGPA with LSAT score were positive (*r* = .33). However, practice minutes showed a stronger relationship with LSAT score than with UGPA (.19 versus .08, respectively). This suggests that students' practice minutes are not entirely determined by their prior achievement, as measured by UGPA; rather, students across a variety of achievement levels are engaging with Khan practice activities.

TABLE 3: Pearson's correlations among LSAT score, UGPA, and practice minutes

Variable	LSAT Score	UGPA
UGPA	0.33*	—
Practice minutes	0.19*	0.08*

**p < .0001*

Methods

A series of least-squares linear regressions were used to model the effects of Khan usage while controlling for UGPA, Pell Grant status (a proxy for socioeconomic status), and test-taker age. Regressions were performed separately to test the effect of each of two engagement variables on LSAT score: log practice minutes (LPM) and number of unique practice exams completed. These two engagement variables were modeled separately. Both were considered to have value for describing the potential effects of Khan Academy LSAT preparation; however, it was determined that including both variables in the model simultaneously would create multicollinearity problems. Regressions were conducted on the total sample as well as within race, ethnicity, and gender subgroups—demographic classifications particularly important to LSAC for understanding diversity issues.

Results

Table 4 presents regression results for the effect of LPMs on LSAT score. A separate regression was also performed that restricted the sample to individuals who had LPMs of at least 2 (corresponding to approximately 6 minutes of practice time). Both regressions controlled for UGPA, Pell Grant status (binary: 1 or 0), and age. This regression model is depicted in the following equation using $Y = LSAT$ score as the dependent variable:

Regression 1

$$Y = a + \sum_{i=1}^{4} b_i X_i + E$$

where

a = Model intercept

b_1 = Slope for log practice minutes (X_1)

b_2 = Slope for UGPA on 4.0 scale (X_2)

b_3 = Slope for Pell Grant status (X_3)

b_4 = Slope for test-taker age (X_4)

E = Residual

TABLE 4: Regression statistics by total sample and by subgroup with LPM ≥ 2

Regression Statistic	Total Sample LPM	LPM ≥ 2
Standardized slope *β*	.17 [a]	.18 [a]
R^2	.21	.21
Sample size	6,938	6,550

[a] Unstandardized regression coefficient significant at $p < 0.01$.

As shown in Table 4, the standardized regression slope was designated generically as *β*, and it has a scale similar to that of correlation coefficients. The standardized slopes can be interpreted as moderate, but positive and statistically significant, suggesting a positive effect on LSAT scores with more time spent on Khan Academy's Official LSAT Prep practice activities.[7] Little difference was observed between the regression results for the full sample and the regression results for the sample that was restricted to participants spending at least 7 practice minutes on the Khan platform. Similar results were also observed when analyses were repeated across racial, ethnic, and gender subgroups.

Table 5 provides a practical guide for interpreting the results of the regression for the total sample, showing the increase in LSAT score corresponding to practice time. To facilitate interpretation, all comparisons below are made relative to the 10th percentile of practice time (26 minutes). Accordingly, students at the 90th percentile of practice time

[7] A standardized regression coefficient of 0.17 indicates that—holding UGPA, Pell Grant status, and test-taker age constant—every increase of 1 standard deviation in LPM on average results in a 0.17 increase in the standard deviation of LSAT score.

(47 hours) had scores that, on average, were 4.3 points higher than students at the 10th percentile. Analyses of racial, ethnic, and gender subgroups showed similar increases, with all subgroups having standardized regression coefficients within 0.05 points of one another. These subgroup breakdowns can be found in Tables A-2 and A-3 in the Appendix. Note that the results reported in Table 5 are not gains for students who took the LSAT twice but rather increments for independent groups of students who spent different amounts of practice time on Khan.

TABLE 5. Average LSAT scores for practice time percentiles [a]

Percentile	Time	Increase	Effect Size [b]
10th	26 m	Baseline	—
25th (Q1)	3 h	1.7	0.16
50th (Q2)	11 h	2.9	0.27
75th (Q3)	27 h	3.8	0.35
90th	47 h	4.3	0.40

[a] LSAT score gains from increased practice time compared to baseline 10th percentile. Practice time in minutes (m) and hours (h). Times over one hour rounded to nearest whole hour.

[b] To compute effect sizes, each score increase was divided by the sample standard deviation for the LSAT, 10.7.

We also considered whether the effect of LPM on LSAT scores might differ for students starting at different baselines, which we defined as the score on their first practice exam (FPE). To examine this potential association, a set of interaction variables were developed in two steps. First, a set of dummy variables was obtained for FPE score quartiles (*FPE1–FPE4*) and then mean centered (C*FPE1*–C*FPE4*); *LPM* was also mean centered (*CLPM*). Second, a set of interaction indicators was obtained: *INT2 = CFPE2*CLPM*, *INT3 = CFPE3*CLPM*, and *INT4 = CFPE4*CLPM*, where omitting *INT1* (and C*FPE1*) results in the first quartile being designated as the reference category. These interaction effects were then estimated controlling for UGPA, Pell Grant status, and age for test takers with *LPM* > 0 and who took at least one practice exam, as depicted in the regression model below using *Y = LSAT* score as the dependent variable:

Regression 2

$$Y = a + \sum_{i=1}^{4} b_i X_i + \sum_{i=2}^{4} c_i FPE_i + \sum_{i=2}^{4} d_i INT_i + E$$

where the regression coefficients (*a*, *b*) are defined above in Regression 1 and

$c_2 – c_4$ = *Slopes for FPE2 – FPE4*

$d_2 – d_4$ = *Slopes for INT2 – INT4*

As shown in Table 6, we observed larger standardized effects of LPM for individuals in the second and third FPE quartiles (.24) than for those in the fourth FPE quartile (.18). While individuals with lower FPE scores tend to score lower on the actual LSAT than those with higher FPE scores, these interactions also suggest that lower scoring individuals have a higher rate of return for practice minutes, whereas higher scoring students have less room to improve due to a potential ceiling effect.

TABLE 6. Regression statistics for model including interactions between LPM and FPE category

Predictor	Standardized Slope
LPM	.02
FPE2	.06
FPE3	.30
FPE4	.68
INT2	.24
INT3	.24
INT4	.18
R^2	.64
Sample size	4,607

Note. All regression slopes ≥ .18 were significant at $p < .01$.

We also found that the number of practice exams completed was associated with increased LSAT score increments, both overall and across various demographic subgroups. For the number of practice exams taken, we created a set of six dummy categorical variable bins *BPE1–BPE6* corresponding to six bins (0, 1–2, 3–4, 5–6, 7–8, and 9–10). Sample sizes for these bins varied across subgroups, as shown in Table 7.

TABLE 7. Sample size breakdown for race, ethnicity, and gender subgroups by bins indicating number of practice exams taken

Binned Practice Test Category Statistic	Black	Hispanic	White	Asian	Male	Female	Total Sample
0 Count	406	289	1,118	159	639	1,287	1,955
0 Column percent	36.5	35	26.8	33	27.8	30.9	29.8
1-2 Count	282	187	995	93	545	963	1,530
1-2 Column percent	25.4	22.6	23.8	19.3	23.7	23.1	23.4
3-4 Count	172	129	713	80	338	726	1,079
3-4 Column percent	15.5	15.6	17.1	16.6	14.7	17.4	16.5
5-6 Count	98	95	531	46	286	484	781
5-6 Column percent	8.8	11.5	12.7	9.5	12.5	11.6	11.9
7-8 Count	88	61	360	54	197	355	556
7-8 Column percent	7.9	7.4	8.6	11.2	8.6	8.5	8.5
9-10 Count	66	65	458	50	291	350	649
9-10 Column percent	5.9	7.9	11.0	10.4	12.7	8.4	9.9
Sample size	1,112	826	4,175	482	2,296	4,165	6,550

The following regression model was then applied to each demographic subgroup separately using the LSAT score (labeled *Y*) as the dependent variable:

Regression 3

$$Y = a + \sum_{i=1}^{3} b_i X_i + \sum_{i=2}^{6} c_i BPE_i + E$$

where

a = Model intercept

b_1 = Slope for UGPA (X_1)

b_2 = Slope for Pell Grant status (X_2)

b_3 = Slope for test taker age (X_3)

c_2 – c_6 = Slopes for BPE2 – BPE6

E = Residual

Both the raw and the standardized coefficients are given in Table 8, where each column shows results from a separate regression for each demographic subgroup. The coefficients c_2–c_6 can be interpreted as the average point increase in LSAT score for bins 2–6, respectively, relative to having taken no practice exams (referred to here as the zero-exam reference group), controlling for UGPA, Pell Grant status, and age (this

interpretation follows from omitting *BPE1* from the regression model). For practical purposes, the magnitudes of standardized regression slopes can be interpreted along a scale from −1 to +1. For example, Black Khan users who took 1–2 practice exams scored 2.84 points higher, on average, compared to Black Khan users who took no practice exams, while Black Khan users who took 9–10 practice exams scored, on average, about 8 points higher than Black Khan users who took no practice exams. Note that the coefficients are all expressed in terms of a comparison to the same demographic subgroup of Khan users who took no practice exams but who spent at least 6 minutes of practice time on Khan. Regressions include UGPA, Pell Grant Status, and test-taker age. The final two rows show the R-square and adjusted R-square values for separate regression models that only include the covariates UGPA, Pell Grant status, and age for anyone in the sample who took a complete practice exam. Results for Native American test takers are not reported due to the small sample size but will be reported in future research after more Khan usage data are available.

TABLE 8. Regression statistics for racial, ethnic, and gender subgroups, using binned number of practice exams to predict LSAT scores[a]

Binned Practice Exam Regression Statistic	Black	Hispanic	White	Asian	Male	Female	Total Sample
1-2 Slope	2.84*	2.46*	.95**	−1.39	.75	1.76*	1.59*
1-2 *β*[b]	.13	.11	.04	−.05	.03	.07	.06
3-4 Slope	4.63*	3.81*	2.98*	1.90	2.55*	4.20*	3.58*
3-4 *β*	.18	.14	.12	.07	.09	.15	.13
5-6 Slope	5.35*	6.59*	3.73*	1.90	3.23*	4.88*	4.39*
5-6 *β*	.16	.22	.13	.05	.10	.15	0.14
7-8 Slope	5.33*	4.32*	5.34*	5.01*	4.80*	5.60*	5.59*
7-8 *β*	.15	.12	.16	.16	.13	.15	.15
9-10 Slope	8.00*	7.05*	6.63*	4.73*	5.70*	7.61*	7.26*
9-10 *β*	.21	.20	.22	.14	.18	.20	.21
R^2	.21	.21	.17	.23	.18	.29	.23
Adj R^2	.21	.21	.16	.21	.18	.29	.23
R^2 [c]	0.14	0.15	0.11	0.18	0.15	0.23	0.18
Adj R^2 [c]	0.14	0.15	0.11	0.18	0.14	0.23	0.18

[a] Each exam group comparison is made to the zero-exam reference group. Regressions are restricted to individuals who spent ≥ 2 log practice minutes (about 6 minutes).

[b] Standardized regression coefficient.

[c] Separate regression includes only the covariates UGPA, Pell Grant status, and test-taker age.

*$p < 0.01$, ** $p < 0.05$.

In an effort to rule out alternative explanations for the influence that the number of practice exams may have on LSAT scores, we examined whether individuals who took more practice exams were also more likely to take a commercial LSAT prep course. If this were the case, the effects of Khan and commercial test prep would be confounded. Looking at the entire Khan sample, 31% reported having taken a commercial test-prep course. Table 9 displays the percentages for the subset who completed practice exams; those who completed more practice exams were not more likely to have also taken a commercial test-prep course. This indicates that it is unlikely that the boost associated with using Khan can alternatively be attributed to commercial test prep.

TABLE 9. Commercial test-prep usage for Khan users who completed at least one practice exam on the Khan platform

Binned Practice Exams	Percentage	*n*
1–2	30	598
3–4	26	378
5–6	29	297
7–8	30	222
9–10	31	273

Several analyses were also conducted to examine regression assumptions. First, despite the presence of a few moderately large residuals, regression coefficients were robust with respect to removal of these outliers. Second, comparative residual analysis was conducted, and it was determined that LSAT performance was predicted about equally well across most demographic subgroups.

Discussion

Our analyses showed that greater usage of LSAT test prep on Khan Academy was associated with higher performance on the LSAT. This finding holds true for demographic subgroups as well as for the overall population. Increased time spent on Khan Academy's Official LSAT Prep platform was associated with higher LSAT scores, and increases in both practice time and number of completed practice exams were associated with significantly higher LSAT performance. Our research also showed that the performance of test takers with lower initial practice exam scores was associated with slightly higher LSAT score gains per practice minute, indicating that students who might have otherwise scored lower on the LSAT may have benefitted the most from practice activities, a finding that could indicate that Khan has the potential to decrease

the score gap among demographic groups. Additionally, the association between a greater number of completed practice exams and higher LSAT scores was most pronounced among test takers who identified as Black, Hispanic, and female.

Because the Khan LSAT platform is free of charge and easily accessible online, individuals preparing to take the LSAT may want to consider incorporating Khan Academy's Official LSAT Prep activities when preparing for the LSAT. In fact, our sample of Khan users were less likely to indicate that they had taken a commercial test-prep course than were test takers generally (31% of our sample versus 43% of test takers who responded to the Post-LSAT Questionnaire), suggesting that many students are using Khan Academy as their core preparation, not simply as supplemental practice. Khan Academy LSAT preparation is likely a pragmatic alternative to commercial test-prep courses and private tutoring.

Future research is needed to continue evaluating the effectiveness of the Khan Academy LSAT prep platform. While students who took at least one practice exam took an average of 4.5 practice exams, 51% of students in the sample did not complete a full practice exam. Additional research may suggest how to engage students with the full set of features on Khan Academy, leading to broader impacts on LSAT performance. It would be useful for other analyses to focus on the effects of practice time spent on specific LSAT question types and more sensitive measures of engagement. As Khan Academy increases its reach and more individuals utilize the platform for test prep, additional research will benefit from larger sample sizes, allowing us to study the effects of Khan usage on smaller racial and ethnic subgroups (e.g., Native American) or gender subgroups within racial and ethnic subgroups. Other important demographic subgroups to study include disabled individuals as well as first-generation college students and individuals identifying as LGBTQIA+. Briggs (2001) also underscored the importance of considering a control group of students who did not engage in test prep as a comparison. While this current analysis focused on the intensity of Khan Academy prep for the LSAT and its associated effects on LSAT score, a future study examining different criteria for defining control groups of non-Khan users may be helpful.

One limitation accompanying any quasi-experimental study such as this one is that while engagement with Khan Academy LSAT practice activities and exams is associated with higher LSAT scores, we cannot establish that Khan Academy practice activities actually *caused* the increase in LSAT scores. We proportionally weighted the

sample (using post-stratification weights[8]) to reflect the population of all test takers who indicated using Khan in terms of race, ethnicity, and gender. We also controlled for important potentially confounding factors such as prior achievement (UGPA), Pell Grant status, and test-taker age, eliminating some alternative causes of the engagement effects. However, individuals who self-selected into Khan participation may differ in terms of other unobserved characteristics such as motivation, academic support, or life circumstances.

[8] Adjusting for non-response by weighting at https://www.restore.ac.uk/PEAS/nonresponsetxt.php#cation.

References

Abdulaziz, A. A. (2018). The changes in massive open online courses (MOOCs) studies between 2012 and 2017—a review of literature. *World Journal of Education, 8*(4), 59–71.

Al Shawwa, L., Abulaban, A. A., Abulaban, A. A., Merdad, A., Baghlaf, S., Algethami, A., Abu-Shanab, J., & Balkhoyor, A. (2015). Factors potentially influencing academic performance among medical students. *Advances in Medical Education and Practice, 6*, 65–75. http://dx.doi.org/10.2147/AMEP.S69304

Briggs, D. C. (2001). The effect of admissions test preparation: Evidence from NELS:88. *Chance, 14*(1), 10–18. https://doi.org/10.1080/09332480.2001.10542245

Buchmann, C., Condron, D. J., & Roscigno, V. J. (2010). Shadow education, American style: Test preparation, the SAT and college enrollment. *Social Forces, 89*(2), 435–461.

Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs. *User Modeling and User-Adapted Interaction, 28,* 127–203. https://doi.org/10.1007/s11257-018-9203-z

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspectives on Psychological Science, 15*(5), 1243–1255. https://doi.org/10.1177/1745691620921521

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review, 12*, 45–58. http://dx.doi.org/10.1016/j.edurev.2014.05.001

Powers, D. E. (1993). Coaching for the SAT: A summary of the summaries and an update. *Educational Measurement: Issues and Practice, 12*(2), 24–30. http://doi.org/10.1111/j.1745-3992.1993.tb00530.x

Powers, D. E. (1998). Preparing for the SAT®I: Reasoning Test—An update (College Board Report No. 98-5). http://files.eric.ed.gov/fulltext/ED562813.pdf

Powers, D. E., & Camara, W. J. (1999). Coaching and the SAT®I (College Board, Office of Research and Development Report No. RN-06). http://files.eric.ed.gov/fulltext/ED562660.pdf

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton, Mifflin and Company.

Sun, Y., Ni, L., Zhao, Y., Shen, X.-L., & Wang, N. (2019). Understanding students' engagement in MOOCs: An integration of self-determination theory and theory of relationship quality. *British Journal of Educational Technology, 50*(6), 3156–3174. https://doi.org/10.1111/bjet.12724

Sweeney, A. T., Fox, C., & Reese, L. M. (2019). Summary of self-reported methods of test preparation by LSAT takers for testing years 2014–2015 through 2017–2018 (TR 18-03). Law School Admission Council.

Terras, M. M., & Ramsay, J. (2015). Massive open online courses (MOOCs): Insights and challenges from a psychological perspective. *British Journal of Educational Technology, 46*(3), 472–487. https://doi.org/10.1111/bjet.12274

Appendix

TABLE A-1. Demographic breakdown of Post-LSAT Questionnaire respondents

	Consenting Khan Users		Non-Khan User [a]
Group	**Yes [b]**	**No**	**Total**
American Indian	0.29	0.39	0.33
Asian	7	10	13
Black	15	13	11
Hispanic	8	8	9
White	54	53	50
Two or more races/ethnicities	10	9	9
Female	64	59	56
Male	36	41	44
Pell Grant recipient	37	26	24
First generation	35	33	27
Average age	26	25	25
Average LSAT score	151	151	151
Sample size	12,471	37,450	71,368

[a] 2018–2019 testing year.

[b] Consenting Khan users who were located in the Khan database were included in sample.

TABLE A-2. Average LSAT scores and effect sizes for practice time percentiles by racial and ethnic subgroup

	Black			Hispanic			White			Asian		
Percentile	**Time**	**Increase**	**Effect Size**	**Time**	**Increase**	**Effect Size**	**Time**	**Increase**	**Effect Size**	**Time**	**Increase**	**Effect Size**
10th	14 m	Baseline	—	23 m	Baseline	—	34 m	Baseline	—	20 m	Baseline	—
25th (Q1)	2 h	1.5	0.1	2 h	1.8	0.2	3 h	1.6	0.1	2 h	1.4	0.1
50th (Q2)	9 h	2.7	0.3	9 h	3	0.3	12 h	2.8	0.3	11 h	2.6	0.2
75th (Q3)	23 h	3.4	0.3	24 h	4	0.4	28 h	3.6	0.3	28 h	3.3	0.3
90th	46 h	3.9	0.4	44 h	4.6	0.4	46 h	4.1	0.4	49 h	3.7	0.3

TABLE A-3. Average LSAT scores and effect sizes for practice time percentiles by gender subgroup

	Male			Female		
Percentile	**Time**	**Increase**	**Effect Size**	**Time**	**Increase**	**Effect Size**
10th	31 m	Baseline	—	24 m	Baseline	—
25th (Q1)	3 h	1.3	0.1	3 h	1.8	0.2
50th (Q2)	12 h	2.3	0.2	10 h	3.2	0.3
75th (Q3)	29 h	2.9	0.3	26 h	4.1	0.4
90th	48 h	3.3	0.3	46 h	4.7	0.4