ESSA RESEARCH REPORT

March 15, 2019

The Impact of IXL Math and IXL ELA on Student Achievement in Grades Pre-K to 12

IXL LEARNING 777 Mariners Island Blvd., Suite 600, San Mateo, CA 94404
650-372-4040 | www.ixl.com

TABLE OF CONTENTS

ESSA Evidence for IXL Math and ELA

This evaluation of IXL's impact on mathematics and English language arts achievement meets the required rigor of the What Works Clearinghouse (WWC) standards for quasi-experimental studies with reservation and the Every Student Succeeds Act (ESSA) Tier II standard for evidence-based interventions. In accordance with these standards, this study used a pretest-posttest quasi-experimental design and implemented propensity score matching to reduce or eliminate selection bias. Our treatment and control groups were well matched for analysis following ESSA and WWC guidelines.

As required by ESSA Tier II standards, this study also: included a large sample size (4,000 students across multiple sites); measured outcomes using a reliable benchmark assessment (NWEA MAP); applied multilevel models to account for sample clustering effects; and controlled for potentially confounding factors in the analysis including prior performance and background (e.g., gender, student status, race/ethnicity, grade level, English language learner status, special education status, and teacher background or experience) (ESSA n.d.).

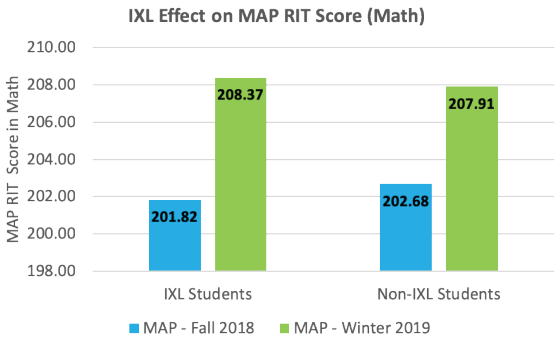
This study found that the use of IXL had a positive and statistically significant effect on student academic achievement in math and reading. In addition, students with higher levels of IXL usage were more likely to have greater growth in both subjects.

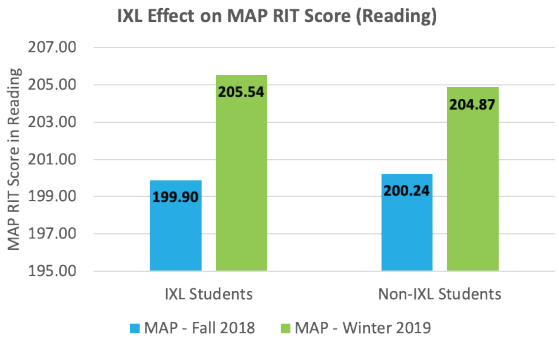
Executive Summary

This study took place in a large virtual public charter school in the United States. The school provides internet-based individualized instruction to students in grades Pre-K to 12. IXL, a personalized online learning platform, has been provided to students in the charter school since 2017.

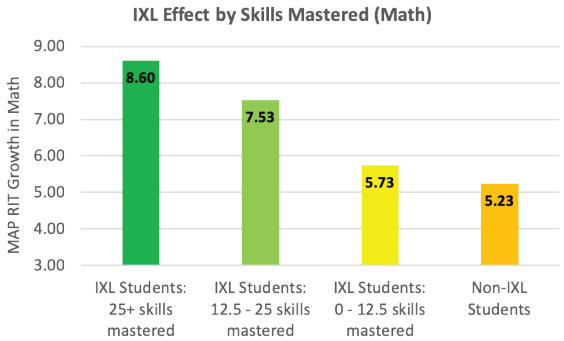
This study focused on approximately 4,000 students in grades Pre-K to 12 who began using IXL for the first time during the fall semester of the 2018-19 school year. Matched students from the same school without access to IXL were treated as a control group. The duration of the IXL implementation was one semester (about 17 school weeks). IXL usage by the students in this study ranged from less than one minute per week to over five hours per week. Even with the short implementation time and the wide range in usage, we found a positive correlation between IXL usage and student academic achievement, as measured by the NWEA MAP tests. The key findings of this study include:

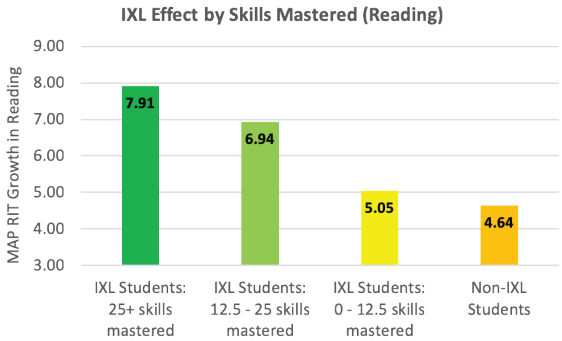
- **IXL has a positive effect on student learning.** Students using IXL outperformed students without IXL by approximately 1 point on the MAP math and reading tests across grades Pre-K to 12.





- **More IXL usage leads to higher achievement.** The IXL effect was larger for students with more questions answered and more skills mastered on IXL. Students who mastered 25 or more IXL skills outperformed non-IXL students by more than 2 points on MAP math and reading tests.





The Impact of IXL Math and IXL ELA on Student Achievement in Grades Pre-K to 12

Introduction

The school participated in this study is a large virtual public charter school that provides individualized instruction to Pre-K to 12th grade students seeking a non-traditional educational setting. The school implements a blended learning model that allows students to set their own pace with guidance and instruction from their assigned teacher. IXL has been offered in this charter school as one option for supplemental curricula since 2017.

To help inform the ongoing development of the school's blended learning model, IXL Learning researchers conducted an evaluation of the impact of IXL on student learning. The evaluation was based on a quasi-experimental design, which was designed to meet the What Works Clearinghouse (WWC) standards with reservation and the Tier 2 evidence standards of the Every Student Succeeds Act (ESSA). These two designations are the highest standards for quasi-experimental design.

The purpose of this study is to evaluate the effectiveness of IXL on student mathematics and reading achievement as measured by the NWEA MAP tests. The research questions are:

- **(IMPACT)** Does student performance on the NWEA MAP test differ for students who use IXL and similar students who do not use IXL?
    - What is the impact of IXL Math on student mathematics achievement?
    - What is the impact of IXL ELA on student reading achievement?
    - To what extent do impacts vary by student subgroups (including grade level, prior achievement, gender, English language learners, economically disadvantaged students, racial/ethnic minority students, and those receiving special education services)?
- **(USAGE)** For students exposed to IXL, what is the relationship between their IXL usage and their academic achievement?

Methods

STUDY DESIGN

A quasi-experimental pretest-posttest control group design (see Figure 1) was used to measure the effectiveness of IXL. This type of study evaluates the treatment effect by comparing the performance of the treatment group and the control group on the posttest, after adjusting for their background and pretest performance.

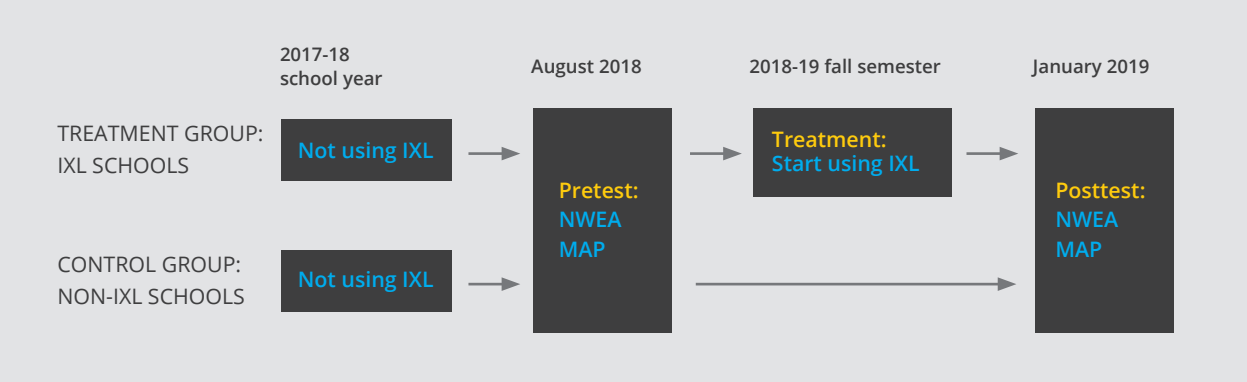


Figure 1. Study design

PARTICIPANTS

This study took place in the fall semester of the 2018-19 school year at the charter school. Each enrolled student is assigned to a teacher. Students work with their assigned teacher to create an Individual Learning Plan for the school year and choose their core and supplemental curricula based on their needs and interests. The charter school used the NWEA MAP as their benchmark assessment to track students' progress. The treatment group of this study includes tested students[1] who chose IXL for the first time during the fall semester of the 2018-19 school year. The treatment group consisted of 3,678 students for math and 2,929 students for reading.

The control group includes matched students who did not use IXL in the 2017-18 or 2018-19 school years. One-to-one matching was used to match each student in the treatment group with a peer student who did not use IXL and had an identical or very similar background. The matching criteria include grade level, gender, ethnicity, English language learner status, special education status, economically disadvantaged status, pretest score, and the background of the assigned teacher (i.e., the number of students assigned to the teacher and whether the teacher was a new teacher[2]). Details of the matching method are presented in Appendix A. All students in the treatment group were matched; therefore, the control group consisted of 3,678 students for math and 2,929 students for reading.

Table 1 shows changes in the samples between the point of matching and the analysis. There was a loss of students due to the lack of posttest (see Table 1). For math, the attrition rate was 14% for IXL students and 21% for non-IXL students. For reading, the attrition rate was 14% for IXL students and 20% for non-IXL students. The attrition rate differences between the IXL students and non-IXL students were 7% for math and 6% for reading. Attrition rate differences within 15% are considered to be acceptable according to the ESSA Standards (ESSA, n.d.). The on-track students (i.e., students with both pretest and posttest results) in the last row of Table 1 were the sample used to evaluate the effect of IXL in this study.

[1] *Tested students: students who took the NWEA MAP in fall 2018.*

[2] *A new teacher is a teacher who started at the charter school in the fall semester of the 2018-19 school year.*

Table 1. Number (percentage) of students in IXL group and control group

Values	Math		Reading	
	IXL	Non-IXL	IXL	Non-IXL
Matched students	3,678 (100%)	3,678 (100%)	2,929 (100%)	2,929 (100%)
Loss due to lack of posttest	525 (14%)	764 (21%)	406 (14%)	583 (20%)
On-track students	3,153 (86%)	2,914 (79%)	2,523 (86%)	2,346 (80%)

Table 2 presents the equivalence at pretest for on-track students. Column "Diff" is the average difference in standard deviation units between IXL students and non-IXL students. The difference for prior achievement did not exceed 0.05 standard deviations and none of the background differences exceeded 0.25 standard deviations, which indicates that IXL students and non-IXL students are two equivalent groups in both math and reading according to the WWC standards (WWC, 2017) and ESSA standards (ESSA, n.d.).

Table 2. IXL and Non-IXL Equivalence at Pretest for On-track Students

	Math					Reading				
	IXL (N = 3,153)		Non-IXL (N = 2,914)		Diff[a]	IXL (N = 2,523)		Non-IXL (N = 2,346)		Diff[a]
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
MAP RIT Fall 2018[b]	0.04	0.92	0.09	1.01	-0.05	0.03	0.94	0.06	1.01	-0.03
Gender										
Male	51%	0.50	50%	0.50	0.02	50%	0.50	50%	0.50	0.01
Female	49%	0.50	50%	0.50	-0.02	50%	0.50	50%	0.50	-0.01
Status										
Econ. disadv.	61%	0.49	61%	0.49	0.01	60%	0.49	62%	0.49	-0.03
Special education	19%	0.39	19%	0.39	-0.02	20%	0.40	21%	0.41	-0.03
ELL	1%	0.09	1%	0.09	0.01	1%	0.08	0%	0.07	0.02

Race/Ethnicity										
White	65%	0.48	67%	0.47	-0.04	65%	0.48	67%	0.47	-0.05
American Indian	14%	0.34	13%	0.34	0.01	14%	0.35	14%	0.34	0.02
African American	10%	0.30	9%	0.29	0.05	10%	0.30	9%	0.28	0.05
Hispanic or Latino	10%	0.30	10%	0.29	0.01	9%	0.29	9%	0.29	0.00
Asian	1%	0.10	1%	0.10	-0.01	1%	0.10	1%	0.08	0.02
Native Hawaiian	1%	0.08	1%	0.07	0.02	0%	0.07	1%	0.07	-0.01
Grade level										
Pre-K and K	3%	0.16	3%	0.17	-0.01	3%	0.16	2%	0.16	0.01
Grade 1	4%	0.19	4%	0.19	0.00	4%	0.20	5%	0.21	-0.03
Grade 2	9%	0.29	9%	0.28	0.01	9%	0.29	9%	0.28	0.02
Grade 3	13%	0.34	13%	0.33	0.01	14%	0.35	14%	0.34	0.01
Grade 4	12%	0.32	12%	0.32	0.01	12%	0.33	12%	0.32	0.01
Grade 5	13%	0.34	13%	0.34	-0.01	13%	0.34	15%	0.35	-0.04
Grade 6	15%	0.35	15%	0.36	-0.02	15%	0.35	16%	0.36	-0.03
Grade 7	16%	0.37	17%	0.37	-0.01	16%	0.36	15%	0.35	0.03
Grade 8	15%	0.36	15%	0.35	0.01	14%	0.34	13%	0.34	0.02
Grade 9 and up	1%	0.09	1%	0.07	0.03	1%	0.08	0%	0.07	0.04
Teacher background										
# of students[c]	0.36	0.68	0.35	0.68	0.00	0.38	0.69	0.38	0.68	0.00
New teacher	39%	0.49	39%	0.49	0.01	39%	0.49	38%	0.48	0.02

[a] Diff is the difference between IXL students and non-IXL students in standard deviation units. It is computed as the mean difference divided by the standard deviation for non-IXL students.

[b] MAP RIT score was standardized within each grade level.

[c] The number of students assigned to each teacher was standardized across all teachers.

IXL

IXL is a personalized learning platform designed to help students build academic skills that are fully aligned to state standards. It offers thousands of skills in math, English language arts (ELA), science, and social studies from Pre-K to 12th grade. As students practice on IXL, they receive questions that automatically adapt to their skill level, and get progressively more challenging as they work. As of 2019, IXL is being used by over 350,000 teachers worldwide. Teachers have used IXL to introduce new topics, help students to reinforce concepts, prepare for standardized tests, and provide personalized instruction to students. Teachers can also track progress for individual students or entire classes on IXL and adjust their classroom instruction to meet student learning needs.

Throughout IXL, student progress is measured by the program's proprietary SmartScore. The SmartScore starts at 0, increases as students answer questions correctly, and decreases if questions are answered incorrectly. A student is considered proficient in a skill when they reach a SmartScore of 80. A student is considered mastery in a skill when they reach a SmartScore of 100. SmartScore measures are used throughout this analysis to assist in the interpretation of the IXL usage effect.

NWEA MAP

In this study, students' academic achievement in math and reading were assessed using the math and reading sections of the NWEA MAP, respectively. MAP is a collection of computer-based adaptive assessments administered to students in grades Pre-K to 12. Students below 2nd grade take the MAP Growth K-2, students in grades 2 to 5 take the MAP Growth 2-5, and students at or above 6th grade take the MAP Growth 6+. MAP is administered to students three times throughout the school year: August, January, and May. The August 2018 MAP tests were used as the pretest and the January 2019 MAP tests were used as the posttest in this study.

Each MAP test reports a RIT score, which is a Rasch Unit scale score that measures student performance, regardless of age or grade level. The RIT scale scores typically range between 150 and 300. The higher the RIT score, the higher achievement the student has shown in the subject. In MAP math and reading tests, RIT scores are also reported in different goal areas to show students' relative strength and concern areas. Table 3 shows the goal area names for the MAP math and reading tests.

Table 3. MAP Tests Goal Area Names

Test name	Goal area	Math	Reading
MAP Growth K-2	Goal 1	Number Sense	Reading Foundations
	Goal 2	Algebraic Reasoning and Algebra	Comprehension, Critical Reading, and Research
	Goal 3	Geometry and Measurement	Vocabulary
	Goal 4	Data and Probability	Writing and Language
MAP Growth 2-5, MAP Growth 6+	Goal 1	Number and Operations	Reading Process: Read and Comprehend Texts
	Goal 2	Algebraic Reasoning and Algebra	Critical Reading: Interpret and Evaluate Texts
	Goal 3	Geometry and Measurement	Vocabulary
	Goal 4	Data and Probability	N/A

To measure student growth, MAP reports Met Projected Growth and Conditional Growth Index (CGI). Met Projected Growth indicates whether students met growth projections (Yes) or fell short (No). CGI shows how much individual growth deviates from the student growth norms. CGI is expressed in standard deviation units and can be used to compare students across grades and achievement levels. A CGI of zero means a student showed gains that were equivalent to the growth norms. A CGI of 1.0 indicates that a student's growth was one standard deviation above the norm, which would represent a high level of growth. By contrast, a CGI of -1.0 indicates that a student's growth was 1 standard deviation below the norm. This study used Met Projected Growth and CGI from fall to winter to measure students' growth during the fall semester of the 2018-2019 school year.

ANALYSIS

Because matching may not result in identical treatment and control groups, a "doubly robust" approach (Funk, et al., 2011) was applied to evaluate the effectiveness of IXL. The "doubly robust" approach combines the benefits of matching and regression adjustment. The regression-based adjustment was used to account for residual differences between IXL students and matched non-IXL students.

Impact Analysis (Research Question 1). We applied a series of two-level hierarchical regression models to calculate the IXL effect—i.e., the performance difference between IXL students and non-IXL students on the NWEA MAP, controlling for factors such as prior performance, gender, student status, race/ethnicity, grade level, and teacher background. Separate regression models were used to estimate the IXL effect in math and reading. To examine the extent to which the effect of IXL differs across student subgroups, we included an additional term in the regression model, separately for each subgroup category, that captures the interaction between IXL access and a particular student subgroup.

Usage Analysis (Research Question 2). We conducted two types of analyses to examine the relationship between IXL usage and student achievement. The first analysis built off of the impact analysis model to look at the relationship between different levels of IXL usage and student achievement, relative to non-IXL students. For this analysis, we set benchmarks for low, medium, and high IXL usage and substituted these student usage indicators into the regression model. The second analysis examined the relationship between different levels of IXL usage and student achievement among IXL students. For this analysis, we ran a different set of two-level hierarchical regression models that estimate the within-teacher relationship between the student-level IXL usage and achievement, taking into account students' prior performance and background. The two types of analyses demonstrates whether higher usage of IXL is associated with better achievement. (See Appendix B for a detailed explanation of analytical methods.)

Results

IXL USAGE SUMMARY

Students started to use IXL on different dates across the fall semester of the 2018-19 school year. Table 4 presents an overview of student start time on IXL. About half of the students started using IXL in September. Nearly 20% of the students did not start on IXL until November.

Table 4. Start Date on IXL

Start date	IXL Math		IXL ELA	
	N	%	N	%
August	11	0%	7	0%
September	1,657	53%	1,184	47%
October	967	31%	818	32%
November	338	11%	330	13%
December	180	6%	184	7%
Total	3,153	100%	2,523	100%

The amount of usage on IXL varied across students. Table 5 shows the IXL usage for students at the 25th, 50th, and 75th percentiles from 08/01/2018 to 12/31/2018. An average student (at the 50th percentile) spent 221 minutes on IXL Math and 164 minutes on IXL ELA, which is approximately 13 minutes per week[3] on IXL Math and 10 minutes per week on IXL ELA. The average number of questions answered is approximately 27 per week on IXL Math and 20 per week on IXL ELA. The majority of the students (75%) achieved mastery on less than one skill per week on IXL Math and IXL ELA.

Table 5. IXL Usage for Students at the 25th, 50th, and 75th Percentiles

Usage measure	IXL Math			IXL ELA		
	25th	50th	75th	25th	50th	75th
Time spent (in minutes)	78	221	505	54	164	424
Questions answered	160	460	962	106	345	856
Skills practiced	6	15	32	3	10	24
Skills proficient	4	12	27	2	8	20
Skills mastered	2	7	18	1	4	12

Note: the duration of the IXL usage in this table is from 08/01/2018 to 12/31/2018.

IXL EFFECT

Because students may choose to use IXL Math, IXL ELA, or both subjects, we first compared the performance difference between students with both IXL subjects and students with only one IXL subject. The analysis showed no difference between the usage of two subjects and one subject in math (β = -0.03, p = 0.96) and reading (β = 0.38, p = 0.70). Therefore, we combined students with IXL Math only and students with both IXL Math and IXL ELA in the math analysis. We also combined students with IXL ELA only and students with both IXL Math and IXL ELA in the reading analysis.

IXL Effect on MAP RIT Score. The use of IXL showed a statistically significant effect on students' performance on the NWEA MAP in both math and reading. Figure 2 shows the MAP RIT scores in fall and winter for IXL students and non-IXL students in math and reading. The IXL effect is 1.15 in math and 0.90 in reading (see Appendix C, Table C1 for details). That is, if an average non-IXL student had used IXL in the fall semester of the 2018-19 school year, the student would be expected to score 1.15 points higher in math and 0.90 points higher in reading on the NWEA MAP.

[3] *This study assumed there were 17 school weeks during the fall semester of the 2018-19 school year.*

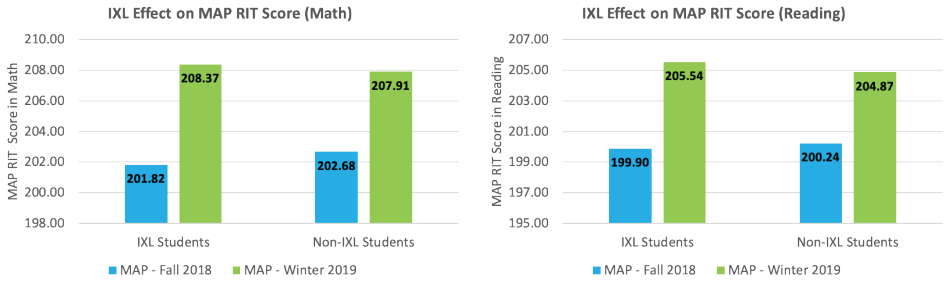


Figure 2. The IXL Effect on the MAP RIT Scores

IXL Effect on MAP Fall to Winter Growth. The use of IXL also showed a statistically significant effect on students' MAP Conditional Growth Index (CGI) in both math and reading. Figure 3 shows that the CGI in math is 0.12 for IXL students and -0.13 for non-IXL students, and the CGI in reading is 0.12 for IXL students and -0.04 for non-IXL students. IXL students made more improvement than the national norm in both math and reading, while non-IXL students made less improvement than the national norm. The IXL effect is 0.21 for math and 0.13 for reading (see Appendix C, Table C2 for details).

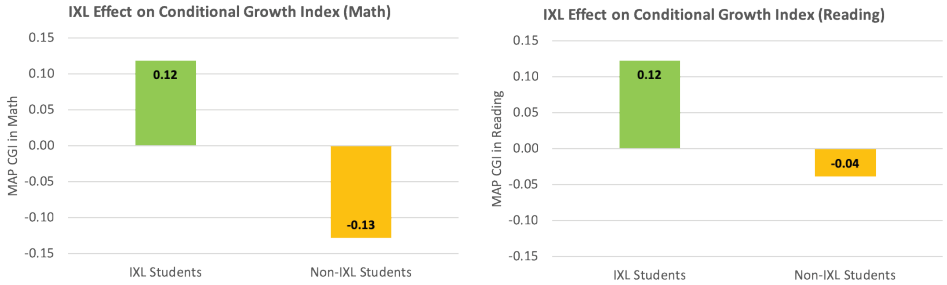


Figure 3. The IXL Effect on the MAP Conditional Growth Index

The IXL effect was also observed in the percentage of students who Met Projected Growth on MAP. As shown in Figure 4, a higher percentage of IXL students met the projected growth from fall to winter than non-IXL students. The percentage difference is 5 percent for math and 2 percent for reading.

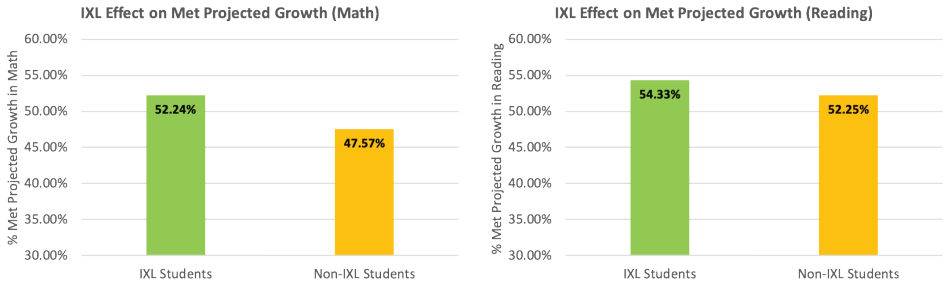


Figure 4. The IXL Effect on Percentage of Met Projected Growth on MAP

IXL Effect in MAP Goal Areas. In all goal areas of the MAP math and reading tests, the IXL effect was also found to be positive and statistically significant. Figure 5 shows the MAP RIT growth, which is computed as the RIT score in winter 2019 minus the RIT score in fall 2018. IXL students performed better than non-IXL students in all goal areas in both math and reading (see Appendix C, Tables C3 and C4 for details).

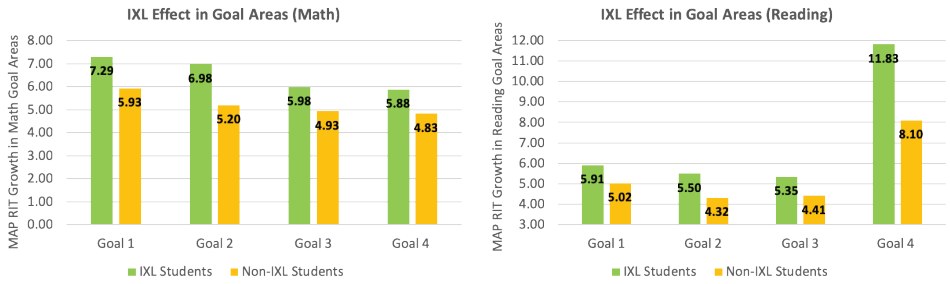


Figure 5. The IXL Effect on MAP RIT in Different Goal Areas

Note: For reading, Goal 4 only applies to students who took the MAP Growth K-2 test (see Table 3).

IXL Effect in Different Grade Levels. There is a statistically significant interaction effect between IXL usage and student grade level (see Figure 6). For math, the IXL effect is 2.83 for grades pre-K to 2, 1.12 for grades 3 to 5, and 0.63 for grades 6 to 12. For reading, the IXL effect is 3.46 for grades pre-K to 2, 0.67 for grades 3 to 5, and 0.44 for grades 6 to 12. The IXL effect is higher at the lower elementary level than the upper elementary and middle/high school levels in both math and reading (see Appendix C, Tables C5 and C6 for details).

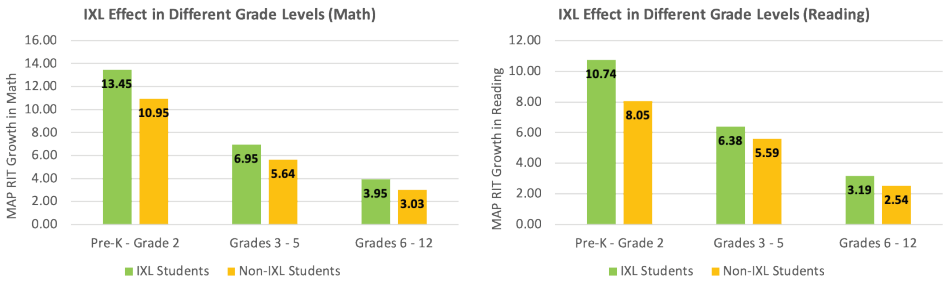


Figure 6. The IXL Effect in Different Grade Levels

IXL Effect for Other Subgroups. No interaction effect was found between the IXL effect and other student subgroups (i.e., prior achievement, gender, English language learners, economically disadvantaged students, racial/ethnic minority students, and those receiving special education services). This indicates that the IXL effect is similar across all these subgroups.

THE USAGE EFFECT OF IXL

IXL Effect by Questions Answered. A positive and statistically significant association was found between the number of questions answered on IXL and student MAP performance. Figure 7 shows the MAP RIT growth for non-IXL students and IXL students with different numbers of questions answered on IXL within the fall semester. For math and reading, the IXL effect is statistically significant when students answered at least 250 questions (about 15 questions per week). More questions answered is associated with a greater IXL effect. For students who answered 850 or more questions (about 50 questions per week), the IXL effect is 1.74 in math and 1.61 in reading (see Appendix C, Tables C7 and C8 for details).

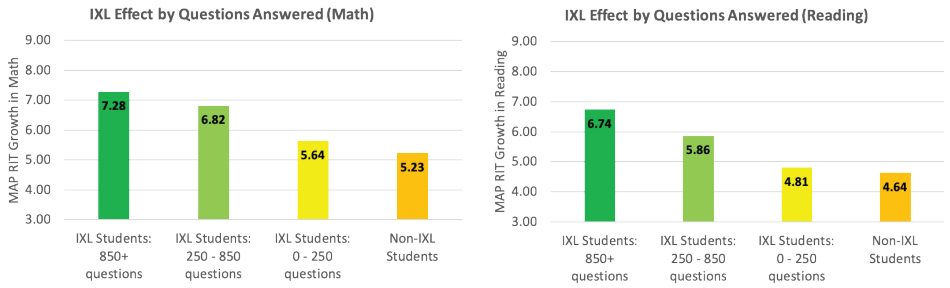


Figure 7. The IXL Effect by Questions Answered per Student

IXL Effect by Skills Mastered. A positive and statistically significant association was also found between the number of skills mastered on IXL and student MAP performance. Figure 8 shows the MAP RIT growth for non-IXL students and IXL students with different numbers of IXL skills mastered within the fall semester. For math, the IXL effect is statistically significant for IXL students in all three usage groups. For reading, the IXL effect is statistically significant when students mastered at least 12.5 skills (about 0.7 skills per week). More skills mastered is associated with higher IXL effect. For students who mastered 25 or more skills (about 1.5 skills per week), the IXL effect is 2.20 in math and 2.17 in reading (see Appendix C, Tables C9 and C10 for details).

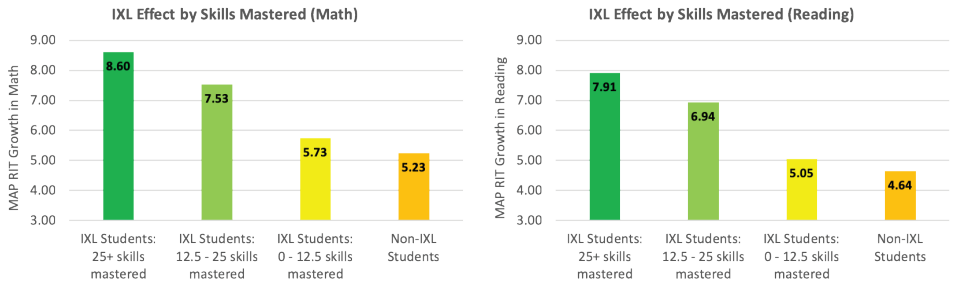


Figure 8. The IXL Effect by Skills Mastered per Student

The Effect of Additional IXL Usage. Our analysis also revealed that the amount of IXL usage is positively correlated with student performance on MAP. Figure 9 shows the expected MAP RIT score improvement if there were additional usage of IXL each week. If a student mastered one additional IXL Math skill per week during the fall semester, the student could expect to improve 0.42 points on the MAP RIT score in math in winter 2019. If a student mastered one additional IXL ELA skill per week during the fall semester, the student could expect to improve 0.49 points on the MAP RIT score in reading in winter 2019 (see Appendix C, Table C11 for details).

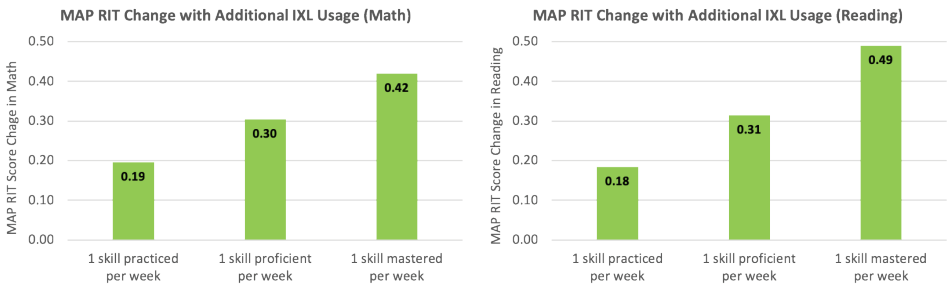


Figure 9. MAP RIT Score Improvement with Additional IXL Usage

Conclusion

This study observed the implementation of IXL during a short time frame (the fall semester of the 2018-19 school year), and students' average weekly usage during this time was approximately 13 minutes on IXL Math and 10 minutes on IXL ELA. Even with this short implementation, analysis of the data showed that the use of IXL had a small positive effect on student academic achievement in both math and reading. IXL students made more improvement from fall to winter as compared to the national norm than non-IXL students. These effects were statistically significant, indicating there is a high probability that similar students using IXL would achieve similar results. The IXL effect was also observed in all goal areas of the MAP math and reading tests.

The analysis also showed a positive correlation between IXL usage and student academic achievement. In particular, the IXL effect for students with more than 25 skills mastered (1.5 skills per week) is about three times higher than the IXL effect for students with less than 12.5 skills mastered (about 0.7 skills per week). One additional skill mastered per week was associated with an expected 0.5 point increase on MAP RIT scores in both math and reading.

References

Every Student Succeeds Act (n.d.). *Evidence for ESSA: Standards and Procedures*. Retrieved from https://content.evidenceforessa.org/sites/default/files/On%20clean%20Word%20doc.pdf

Funk, M. J., Westreich, D., Wiesen, C., Sturmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology*, 173(7), 761–767.

Ho, D., Imai, K., King, G., & Stuart, E. A.(2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28.

What Works Clearinghouse (2017). *What Works Clearinghouse standards handbook* (Version 4.0). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Appendix A: Propensity Score Matching

A propensity score is the probability that a student with certain characteristics will be assigned to the treatment group (as opposed to the control group). Propensity score matching is a statistical method to create equivalent treatment and control groups in order to reduce or eliminate selection bias between the two groups. The most commonly used propensity score matching method, one-to-one matching, was applied in this study. This method forms pairs of treatment student and control student, such that matched students have identical or very similar values of the propensity score. That is, each IXL student in the treatment group is matched with a non-IXL student with identical or very similar characteristics.

The matching criteria in this study include grade level, gender, ethnicity, English language learner status, special education status, economically disadvantaged status, prior achievement as measured by the MAP tests in August 2018, and the background of the assigned teacher (i.e., the number of students assigned to the teacher and if the teacher was a new teacher in the fall semester of the 2018-19 school year). The matching criteria were only based on the data collected during the pretest and before the treatment. R package Matchit (Ho, et al., 2011) was used to carry out the matching. Following Every Student Succeeds Act (ESSA) and What Works Clearinghouse (WWC) guidelines, we targeted a matched sample that results in group differences of no more than 0.05 standard deviations for the prior achievement measure and no more than 0.25 standard deviations for the student background measures (ESSA, n.d.; WWC, 2017).

Table A1 shows the number of students in the treatment group (IXL) and the control group (non-IXL) before and after matching. Every student in the IXL group was matched with one student in the non-IXL group. No student was discarded during the matching process.

Table A1. Number of matched and unmatched students

Values	Math		Reading	
	IXL	Non-IXL	IXL	Non-IXL
All tested students	3,678	7,092	2,929	7,941
Matched students	3,678	3,678	2,929	2,929
Unmatched students	0	3,414	0	5,012
Discarded students	0	0	0	0

Appendix B: Analytical Methods

Because matching may not result in identical treatment and control groups, this study applied a "doubly robust" approach (Funk, et al., 2011) that combines the benefits of both matching and regression-based adjustment to evaluate the effectiveness of IXL. Any residual differences between IXL students and matched non-IXL students would be accounted for by the regression adjustment. In this section, we outlined the analytic approach that we used to address each of the two research questions.

1. Impact Analysis (Research Question 1)

The IXL effect on student academic achievement was evaluated by two-level hierarchical regression models. The model accounted for the nesting effect of students under teachers and the differences between the IXL students' and matched non-IXL students' characteristics. The model takes the following general form:

Level 1 (students):

Equation 1a. $Y_{ij} = \beta_{0j} + \beta_{1j}T_{ij} + \beta_{2j}X_{ij} + e_{ij}$,
where Y_{ij} is the MAP RIT score for student i assigned to teacher *j*; T_{ij} is a dichotomous indicator for whether the student had access to IXL (T_{ij}= 1) or not (T_{ij}= 0); and X_{ij} is a vector of student background characteristics, including prior achievement (RIT score from fall 2018 MAP tests, standardized within each grade level) and dichotomous indicators of student background, including grade level, gender, ethnicity, English language learner status, special education status, economically disadvantaged status. The main parameter of interest is β_{1j}, which is the IXL effect (i.e., the effect of using IXL).

Level 2 (teachers):

Equation 1b. $\beta_{oj} = Y_{00} + Y_{01}C_j + u_{ij}$,
where C_j is a vector of teacher characteristics, including the number of students assigned to the teacher in the fall semester (centered on the average number of students for all teachers) and a dichotomous indicator of whether the teacher was a new teacher in the 2018-19 school year.

The model was run separately for math and reading to estimate the effect of IXL Math and IXL ELA, respectively. To examine the extent to which the IXL effect differs across student subgroups, we included an additional term in Equation 1a, separately for each subgroup category, to capture the interaction between IXL access and a particular student subgroup. For example, to test whether the IXL effect differs for males and females, we ran a model that includes an interaction term between IXL access and whether the student is female or not.

2. Usage Analysis (Research Question 2)

The impact analysis examined the IXL effect, but the magnitude of the IXL effect could depend on the extent to which students actually used IXL. To examine the relationship between IXL usage and student achievement, we conducted two types of analyses. The first analysis built off of the impact analysis model to look at the relationship between different levels of IXL usage and student

achievement, relative to students with no access to IXL (i.e., the control group). For this analysis, we constructed benchmarks for low, medium, and high IXL usage and substituted these student usage indicators into Equation 1a instead of the dichotomous IXL access indicator. We selected two usage measures that were the best representation of IXL usage to construct the benchmarks. The two usage measures are number of questions answered on IXL and number of skills mastered on IXL. This analysis tells us whether the effect is larger for students who answered more questions and mastered more skills on IXL.

The second analysis examined the relationship between different levels of IXL usage and student achievement among IXL students. For this analysis, we ran a series of two-level hierarchical regression models that estimate the within-teacher relationship between the student-level measure of IXL usage and achievement, taking into account student characteristics and prior achievement. The usage measures include number of skills practiced (skills with at least one question answered) per week, number of skills proficient (SmartScore >= 80) per week, and number of skills mastered (SmartScore = 100) per week. The model takes the following general form:

Level 1 (students):

Equation 2a. $Y_{ij} = \beta_{0j} + \beta_{1j}Z_{ij} + \beta_{2j}X_{ij} + e_{ij}$,
where Y_{ij} is the MAP RIT score for student *i* assigned to teacher *j*; Z_{ij} is a measure of IXL usage for student *i*, centered on the mean level of usage for teacher *j* ; and X_{ij} is a vector of student background characteristics the same as Equation 1a. The main parameter of interest is β_{1j}, which is the usage effect of IXL.

Level 2 (teachers):

Equation 2b. $\beta_{oj} = Y_{00} + Y_{01}\bar{Z}_j + Y_{02}C_j + u_{ij}$,
where $\bar{Z}_j$ is a measure of average IXL usage of all students assigned to teacher *j*, centered on the mean level of usage for all teachers; and C_j is a vector of teacher characteristics the same as Equation 1b.

To assist in the interpretation of the IXL effect and the usage effect of IXL, we reported statistical significance and effect size. Statistical significance, also referred to as *p*-value, is the probability that the IXL effect is zero. A small *p*-value (e.g., less than 0.05) indicates strong evidence that the IXL effect is not zero. Effect size is the mean difference in standard deviation units and is known as Hedges' *g*. In this study, effect size is computed using adjusted mean and unadjusted standard deviations. More details about these analytical methods can be found in What Works Clearinghouse (2017).

Appendix C: Data Tables

Table C1. IXL Effect on MAP RIT Score

Values	Math		Reading	
	IXL	Non-IXL	IXL	Non-IXL
Number of students	3,153	2,914	2,523	2,346
MAP RIT score 08/2018	201.82	202.68	199.90	200.24
MAP RIT score 01/2019	208.37	207.91	205.54	204.87
IXL effect	1.15***		0.90**	
Effect size	0.05		0.04	

*Note: ***: significant at .001 level; **: significant at .01 level.*

Table C2. IXL Effect on MAP Fall to Winter Growth

Values	Math		Reading	
	IXL	Non-IXL	IXL	Non-IXL
Number of students[a]	3,122	2,876	2,523	2,346
MAP fall to winter CGI	0.12	-0.13	0.12	-0.04
IXL effect	0.21***		0.13*	
Effect size	0.11		0.01	

Note: [a]The CGI for a few students was not reported. The sample only included students with CGI.
****: significant at .001 level; *: significant at .05 level.*
CGI: Conditional Growth Index.

Table C3. IXL Effect on MAP Fall to Winter Growth in Math

Values	Math Goal 1		Math Goal 2		Math Goal 3		Math Goal 4	
	IXL	Non-IXL	IXL	Non-IXL	IXL	Non-IXL	IXL	Non-IXL
Number of students	3,153	2,914	3,152[a]	2,914	3,152[b]	2,914	3,153	2,914
MAP RIT score 08/2018	201.86	202.80	201.14	202.16	201.71	202.56	202.71	203.41
MAP RIT score 01/2019	209.15	208.73	208.12	207.36	207.69	207.49	208.59	208.24
RIT fall to winter growth	7.29	5.93	6.98	5.20	5.98	4.93	5.88	4.83
IXL effect	1.11***		1.48***		0.83*		1.17***	
Effect size	0.05		0.07		0.04		0.05	

Note: Goal 1: Number Sense for MAP Growth K-2; Number and Operations for MAP Growth 2-5 and 6+.

Goal 2: Algebraic Reasoning and Algebra.

Goal 3: Geometry and Measurement.

Goal 4: Data and Probability.

[a] One student in Grade 3 has RIT Goal 2 score missing.

[b] One student in Grade 1 has RIT Goal 3 score missing.

****: significant at .001 level; *: significant at .05 level.*

Table C4. IXL Effect on MAP Fall to Winter Growth in Reading

Values	Reading Goal 1		Reading Goal 2		Reading Goal 3		Reading Goal 4	
	IXL	Non-IXL	IXL	Non-IXL	IXL	Non-IXL	IXL	Non-IXL
Number of students	2,523	2,345[a]	2,522[b]	2,345[a]	2,523	2,346	357[c]	327[c]
MAP RIT score 08/2018	198.69	198.83	200.31	200.77	201.40	201.86	166.61	167.05
MAP RIT score 01/2019	204.60	203.85	205.81	205.09	206.75	206.27	178.44	175.15
RIT fall to winter growth	5.91	5.02	5.50	4.32	5.35	4.41	11.83	8.10
IXL effect	1.03*		0.92*		0.77*		3.26*	
Effect size	0.05		0.04		0.03		0.16	

Note: Goal 1: Reading Fundations for MAP Growth K-2; Reading Process - Read and Comprehend Texts for MAP Growth 2-5 and 6+.

Goal 2: Comprehension, Critical Reading, and Research for MAP Growth K-2; Critical Reading - Interpret and Evaluate Texts for MAP Growth 2-5 and 6+.

Goal 3: Vocabulary.

Goal 4: Writing and Language for MAP Growth K-2.

[a]One student in Grade 6 has RIT Goal 1 and Goal 2 score missing.

[b]One student in Grade 1 has RIT Goal 2 score missing.

[c]Goal 4 only included students who took MAP Growth K-2.

**: significant at .05 level.*

Table C5. IXL Effect by Grade Levels in Math

Values	Math Pre-K to 2		Math grades 3 to 5		Math grades 6+	
	IXL	Non-IXL	IXL	Non-IXL	IXL	Non-IXL
Number of students	484	445	1,202	1,102	1,467	1,367
MAP RIT score 08/2018	172.87	171.72	196.65	197.31	215.61	217.09
MAP RIT score 01/2019	186.32	182.67	203.60	202.95	219.56	220.12
RIT fall to winter growth	13.45	10.95	6.95	5.64	3.95	3.03
Interaction effect	1.71*		Reference group		-0.49	
IXL effect	2.83 (1.12 + 1.71)		1.12*		0.63 (1.12 - 0.49)	

*Note: *: significant at .05 level.*

Table C6. IXL Effect by Grade Levels in Reading

Values	Reading Pre-K to 2		Reading grades 3 to 5		Reading grades 6+	
	IXL	Non-IXL	IXL	Non-IXL	IXL	Non-IXL
Number of students	401	373	994	942	1,128	1,031
MAP RIT score 08/2018	172.30	171.82	196.15	196.23	213.00	214.18
MAP RIT score 01/2019	183.04	179.87	202.53	101.82	216.71	216.72
RIT fall to winter growth	10.74	8.05	6.38	5.59	3.19	2.54
Interaction effect	2.78**		Reference group		-0.23	
IXL effect	3.45 (0.67 + 2.78)		0.67		0.44 (0.67 - 0.23)	

*Note: **: significant at .01 level.*

Table C7. IXL Effect by Questions Answered on IXL Math

Values	IXL			Non-IXL
	850+ questions	250 - 850 questions	0 - 250 questions	
Number of students	925	1,162	1,066	2,914
MAP RIT score 08/2018	201.10	201.94	202.31	202.68
MAP RIT score 01/2019	208.38	208.76	207.95	207.91
RIT fall to winter growth	7.28	6.82	5.64	5.23
IXL effect	1.74***	1.38***	0.47	N/A
Effect size	0.08	0.06	0.02	

*Note: ***: significant at .001 level.*

Table C8. IXL Effect by Questions Answered on IXL ELA

Values	IXL			Non-IXL
	850+ questions	250 - 850 questions	0 - 250 questions	
Number of students	641	821	1,061	2,346
MAP RIT score 08/2018	198.17	199.72	201.08	204.87
MAP RIT score 01/2019	204.91	205.58	205.89	200.23
RIT fall to winter growth	6.74	5.86	4.81	4.64
IXL effect	1.61**	1.14*	0.32	N/A
Effect size	0.07	0.05	0.01	

*Note: **: significant at .01 level; *: significant at .05 level.*

Table C9. IXL Effect by Skills Mastered on IXL Math

Values	IXL			Non-IXL
	25+ skills mastered	12.5 - 25 skills mastered	0 - 12.5 skills mastered	
Number of students	567	530	2,056	2,914
MAP RIT score 08/2018	195.33	200.52	203.95	202.68
MAP RIT score 01/2019	203.93	208.05	209.68	207.91
RIT fall to winter growth	8.60	7.53	5.73	5.23
IXL effect	2.20***	1.53**	0.77*	N/A
Effect size	0.10	0.07	0.03	

*Note: ***: significant at .001 level; **: significant at .01 level; *: significant at .05 level.*

Table C10. IXL Effect by Skills Mastered on IXL ELA

Values	IXL			Non-IXL
	25+ skills mastered	12.5 - 25 skills mastered	0 - 12.5 skills mastered	
Number of students	322	307	1,894	2,346
MAP RIT score 08/2018	193.05	196.81	201.56	200.24
MAP RIT score 01/2019	200.96	203.75	206.61	204.87
RIT fall to winter growth	7.91	6.94	5.05	4.64
IXL effect	2.17**	1.79**	0.59	N/A
Effect size	0.10	0.08	0.03	

*Note: **: significant at .01 level; *: significant at .05 level.*

Table C11. The Effect with Additional IXL Usage

Values	Math	Reading
Number of students	3,153	2,523
1 skill practiced per week	0.19*	0.18
1 skill proficient per week	0.30**	0.31
1 skill practiced per week	0.42**	0.49*

*Note: **: significant at .01 level; *: significant at .05 level.*