



RESEARCH REPORT

December 2021

The Impact of IXL on ELA Learning in Iowa

Yao Xiong, Ph.D.

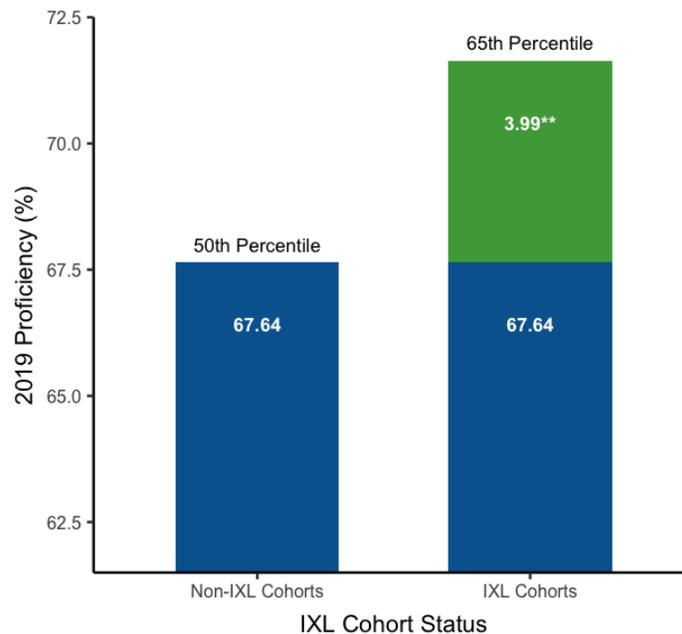
IXL LEARNING 777 Mariners Island Blvd., Suite 600, San Mateo, CA 94404
650-372-4040 | www.ixl.com

Executive Summary

IXL is a personalized learning platform designed to help students build academic skills in core subject areas. As of now, 1 in 5 students in the U.S. and more than 13 million students worldwide use IXL. Previous research at both the student level (e.g., Empirical Education, 2013) and the school level (e.g., An, 2021; Schonberg, 2021) supports the positive impact of IXL on student academic achievement.

The current study aims to further evaluate the impact of IXL on student academic achievement in English Language Arts (ELA). We used a quasi-experimental design with one-to-one propensity score matching to analyze data from 327 school districts across the state of Iowa. Specifically, students from 98 school districts had used IXL ELA during the 2018-2019 school year. IXL usage in these school districts ranged from less than one minute per student per week to more than sixty minutes per student per week. Upon screening out grade cohorts with no or very low usage, we found a positive and statistically significant effect of IXL on student achievement on the 2019 Iowa Statewide Assessment of Student Progress in ELA with a large effect size.

- IXL has a positive effect on students' achievement in ELA.** The proficiency rate among student cohorts in grades 3 through 8 that used IXL ELA was on average 3.99 points¹ higher than that of comparable non-IXL cohorts on the state assessment, which translates to a 15-point gain in percentile rank.



¹ ** in the figure indicates significance at the .01 level.

The Impact of IXL on ELA Learning in Iowa

Study Design

The purpose of this study was to evaluate the impact of IXL on student academic achievement in ELA across different grades and school districts in the state of Iowa. We used a quasi-experimental pretest-posttest design to evaluate the impact of IXL. Specifically, we used one-to-one propensity score matching (PSM) to identify a comparison group (i.e., non-IXL group) that was very similar to the treatment group (i.e., IXL group) in baseline performance and key demographic characteristics. Then, we evaluated IXL impact by comparing the treatment and comparison groups at posttest, controlling for their baseline performance and demographic characteristics.

We analyzed data from 327 school districts across Iowa. Students in 98 of these school districts used IXL ELA during the 2018-2019 school year. Grade cohorts within the school districts that used IXL ELA served as the treatment group, while a matched set of grade cohorts that did not use IXL ELA served as the comparison group. The study design is presented in Figure 1.

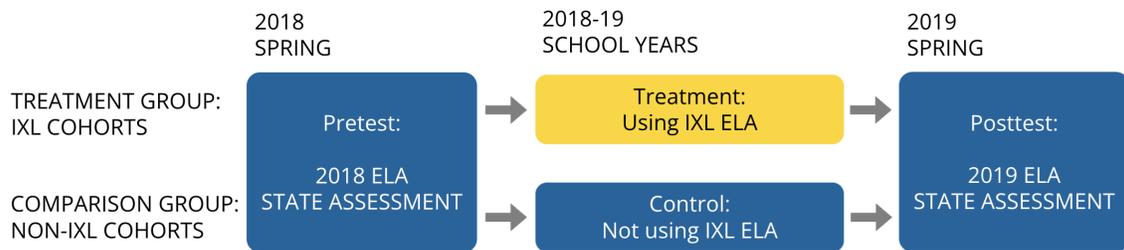


Figure 1. Study Design

Methodology

DATA SOURCES

For this study, we collected and merged data from three different sources:

- IXL usage data for the school districts in Iowa that used IXL ELA. These data were obtained from IXL’s internal database.
- Aggregate student achievement data in ELA and demographic information at the grade cohort level (or at the district level if the grade-cohort-level data were not available) from the Iowa Department of Education (IDOE).
- Demographic information at the district level from the National Center for Education Statistics (NCES).

IXL ELA USAGE

IXL usage data provided information on the adoption of IXL products (e.g., the number of students who used IXL ELA) for each grade level in a given school year. Among IXL cohorts, we examined three main usage indicators: a) the amount of time spent on IXL, b) the number of questions answered, and c) the number of skills in which students reached proficiency (i.e., “skills proficient”).

In this study, we focused on “IXL school districts” instead of “IXL schools” because most of the demographic information was available only at the district level. Further, as the number of students who practiced IXL ELA within a district ranged from a single student to an entire grade cohort, this study defined a school district as an “IXL school district” at each grade level rather than at the district level.

Specifically, a school district was identified as an IXL school district at a certain grade level if: a) at least 10% of the students in that grade cohort were using IXL ELA in the 2018-2019 school year; and b) students reached proficiency in at least two IXL skills per month (i.e., 0.5 skill proficiency per week). After applying these two criteria, a total of 96 IXL grade cohorts from 48 school districts were retained and made up the treatment group. Table 1 presents the usage descriptive statistics for the IXL cohorts.

Table 1. IXL ELA Usage

IXL usage (per student per week)	IXL ELA (<i>n</i> = 96)			
	<i>M</i>	<i>SD</i>	Min	Max
Time spent (in minutes)	16.28	10.67	4.07	68.14
Questions answered	47.78	26.24	18.10	171.53
Skills proficient	0.94	0.50	0.50	3.96

MEASURES

Achievement Measure

Student academic achievement was measured by the Iowa state assessment program. Specifically, the proficiency rate or the percentage of students meeting or exceeding the proficiency standard in each grade cohort served as the outcome measure. To be noted, the Iowa state assessment program became the Iowa Statewide Assessment of Student Progress (ISASP) in the 2018-2019 school year, with more stringent criteria for a student to be classified as proficient as compared to the state assessment program prior to 2018. This shift was reflected in the drop in proficiency rates from 2018 to 2019. Notwithstanding, the content evaluated in the state programs remained fairly consistent across years and was aligned with the Iowa Core standards. For a comprehensive overview of the Iowa state assessment program, please refer to the assessment blueprint (<https://iowa.pearsonaccess.com/bulletins/>).

Covariates

Per What Works Clearinghouse (WWC) protocols (WWC, 2020), covariates for both PSM and analysis included baseline performance and demographic characteristics as follows:

- 2017 ELA proficiency rate (baseline; % students meeting or exceeding the proficiency standard)
- Grade level (ranging from 3 to 8)
- Grade size (# of students enrolled)
- Special education (% students enrolled in the Individualized Education Program)
- English language learner (% students with limited English proficiency)
- Student-teacher ratio
- Socioeconomic status (% economically disadvantaged students measured by eligibility for free/reduced lunch program)
- Gender (% male students)

The descriptive statistics of all variables used in the current study are shown in Appendix A. The treatment and comparison groups had very similar characteristics. Note that we included only one in each set of two highly correlated demographic variables to eliminate redundancy and to prevent multicollinearity issues. Specifically, between school district location and grade size ($r = .72$), we retained grade size, as it provided more information as a continuous variable, given almost all districts were in non-city locations; between % English language learners (ELL; $r = -.82$) and race measured by % white students, we retained ELL, given the focus of the study was ELA achievement.

ANALYTIC APPROACH

To address the main research question about the impact of IXL ELA on student learning, we used a two-step approach. First, we identified a comparison group of cohorts that was very similar to the treatment group on the measured covariates. Second, we specified and estimated a multilevel regression model to examine the IXL effect.

In the first step, we used one-to-one PSM without replacement (Ho et al., 2011) to identify equivalent treatment and comparison groups. The propensity score is the probability of assigning a cohort with certain characteristics to the treatment group. Using the set of covariates (baseline and demographics), each treatment cohort was matched to a comparison cohort with the nearest propensity score value. Comparison cohorts were identified from the 1,486 non-IXL cohorts in the state. We used the R package MatchIt (Ho et al., 2011) for the PSM step. As a result, 96 treatment and 96 comparison cohorts were identified. The matched comparison group had much more similar characteristics with the treatment group than the original unmatched comparison group, as shown in Appendix C. For example, the 2018 proficiency rate (baseline) had an absolute standardized mean difference (ASMD) of less than 0.1 between the matched comparison group and the treatment group, while the original ASMD was above 0.25. Importantly, all covariates had ASMD values lower than 0.25 with most of the ASMD values being below 0.05 (WWC, 2020). The ASMD results indicated a high degree of similarity between the treatment and control groups.

In the multilevel regression model, grade cohorts were assumed to be nested within school districts, with each school district's mean proficiency rate being freely estimated (i.e., a random intercept model). At Level 1 (i.e., the grade cohort level), the outcome variable was the grade-level proficiency rate on the 2019 Iowa state assessment. The main covariate was baseline performance measured

as the proficiency rate on the 2018 Iowa state assessment. Other covariates included: a) grade level, grade size, % special education students, and % ELL at Level 1 (e.g., the grade cohort level); and b) student-teacher ratio, % economically disadvantaged students, and % male students at Level 2 (i.e., the district level). We were mostly interested in the regression coefficient of the IXL treatment variable (i.e., IXL effect) at Level 1. The model was estimated using the R package lme4 (Bates et al., 2015).

Following What Works Clearinghouse guidelines (WWC, 2020), each effect is accompanied by a statistical significance test with a probability value (p -value) and a measure of effect size. The p -value is the probability of observing the current or more extreme results, assuming the tested effect is zero (i.e., no effect) (Cohen, 1994). As such, a very small p -value means that it is very unlikely to observe the current result under the no-effect assumption. P -values less than .05 are considered statistically significant. Effect size is reported using Hedges' g , measuring the model-adjusted (i.e., estimated) mean difference between treatment and comparison groups on an outcome measure in unadjusted (i.e., observed) standard deviation units, adjusting for sample size (WWC, 2020). For research in the education field, effect sizes of 0.20 or greater would be considered large based on a review of a large number of empirical education interventions (Kraft, 2020). We also estimated the corresponding percentile gain based on the effect size for the IXL effect. Percentile gain, based on the effect size, is the expected change in IXL cohorts' percentile rank relative to non-IXL cohorts at the 50th percentile.

Results

EFFICACY OF IXL ELA

As shown in Figure 2, the ELA proficiency rate was 3.99 points higher among IXL ELA cohorts, as compared to non-IXL cohorts ($p < .01$; Hedge's $g = 0.38$), after accounting for baseline performance and demographic characteristics. This result indicates that IXL cohorts outperformed comparable non-IXL cohorts by having an additional 3.99% of students reach proficiency on the state assessment in ELA, holding all the covariates constant. If a typical non-IXL district started using IXL ELA, it would likely observe similar learning gains. The effect size of 0.38 corresponds to a percentile gain of 15 points—that is, if a non-IXL district at the 50th percentile adopted IXL ELA with students reaching proficiency in at least two skills per month, it would be expected to be at the 65th percentile on the state assessment in ELA after one year of IXL usage. Detailed results of the regression analysis are presented in Appendix B.

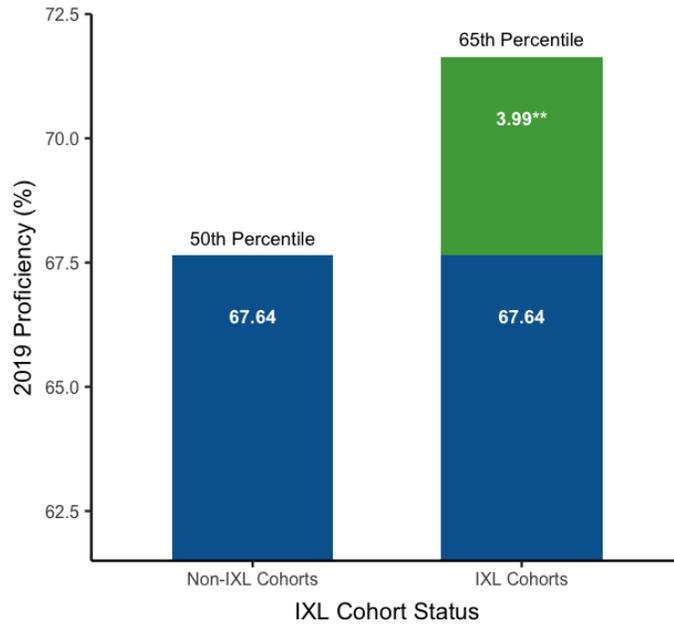


Figure 2. The Efficacy of IXL ELA

Conclusion

In the current study, we examined the effect of IXL ELA implementation on student ELA achievement at the grade cohort level across school districts in the state of Iowa. We adopted a quasi-experimental design to identify the most comparable non-IXL group and accounted for baseline performance and key demographic characteristics.

The results showed that IXL ELA had a positive and statistically significant impact on student academic achievement in ELA with a large effect size. IXL recommends that students aim to reach proficiency in two IXL skills per week to maximize learning gains. The IXL usage observed in the current study was noticeably lower than this recommendation. However, we still observed a sizable treatment effect with this limited use of IXL ELA. We would expect an even larger effect with greater IXL ELA usage. In sum, schools and districts that wish to boost student learning and achievement should consider implementing IXL into their classrooms.

References

An, X. (2021). *Measuring the impact of IXL Math and IXL Language Arts in Kentucky schools*. Research Report. <https://www.ixl.com/research/Impact-of-IXL-in-Kentucky.pdf>

Bates D, Mächler M, Bolker B, Walker S (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.

Empirical Education. (2013). *A study of student achievement, teacher perceptions, and IXL Math*. <https://www.ixl.com/research/IXL-Research-Study-2013.pdf>

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1–28. <https://www.jstatsoft.org/v42/i08/>

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241-253.

Schonberg, C. (2021). *The impact of IXL on math and ELA learning in Colorado: A quasi-experimental study*. Research Report. <https://www.ixl.com/research/Impact-of-IXL-in-Colorado.pdf>

What Works Clearinghouse. (2020). *What Works Clearinghouse procedures and standards handbooks*. <https://ies.ed.gov/ncee/wwc/Handbooks>

Appendix A. Descriptive Statistics for Achievement and Demographic Variables

Table A. Means (Standard Deviations) for IXL (Treatment) and Non-IXL (Comparison) Cohorts

	Treatment	Comparison
# of districts	48	79
# of cohorts (grades)	96	96
ELA Achievement	%	%
2018 Proficiency rate (baseline)	78.89 (7.93)	79.56 (8.86)
2019 Proficiency rate (posttest)	74.70 (8.89)	70.75 (11.97)
Grade-level characteristics		
Grade level:	%	%
grade 3	13.54	16.67
grade 4	11.46	10.42
grade 5	20.83	21.88
grade 6	16.67	13.54
grade 7	18.75	18.75
grade 8	18.75	18.75
Grade size (# of students)	76.38 (62.00)	82.13 (75.82)
% Special education	13.3 (4.72)	13.16 (6.08)
% English Language Learners (ELL)	3.08 (7.70)	3.63 (9.52)
District-level characteristics		
Student-teacher ratio	14.57 (2.97)	14.56 (2.68)
% Economically disadvantaged	37.15 (9.59)	37.10 (14.25)
Gender: % male	51.54 (2.06)	51.39 (2.21)

Appendix B. IXL Efficacy Results

Table B. The Effect of IXL ELA on 2019 ELA Proficiency

Predictor	<i>b</i>	<i>SE</i>	95% CI		β	<i>t</i>	<i>p</i>
(Intercept)	67.64	1.82	64.02	71.01	--	37.18	< .001
2018 ELA proficiency ¹	0.27	0.08	0.11	0.45	0.21	3.20	.002
Grade 4 ²	7.72	2.41	3.25	12.73	0.23	3.20	.002
Grade 5 ²	0.98	2.11	-2.96	5.21	0.04	0.46	.644
Grade 6 ²	0.78	2.31	-3.48	5.53	0.03	0.34	.734
Grade 7 ²	5.34	2.18	1.26	9.66	0.20	2.45	.015
Grade 8 ²	2.60	2.15	-1.41	6.93	0.10	1.21	.228
Grade size ¹	0.00	0.01	-0.01	0.02	0.03	0.48	.631
Student-teacher ratio ¹	-0.20	0.28	-0.72	0.32	-0.05	-0.72	.481
% Economically disadvantaged ¹	-0.32	0.07	-0.44	-0.19	-0.36	-4.81	< .001
% Special education ¹	-0.24	0.12	-0.47	0.00	-0.12	-1.95	.053
% English Language Learners (ELL) ¹	-0.17	0.09	-0.33	0.01	-0.13	-1.80	.074
Gender: % male ¹	-0.17	0.34	-0.80	0.45	-0.03	-0.51	.610
IXL ELA (SP/week >= 0.5)	3.99	1.28	1.61	6.38	0.19	3.12	.003

Note. *b* = unstandardized regression coefficient, *SE* = standard error, CI = confidence interval, β = standardized regression coefficient, SP/week = skills proficient per week (i.e., the average number of skills in which students reached proficiency each week).

¹ Mean-centered.

² Dummy coded; grade 3 as reference group.

Appendix C. Comparability of Matched Samples

