

## Goalbook Pathways: Impact in Lima City Schools

*March 2020*

*This report leverages product usage data and school district data to explore the impact of Goalbook Pathways usage on student outcomes.*

## Introduction and Principal Findings

The developers of Goalbook Pathways contracted with Evidently, Inc. to evaluate the effectiveness of their product in Lima City Schools, in Lima, OH. Goalbook Pathways is a platform that provides teachers with standards-based resources to personalize instruction so that all students can succeed. Evidently previously conducted a correlational study using data from the 2017-18 school year, which explored patterns of usage of Goalbook Pathways by teachers, and measured the association between the product usage and Scholastic Reading Inventory (SRI) outcomes in Spring 2018 for students in upper elementary and middle grades. The correlational study found positive associations of Goalbook Pathways with the SRI assessment, including positive correlations between product usage and SRI outcomes within important student subgroups, and positive associations of SRI outcomes with more powerful use of the Goalbook Pathways platform. Appendix A includes a summary of these results. The current impact study was commissioned to build on these findings. The analytic sample for this impact study pools all Lima City classes across three years of usage, including the 2016-2017, 2017-18, and 2018-19 school years.

This study analyzed the impact of Goalbook Pathways on student achievement in reading, as measured by student growth (defined as moving to a higher proficiency level between fall and spring test administration) on the SRI reading assessment. The principal findings are:

- Goalbook Pathways had an impact on the percentage of students who achieved growth on the SRI proficiency level. The impact is estimated at 5.3 percentage points.
- The strongest impacts were for male students and for black students.
- There was no evidence of differential impact of Goalbook between students with varying pretest levels; Goalbook had a positive impact for students in all proficiency levels.

## Results

We found a positive impact of Goalbook Pathways on student growth on the SRI assessment estimated at 5.3 percentage points; this difference is from the 46.8% probability of growth for comparison students to the 52.1% probability for students of teachers who use Goalbook Pathways. In other words, 11% more students in classes where teachers did not use Goalbook would have showed growth if their teachers had used it.<sup>1</sup>

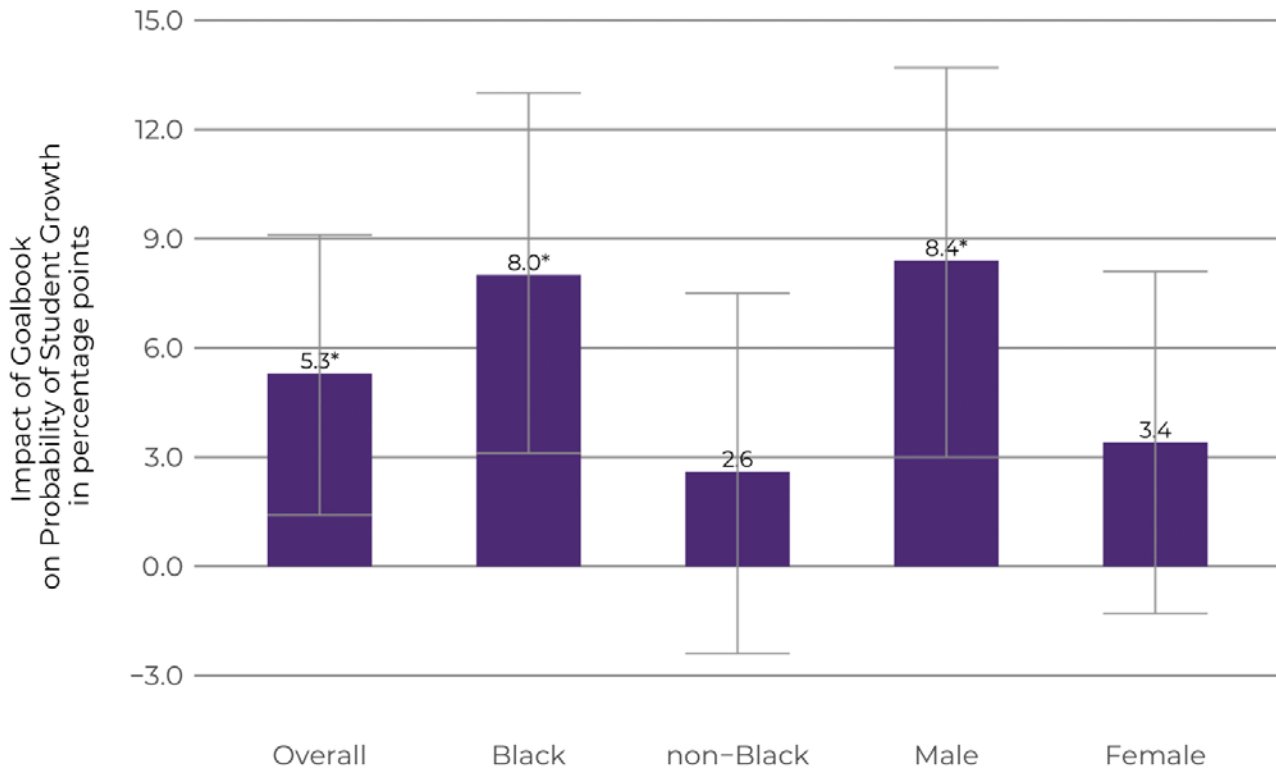
Importantly, this positive impact was found for subgroups of students in each proficiency level on the pre-test, and we did not find any differential impact favoring high-performing students (or any group of students based on pre-test level). Detailed results for each subgroup are found in Table 1.

We did find positive differential impacts of Goalbook Pathways on two categories of students. We found limited evidence that the impact of Goalbook Pathways is (1) greater for male students than for female students (8.0 percentage points vs. 2.6 percentage points,  $p = .09$ ) and (2) greater for black students than for non-black students (8.4 percentage points vs. 3.4 percentage points,  $p = .12$ ).<sup>2</sup> These findings, along with the average effect, are shown in Figure 1.

---

<sup>1</sup> The ratio of the Goalbook probability of student growth, 52.1, to the comparison probability of growth, 46.8, is 1.11, which implies that growth is about 11% more likely in the Goalbook group than in the comparison group.

<sup>2</sup> See Technical Details for interpretation of  $p$ -values.



**FIGURE 1. IMPACT OF GOALBOOK PATHWAYS ON ALL STUDENTS AND SELECTED SUBGROUPS**

Note: Gray bars illustrate the 95% confidence interval. There is strong evidence that the results fall within this range.

There was no evidence that the impact of Goalbook Pathways was different for special education students compared to non-special education students. We also found no evidence of differential impacts of Goalbook Pathways by grade level, but we caution that these findings may be limited by the sample size. Due to the composition of the sample, we could not look at impacts for students based on their socioeconomically disadvantaged status or English Language Learner status.

TABLE 1. DETAILED RESULTS

Category	Effect on probability of student growth	<i>p</i> value	Comparison probability	Goalbook probability	% of students with growth	95% confidence interval for effect on growth	<i>p</i> value for differential impact
Average Effect	5.3	<.01	46.8	52.1	11.3	(1.4, 9.1)	n/a
Below Basic	5.6	<.01	49.0	54.6	11.5	(1.5, 9.8)	<sup>a</sup>
Basic	5.8	<.01	50.8	56.6	11.4	(1.6, 10.0)	
Proficient	4.7	<.01	24.2	28.9	19.4	(1.8, 7.6)	
Advanced	2.8	.06	85.4	88.3	3.3	(-0.2, 5.8)	
Male	8.0	<.01	45.2	53.2	17.7	(3.1, 13.0)	.09
Female	2.6	.31	48.6	51.2	5.3	(-2.4, 7.5)	
Black	8.4	<.01	38.9	47.3	21.5	(3.0, 13.7)	.12 <sup>b</sup>
Non-Black	3.4	.16	52.0	55.4	6.5	(-1.3, 8.1)	
Special Education	4.1	<.01	25.9	30.0	15.8	(1.1, 7.1)	
Non-Special Education	5.7	<.01	51.1	56.8	11.2	(1.7, 9.8)	.41

<sup>a</sup> None of the subgroup impact estimates are pairwise different at  $p < .05$  for varying levels of pre-test; we don't have strong evidence of differential impact across levels of pre-test even though the means differ.

<sup>b</sup> We have limited evidence that the differential impact for black and non-black students was significant, but there was no significant evidence of differential impact between other racial subgroups.

# Technical Details

## DATA PREPARATION

This study of Goalbook Pathways is based on student data from Lima City Schools and teacher-level application usage data from Goalbook from the 2016-17, 2017-18, and 2018-19 school years. Student data from Lima City Schools included demographics, school and teacher identifiers, Fall pretest score, and Spring outcome score from the Scholastic Reading Inventory (SRI) for all students in grades 3 through 8. The SRI is a formative assessment administered in the fall, winter, and spring to measure reading comprehension. Goalbook data included the number of days of Pathways usage for each teacher.

## STUDY DESIGN

This study uses a quasi-experimental comparison group design. Students enrolled in classes taught by teachers who used Goalbook Pathways for at least two days in a given academic year were included in the treatment group. The remaining students were included in the comparison group. Inclusion of a student into the treatment group in a given year was made regardless of whether this student was or was not enrolled in classes where the teacher used Goalbook Pathways in a previous year, under an assumption that Goalbook has no direct cumulative effect on the student.

The outcome of interest in this study was student achievement in reading, as measured by the SRI assessment. One noteworthy feature of the available data is a substantial proportion of students in the sample performing at the “Below Basic” level – 15% - were scored as beginning readers (BR) and had no reported scale score. Using scale scores as the outcome would require excluding those lowest performing students from the analysis, which would introduce a substantial bias in the results. The analysis relies instead on the proficiency levels to include as many students as possible and define the student growth metric. Specifically, we define the outcome as the probability of “student growth”: that a student will increase their proficiency level from pre-test (fall administration) to post-test (spring administration), i.e. “Below Basic” to “Basic”, “Basic” to “Proficient”, “Proficient” to “Advanced”, or remain in the “Advanced” category. More information about how the proficiency levels relate to scale scores (Lexile measures) is provided in Appendix B.

## ANALYTIC SAMPLE

The analytic sample in this study included 4,252 students, with 1,061 in the Goalbook Pathways group and 3,191 in the comparison group. 718 students were excluded due to missing pretest or outcome test data. The sample was balanced with less than .25 standard deviations on each covariate and required no additional balancing or matching; the final parameters of the analytic sample are presented in Table 2.

TABLE 2. CHARACTERISTICS OF STUDY SAMPLE (STUDENTS)

Category	Comparison mean	Goalbook mean	Pooled Standard Deviation	Difference in % Standard Deviation
% Male	53	51	50	5
% White	36	35	48	3
% Black	40	49	49	2
% Hispanic	5	5	22	0.1
% Special Education	17	17	38	1
% Below Basic	40	47	49	15
% Basic	35	31	47	10
% Proficient	17	17	38	.7
% Advanced	8	5	26	13
Total Students	3191	1061	n/a	n/a

## ANALYSIS

The analysis was performed using a logistic mixed effect regression model to estimate the impact of Goalbook Pathways on the student growth metric, adjusting for the differences in treatment and comparison group composition, and taking into account clustering of students and teachers in schools.

The ' $p$  value' is the measure of the precision of the results or the strength of evidence that the effect in question is statistically different from zero. Conventional interpretation is that a  $p$  value of .05 or less signifies strong evidence, and  $p$  values above .05 but less than .20 provide limited evidence. Higher  $p$  values imply that our results provide no reliable information about the impact of usage on outcomes, since the probability that the true effect is zero—or even has an opposite sign—is too high. Higher  $p$  values (lower precision of the results) are typical when the sample of students is small and does not necessarily mean that there is no impact.

Disclaimer: It must be taken into account that these results are obtained using a relatively small sample in a particular school district and may not necessarily be replicable elsewhere. Model quality, defined as the percentage of correct predictions, was 65.1. This implies that



one-third of the variation in student outcomes is affected by factors not reflected in the data set, which is a limitation of the results.

## Appendix A: Correlational Study

The Evidently, Inc. research team originally identified Lima City, OH as a potential research site in 2019 by conducting an analysis of usage of the Goalbook Pathways product in schools/districts in the US during the 2017-18 school year. After entering a data-sharing agreement with the district, we obtained student reading outcomes from the Scholastic Reading Inventory as well as student demographic characteristics for students in grades 3 through 8. We matched these to product usage data collected from the Goalbook Pathways product. The study explored patterns of usage of Goalbook Pathways by teachers and measured the association between the product usage and SRI outcomes in Spring 2018. The principal findings were:

- The count of “Viewed Unpacked Standard” action is the single best predictor of student outcomes. Using this metric, Goalbook Pathways shows a strong evidence of promise.
- The strongest association with student outcomes are observed in eighth grade and for male and for black students.
- Teachers, rather than school or district leaders or students, primarily made their own decision to use Goalbook Pathways.
- There is a wide variation in Goalbook Pathways usage among teachers and across schools: the average days of usage range between 1.5 and 20 days across schools and vary between 0 and 73 days for individual teachers.

Out of the original sample of usage events provided by Goalbook, we used principal component analysis (PCA) to reduce the dimensionality of the data set. There were statistically significant positive correlations among all events, and the first component in the PCA accounts for most (63%) of the total variation among usage. Two types of events - "Viewed Standard Resources" and "Showed Item Preview" - stand out as the most widely used and most correlated with overall usage score (first principal component), while "Viewed Unpacked Standard" follows closely in terms of correlation but is less frequently used. In looking at teacher-level usage of Goalbook, we identified a group of “Active/frequent users” and a group of “Occasional” users.

## Appendix B: Scholastic Reading Inventory Levels

The following table displays the Lexile levels associated with proficiency levels, as provided on page 53 of the 2014 SRI (Scholastic Reading Inventory) Technical Guide.

**TABLE 3. PERFORMANCE STANDARD PROFICIENCY BANDS FOR THE READING COMPREHENSION ASSESSMENT, IN LEXILE MEASURES, BY GRADE**

Grade	Below Basic	Basic	Proficient	Advanced
K	N/A	BR	0L to 275L	280L and Above
1	BR	0L to 185L	190L to 530L	535L and Above
2	BR to 215L	220L to 415L	420L to 650L	655L and Above
3	BR to 325L	330L to 515L	520L to 820L	825L and Above
4	BR to 535L	540L to 735L	740L to 940L	945L and Above
5	BR to 615L	620L to 825L	830L to 1010L	1015L and Above
6	BR to 725L	730L to 920L	925L to 1070L	1075L and Above
7	BR to 765L	770L to 965L	970L to 1120L	1125L and Above
8	BR to 785L	790L to 1005L	1010L to 1185L	1190L and Above
9	BR to 845L	850L to 1045L	1050L to 1260L	1265L and Above
10	BR to 885L	890L to 1075L	1080L to 1335L	1340L and Above
11/12	BR to 980L	985L to 1180L	1185L to 1385L	1390L and Above

As of this report date, the SRI Technical Guide can be found here:

[https://www.hmhco.com/products/assessment-solutions/assets/pdfs/sri/SRI\\_TechGuide.pdf](https://www.hmhco.com/products/assessment-solutions/assets/pdfs/sri/SRI_TechGuide.pdf)

Reference this report: Empirical Education Inc. (2020). *Goalbook Pathways: Impact in Lima City Schools*. (Empirical Education Rep. No. Empirical\_Goalbook-6062-EVI-2020-O.2). Empirical Education Inc. [https://www.empiricaleducation.com/pdfs/Empirical\\_Goalbook-6062-EVI-2020-O.2.pdf](https://www.empiricaleducation.com/pdfs/Empirical_Goalbook-6062-EVI-2020-O.2.pdf)