Efficacy Study of Zearn Math in a Large Urban School District

Jennifer R. Morrison, Ph.D.
Betsy Wolf, Ph.D.
Steven M. Ross, Ph.D.
K. L. Risman, M.A.
Caitlin C. McLemore, M.Ed.

April 2019

Johns Hopkins School of Education
Center for Research and Reform in Education

Contents

EXECUTIVE SUMMARY:
Zearn Math Evaluation Report: Year One Findings

Zearn Math provides an integrated approach to math teaching and learning, connecting a print- and software-based curriculum with a rotational classroom model, professional development, and classroom- and school-level reports on student learning. Zearn Math's K-5 curriculum is used by over 1.5 million elementary school students and 50,000 teachers across the United States. Each day with Zearn Math, students learn in a rotational model that allows students to learn new grade-level content in two ways: independently with engaging digital lessons and in small groups with their teacher and classmates. Students learn at their own pace working in *Independent Digital Lessons* through fluency, guided practice, and independent practice. While half the class works in the self-paced station, teachers provide instruction to the other half of the class. To support districts and schools with implementation, Zearn Math offers, at cost, year-long professional development, *School Accounts* with administrator-level implementation services, and print materials.

The purpose of the present study was to gather summative and formative data related to the implementation of Zearn Math during the first year of a two-year study. The sample included students in a large urban school district comprised of predominantly African American (58%) and Hispanic (28%) students. Roughly 90% of students are economically disadvantaged. As of the 2016-17 school year, only 8% of the district's students in Grades 3-8 met mathematics proficiency levels as compared to 40% of all students in the state. The current research was conducted during the district's second year of implementation of Zearn Math. During the 2016-17 school year, the district piloted the curriculum in two schools and provided some implementation support. The following year, 15 elementary schools began learning with Zearn Math and implementing the rotational classroom model, while the remainder of the elementary schools in the district continued to learn with EngageNY. This study explores the treatment schools' implementation of Zearn Math and the impacts of the shift during the first year of implementation.

This study includes findings from multiple data sources, including classroom observations and focus groups with teachers from four schools randomly selected for a site visit, student and teacher questionnaires, interviews with school administrators and Zearn Math personnel, student usage data across all schools, and student achievement data from two standardized mathematics assessments (NWEA MAP and the state assessment). Key findings of the current study include:

Administrators, students, and teachers have, overall, very positive perceptions of Zearn Math. Among the most robust findings from the current study is that teachers and administrators want to use Zearn Math in the future and would recommend it to their peers. Teachers generally perceived student- and teacher-facing features as effective, with few exceptions. The teacher reports and half-class model were viewed as notable strengths of Zearn Math. In addition, students expressed overwhelmingly positive feelings about the curriculum—almost 90% of students agree it is good for learning mathematics, and roughly 80% of students look forward to using it each day.

Teachers in the treatment schools have integrated specific teacher practices prescribed by Zearn Math into their mathematics instruction routines. In the teacher questionnaire, approximately 80% of teachers reported teaching with Zearn Math allows them to differentiate instruction based on the needs of their students. In focus groups, teachers described using individual- and classroom-level data to inform instruction and planning. Additionally, teachers at treatment schools were significantly more likely than comparison school teachers to believe they knew what to do to increase students' engagement in mathematics and felt more confident answering their students' mathematics questions. The reader should exercise caution in interpreting these results, however, due to a low response rate by comparison group teachers.

Administrators and teachers cite Zearn Math's rotational half-class model as a major strength of the curriculum, while also noting the shift away from whole class instruction involved some initial discomfort and a period of adjustment during the first year of implementation. In the teacher questionnaire, approximately 80% of teachers reported teaching with Zearn Math allows them to differentiate instruction based on the needs of their students. Teachers and administrators describe the half-class rotational model, where students rotate between *Independent Digital Lessons* and teacher-led group learning each day, as a key strength of Zearn Math. Teachers explained that they are better teachers in smaller groups in general and that small-group teaching is more effective for mathematics in particular. However, all stakeholder groups described a period of adjustment as teachers shifted from teaching whole class to the half-class format. The change involved some discomfort among teachers and classroom management issues, primarily arising while students were learning independently. Administrators described improved classroom conditions after teachers had opportunities to share effective classroom management techniques with each other.

Multiple data sources indicate that some teachers felt better prepared to implement Zearn Math than others, and stakeholders explained there were limited professional development opportunities prior to the start of the school year. Just under 50% of teachers felt prepared to implement Zearn Math in their classroom. However, teachers and administrators tended to qualify their own perceptions of teacher unpreparedness, describing minimal PD and learning on the job as business-as-usual. Interviews with multiple stakeholders, including Zearn personnel, revealed that the professional development provided to the district was not ideal in terms of timing relative to the beginning of the school year. Additionally, while online training was available to all teachers, there was no way to confirm participation. Deficits in teacher preparation likely explain why teachers most frequently reported not using certain features because they were unaware it existed and may help explain the discomfort teachers described during the adjustment to the half-class rotational format of mathematics instruction.

Administrators and teachers described definitive increases in student engagement in mathematics content. Multiple data sources suggest that teachers and administrators have observed increased student engagement in mathematics in general, and that Zearn Math is widely enjoyed by their students. Both groups described the majority of students as motivated to learn using technology and positively impacted by the competition and "leveling up" embedded in *Individual Digital Lessons*. Administrators notably perceived that students remained engaged during their learning with Zearn Math regardless of skill level.

At the time of qualitative data collection, school personnel expressed cautious optimism related to student learning outcomes. In interviews and focus groups, teachers and administrative teams tended to espouse "wait and see" responses to prompts specifically about student achievement. In the teacher questionnaire, roughly 60% of respondents agreed that Zearn Math is effective for increasing student achievement while 30% of teachers responded neutrally. Respondents were more confident about the degree to which Zearn Math promoted higher-order thinking skills; roughly 70% of teachers agree that *Small Group Lessons* and *Individual Digital Lessons* promote these skills.

After controlling for a number of student-level variables, differences in achievement gains on the NWEA MAP and the state assessment mathematics exams were not significant for the overall treatment and comparison samples but significantly favored Zearn Math for certain student subgroups on each exam. Regression-adjusted NWEA MAP mathematics scores were 187.88 for comparison students and 188.37 for treatment students. This difference reflected an effect size of +0.024, which slightly favored the treatment group, though there was not a statistically significant difference. Regression-adjusted state mathematics assessment scores were 0.03 (in standardized units) for comparison students and -0.053 for treatment students, and the difference was not statistically different. Subgroup analyses related to prior achievement and grade level revealed some statistically significant observations favoring the treatment group; however, taken together, subgroup analyses were difficult to interpret with confidence given the inconsistency of observations between analytic samples, non-equivalence at baseline for some prior achievement subgroups, and the likelihood that usage patterns confound program effects.

Usage of Zearn Math was positively and statistically significantly related with achievement gains on the NWEA MAP and the state mathematics assessment. Usage variables included in the current analysis were associated with achievement gains for treatment students relative to comparison students in both analytic samples. These usage variables include total minutes using Zearn Math and total lessons completed. While gains related to increased usage were all positive and statistically significant, they were small in magnitude.

Multiple data sources indicated that treatment schools did not meet Zearn Math recommendations for weekly usage minutes or averages lessons completed and that implementation varied between and within treatment schools. On average, students in grades 1-5 used the *Independent Digital Lessons* for 72 minutes per week and completed roughly 2 lessons per week, below the Zearn Math recommended 120 minutes per week and 4 completed lessons per week. Average usage among students in grades 1-5 varied widely between and within treatment schools. For example, across all treatment schools, students in grades 3-5 averaged over 1,000 more minutes total in the digital component than students in grades 1 and 2. Within schools, usage data indicate that some grades used the digital component as intended, while others used it infrequently. While a majority of the teachers observed implemented Zearn Math with moderate to high fidelity, multiple data sources indicated that expectations for implementation established by school-level administrators were inconsistent across treatment schools and were inconsistent between district- and school-level administrators. Teachers also identified shortened mathematics blocks and a lack of headphones and devices for their students, as well as students being off-task, as barriers to successful implementation.

Some students may have found Zearn Math's level of rigor and independent learning component very challenging. In the student questionnaire, students in treatment and comparison schools report a relatively high degree of interest in mathematics and mathematics-related self-efficacy. However, students in comparison group schools reported significantly higher agreement with prompts related to mathematics knowledge and mathematics efficacy as compared with treatment students. A potential explanation for these findings is that treatment students may have found Zearn Math material more challenging than previously experienced, which may have affected their feelings toward mathematics in general. Additionally, while teachers and administrators are overwhelmingly satisfied with the curriculum and are optimistic about its use in the future, they also were emphatic about the extent to which the very format of Zearn Math—the demand for individual learning, and the integration of listening, reading, typing and writing—cannot accommodate the learning needs of some students in the district. The most frequent concern from teachers was the perception that Zearn Math is unable to facilitate targeted remedial work to address specific gaps in knowledge.

Recommendations

Based on the findings on this evaluation, we offer the following suggestions for Zearn Math and the district to improve the curriculum and implementation:

Provide district and school administrators with targeted support to lead implementation of Zearn Math.

District and school administrators will benefit from additional professional development and support materials that are geared toward the role of administration during implementation. Professional development for administrators should offer strategies for supporting teachers during the transition to the half-class rotational model, engaging resisters, and should address the importance of fidelity in achieving positive student outcomes.

Provide teachers with resources that are easily accessible and that support understanding and usage of Zearn Math features.

Among the most concerning findings from the current study is the lack of knowledge teachers have about certain Zearn Math features, pervasive perceptions of teacher unpreparedness and the less than favorable feedback related to the professional development teachers did receive prior to implementation. Circumstances may be unique within the district, but Zearn may consider strategies to increase awareness of features and access to as-needed professional development. Zearn may also consider updating features currently perceived as ineffective and centralizing teacher resources that are currently spread across multiple platforms (e.g., YouTube channel, Zearn Facebook page, Zearn website).

Provide teachers and administrators with strategies for supporting students during independent learning.

Multiple data sources suggest students experience both positive and negative emotions extending from the gamified components of Zearn Math, and that teachers experience some difficulty with classroom management in the half-class rotational format. Zearn may consider further research into the negative emotions experienced by students, and how to translate negative emotional states into positive learning experiences. Zearn may also consider including specific guidance about classroom routines and teacher practices that help students persevere during *Independent Digital Lessons*.

Consider providing more flexibility for school-based educators to modify digital components of the curriculum to fit individual school and student needs.

We believe that mathematics teachers and coaches will continue to emphasize the needs of a small and important group of students who need targeted skills interventions. Zearn Math may consider how to increase flexibility within the curriculum to fit unique needs while maintaining cohesiveness of lessons and emphasis on implementation fidelity.

An Evaluation of Zearn Math in a Large Urban School District: Year One Findings

Zearn Math provides an integrated approach to math teaching and learning, connecting a print- and software-based curriculum with a rotational classroom model, professional development, and classroom- and school-level reports on student learning. Zearn Math's K-5 curriculum is used by over 1.5 million elementary school students and 50,000 teachers across the United States. Each day with Zearn Math, students learn in a rotational model that allows students to learn new grade-level content in two ways: independently with engaging digital lessons and in small groups with their teacher and classmates. Students learn at their own pace working in *Independent Digital Lessons* through fluency, guided practice, and independent practice. While half the class works in the self-paced station, teachers provide instruction to the other half of the class. To support districts and schools with implementation, Zearn Math offers, at cost, year-long professional development, *School Accounts* with administrator-level implementation services, and print materials.

The purpose of the present study was to gather year one data as part of a two-year efficacy evaluation of Zearn Math in an urban school district serving a relatively large population of economically disadvantaged and minority students. Interviews, questionnaires, classroom observations, focus groups, digital component usage data and student achievement data yielded summative information related to Zearn Math and formative feedback for future Zearn Math implementations. The current study was designed to address the following research questions:

1. How is Zearn Math implemented in the district overall and across schools?
    - How does implementation vary by grade level, school type, and level of support from school administrators?
    - To what degree do schools implement with high fidelity?
    - What are best implementation practices and barriers to full implementation?

2. Does adoption of Zearn Math lead to differences in instructional practice as measured by teachers' self-reflections and observation of teaching?
    - To what degree are teachers satisfied with Zearn?
    - What types of refinements or resources might improve satisfaction and support changes in practice?

3. Does the adoption of Zearn Math relate to improved student performance on standardized mathematics assessments?
    - To what degree does Zearn Math lead to greater academic growth and achievement in mathematics?
    - To what degree does Zearn Math lead to shifts in student attitudes (engagement, self-confidence in mathematics ability, enjoyment of

mathematics)?

- Do impacts on student outcomes vary by student characteristics (gender, ethnicity, prior achievement)?

4. How do implementation quality and usage levels of Zearn Math relate to teacher and student outcomes, including student performance on standardized mathematics assessments?

Method

Research Design

The current study employed a mixed-methods evaluation design, including qualitative and quantitative data collected from students, teachers, school administrators, and Zearn Math personnel. There are two major components of the current two-year evaluation that begins in 2017-18. The first involved site visits to four case study schools, randomly selected from the group of schools that implemented Zearn Math in the 2017-2018 school year. Site visits involved classroom observations, teacher focus groups, and administrator interviews. The second major component involved an analysis of student achievement data from all schools that implemented Zearn Math in 2017-2018 (treatment schools). Student math achievement data from treatment schools were compared to data from schools not currently implementing Zearn Math (comparison schools) to determine if treatment students demonstrated greater gains in achievement than comparison students.

The evaluation design addressed the summative needs of providing evidence of implementation and the formative needs of providing recommendations for general curriculum improvement. Toward this end, the study sought to achieve the following: (a) collect evidence from multiple data sources that can help explain student achievement outcomes; (b) generated outcomes other than student achievement for both descriptive and comparative analyses; (c) document implementation and application context through school visits and classroom observations; (d) estimate program effects by comparing student math achievement gains from treatment schools (schools implementing Zearn Math) with comparison schools (schools not implementing Zearn Math).

Participants

The districts is a large urban school district serving more than 25,000 students in 50 schools, including 23 grades K-6 elementary schools and 12 grades K-8 combined elementary and middle schools. Three major ethnic groups account for over 95% of the student enrollment: African American (58%), Hispanic (28%), and White (10.1%). Close to 20% are students with disabilities (SWDs), 15% are English language learners

(ELLs), and 90% are economically disadvantaged.[1] In 2017, just 8% of the district's students in grades 3-8 were proficient in grade-level mathematics.

At the time of data collection (Spring 2018), the district was in the second year of implementation of Zearn Math. In 2016-2017, the district piloted the curriculum in two schools and provided some support for implementation. In 2017-2018, the district's mathematics director met with principals across the district and invited them to serve as treatment schools in the present study. This recruitment effort resulted in 15 elementary schools that implemented the curriculum. All schools were provided with up-front professional development from Zearn and ongoing support from the district throughout the year. Prior to Zearn Math, district schools implemented EngageNY (also known as Eureka Math), which served as the foundation for Zearn Math.

Analytic Samples. Because student achievement data were gathered from two different assessments (NWEA MAP and the state assessment), two distinct analytic samples will be described. The NWEA MAP analytic sample includes students in treatment schools in grades 1-5 and students in comparison schools in grades 1-5. Kindergarten students in treatment schools were not included in the analytic sample because these students do not have a pre-test assessment score and are therefore excluded from the analysis. Data from kindergarten students in treatment schools are presented descriptively. The state assessment sample includes students in treatment and comparison schools in grades 4 and 5. Sample sizes vary because the two analytic samples differ on student grade levels. In addition, parents may opt out of student testing and are more likely to do so for the NWEA MAP assessment.

Student characteristics for treatment group students in kindergarten and grades 1-5, and comparison students in grades 1-5 in the NWEA MAP sample are displayed in Tables 1 and 2. The treatment group was comprised of students in kindergarten (N = 517) and grades 1-5 (N = 3,803) in the 15 elementary schools that implemented Zearn Math in the 2017-2018 school year. The comparison group (N = 4,915) was comprised of students in grades 1-5 that attended any of the 20 elementary schools not implementing Zearn Math.

[1] New York State Department of Education, 2017

Table 1
Treatment student characteristics for NWEA MAP sample (kindergarten)

Kindergarten	All
% Female	47.39
% Black	56.29
% White	14.70
% Asian	1.55
% Latino	27.08
% Other race/ethnicity	0.39
% ELL	8.70
% Students with IEP	14.31
N	517

Table 2
Treatment and comparison student characteristics for NWEA MAP analytic sample (grades 1-5)

Grades 1-5	All	Treatment	Comparison
% Female	48.05	48.7	47.55
% Black	57.25	59.19*	55.75*
% White	9.15	11.41*	7.41*
% Asian	3.18	1.79*	4.25*
% Latino	29.88	27.14*	32.00*
% Other race/ethnicity	0.54	0.47	0.59
% ELL	13.08	7.63*	17.29*
% Students with IEP	18.89	19.12	18.72
% students in grade 1	17.96	16.30*	19.25*
% students in grade 2	19.47	18.88	19.92
% students in grade 3	21.87	23.01*	21.00*
% students in grade 4	20.69	21.61	19.98
% students in grade 5	20.00	20.19	19.86
N	8,718	3,803	4,915

*Statistically significantly difference between treatment and comparison group at $p < .05$

As displayed in Table 2, some statistically significant differences emerged between the treatment and comparison group in the NWEA MAP sample. Treatment schools enrolled greater proportions of black and white students and lower proportions of Asian, Latino, and ELL students. The treatment group was comprised of fewer students in first grade and more students in fourth grade. While statistically significant due in part by large student sample sizes, these variations in general seem fairly small in magnitude and not indicative of the treatment and comparison populations differing meaningfully.

Table 3
Baseline achievement characteristics for treatment and comparison students (NWEA MAP sample grades 1-5)

	All	Treatment Mean	Comparison Mean	Standardized Mean Difference
Overall Sample	176.41	178.62	174.70	0.194
Grade 1	146.08	149.15	144.07	0.317
Grade 2	165.71	166.59	165.07	0.058
Grade 3	180.78	182.69	179.17	0.155
Grade 4	188.55	189.21	188.01	0.077
Grade 5	196.71	197.71	195.92	0.152
Low Prior Achievement	167.11	169.35	165.51	0.209
Mid Prior Achievement	189.77	189.65	189.87	0.021
High Prior Achievement	206.01	206.38	205.57	0.001
N	8,718	3,803	4,915	

As displayed in Table 3, baseline equivalence in mathematics achievement between the treatment and comparison groups in the NWEA MAP sample is established because the standardized mean difference for the overall analytic sample and for most of the subgroups is less than 0.25 standard deviations (WWC, 2017). The one exception is that baseline equivalence was violated in grade 1 because treatment students had higher baseline achievement than comparison students by more than 0.25 standard deviations. While the statistical models control for baseline mathematics achievement, they cannot fully account for differences in treatment and comparison students' mathematics abilities.

Student characteristics for treatment and comparison students in grades 4 and 5 in the state assessment analytic sample are displayed in Table 4. The treatment group was comprised of students in grades 4 and 5 (*N* = 1,493) in the 15 elementary schools that implemented Zearn Math in the 2017-2018 school year. The comparison group *(N* = 1,725) was comprised of students in grades 4 and 5 that attended any of the 19 elementary schools[2] not implementing Zearn Math.

[2] The state assessment sample comparison group include 19 elementary schools whereas the NWEA MAP assessment sample included 20 comparison schools. Only students in grade 3 and higher take the state assessment. One school enrolls students in pre-kindergarten through second grade. Students at this school completed the NWEA MAP assessment but not the state assessment.

Table 4
Treatment and comparison student characteristics for state assessment analytic sample (grades 4 and 5)

Grades 4 and 5	All	Treatment	Comparison
% Female	46.74	48.23	45.45
% Black	57.77	59.81	56.00
% White	8.11	10.99*	5.62*
% Asian	3.51	2.34*	4.52*
% Latino	30.24	26.59*	33.39*
% Other race/ethnicity	00.37	00.27	00.46
% ELL	12.68	8.17*	16.58*
% Students with IEP	22.34	21.30	23.25
% students in grade 4 (at posttest)	51.27	52.24	50.43
% students in grade 5 (at posttest)	48.73	47.76	49.57
N	3,218	1,493	1,725

*Statistically significantly difference between treatment and comparison group at $p < .001$

As displayed in Table 5, some statistically significant differences emerged between the treatment and comparison group in the state assessment sample. Comparison schools enrolled lower proportions of white students and greater proportions of Asian, Latino, and ELL students. While statistically significant, these variations are fairly small in magnitude and not indicative of substantial differences in the two populations, especially in light of baseline equivalence of pretest achievement scores between the two groups at pretest (see Table 5).

Table 5
Baseline achievement characteristics for treatment and comparison students (state assessment sample)

	All	Treatment Mean	Comparison Mean	Standardized Mean Difference
Overall Sample	0.000	0.093	-0.080	0.180
Grade 4	0.129	0.220	0.048	0.145
Grade 5	-0.136	-0.046	-0.210	0.185
Low Prior Achievement	-0.565	-0.504	-0.615	0.184
Mid Prior Achievement	0.765	0.766	0.764	0.019
High Prior Achievement	1.684	1.746	1.605	0.290
N	3,218	1,493	1,725	

State assessment scores were standardized because the scale employed by the exam changed from the 2016-2017 to 2017-2018 school year. The scale in 2016-2017 employed a 100-300 range. The scale was changed to a 500-600 range in 2017-2018.

As a result of standardization, the test scores used in the current analysis had a mean of 0 and a standard deviation of 1 and ranged from about -3 to +4.

As displayed in Table 5, baseline equivalence in achievement was satisfied with the overall sample because the difference in prior achievement for treatment and comparison students was 0.18, which is less than the criterion of 0.25 standard deviations (WWC, 2017). However, the treatment group had noticeable higher prior math achievement than the comparison group by an average of 0.18 standard deviations. Baseline equivalence on the state mathematics assessment was also met for specific subgroups, including the grade 4, grade 5, low-pretest, and mid-pretest groups. Baseline equivalence was violated, however, for the high-pretest group because the difference (0.29 standard deviations) was greater than the criterion of 0.25 standard deviations (WWC, 2017). Thus, any findings regarding the high-pretest group on the state math assessment should be interpreted with caution, given that treatment students were higher achieving than comparison students prior to any participation in Zearn Math.

Measures

Data sources for the current study include classroom observations, focus groups with teachers, student and teacher questionnaires, interviews with school administrators and Zearn Math personnel, student usage data, and student achievement data. Observations, focus groups, and interviews with school administrators occurred in four randomly selected case study schools. Student and teacher questionnaires were administered to all treatment and comparison schools. Zearn Math provided student usage data to CRRE. Student achievement data were gathered from two mathematics assessments: the NWEA MAP and the state assessment.

Classroom observations. Classroom observations occurred in all four randomly selected case study schools. Observations lasted approximately 20 minutes each and occurred in four to six classrooms at each school. Classroom observations focused on how and to what extent Zearn Math was implemented in the classroom and specific student and teacher behaviors relative to curriculum use, including instructional practices and student engagement, as well as overall impressions of the classroom (see Appendix A).

Teacher focus groups. Teacher focus groups occurred at all four case study schools. Each focus group included three to six teachers and lasted approximately 45 minutes. The interview protocol (see Appendix B) solicited teachers' descriptions of and reactions to implementation of Zearn Math in their classroom, changes in teaching practices, and perceived impacts of the curriculum on student outcomes.

Administrator interviews. Administrator interviews occurred at all four case study schools. An interview protocol (see Appendix C) was developed to provide opportunity for principals and other administrators to provide descriptions of and reactions to implementation, changes in teacher practice, and student impact. Administrator interviews lasted approximately 45 minutes and were conducted at each of the four visited schools. A total of 10 administrators participated in these group interviews; administrators included principals, vice-principals, and mathematics coaches.

Zearn Math personnel interviews. An interview protocol (see Appendix D) was developed to provide opportunity for Zearn Math personnel to provide descriptions of and reactions to implementation at the district and provide additional context for the evaluation. Personnel interviews were conducted by phone. The interviews lasted approximately 45 minutes and were conducted with two employees at Zearn who have had direct contact with the district prior to and during implementation.

Teacher questionnaire. The CRRE Zearn Math Teacher Reaction Questionnaire was co-developed by CRRE, Zearn Math, and the district. The questionnaire consisted of approximately 20 Likert-type and multiple-choice items focusing on teaching efficacy (e.g., confidence about one's ability to teach mathematics), curriculum usage, perceived impacts of the curriculum on students, and overall impressions of Zearn Math. The questionnaire also collected information about total mathematics instruction time, student-teacher and student-device ratios, and perceived barriers to successful implementation. The teacher questionnaire was administered to teachers in treatment and comparison schools; a modified version including only items related to teaching efficacy was administered to teachers in comparison schools. A total of 198 teachers completed the teacher reaction questionnaire for a response rate of 37.2%; 54 responses are from teachers in comparison schools (33.8% response rate), and 144 responses are from teachers in treatment schools (38.7% response rate). Descriptive statistics and frequencies for the questionnaire are presented in Appendix E.

Student questionnaire. The CRRE Zearn Math Student Reaction Questionnaire (see Appendix F) was co-developed by CRRE, Zearn Math, and the district. The questionnaire consisted of seven Likert-type items focusing on students' mathematics efficacy (e.g., confidence about one's ability to succeed in mathematics) and motivation to learn mathematics in general, and their current experiences regarding Zearn Math. The student questionnaire concluded with two open-ended items to further capture student reactions to the curriculum. A modified version, which did not include questions specifically related to Zearn Math, was administered to students in comparison schools. A total of 1,790 students completed the student reaction questionnaire (55.6% response rate); 508 responses are from students in comparison schools (29.4% response rate), and 1,282 responses are from students in treatment schools (85.9% response rate).

Usage data. Zearn Math provided student usage data to CRRE via a secure web-based service provided by Johns Hopkins University. Metrics included students' total minutes using Zearn Math, total lessons completed (by grade level), and total weeks active in the digital component. The research team conducted basic equations to determine students' average minutes per week, average lessons completed per week, and proportion of lessons completed at/below grade level. Usage data are presented in full in a later section of this report. Additional usage data are presented in Appendix H.

Student achievement data. Student achievement data were the standardized end-of-grade mathematics exams administered to all district students (NWEA MAP), and the New York State Assessment in mathematics. Achievement scores from both assessments in spring of 2017 was used to establish baseline equivalence. Achievement scores in mathematics from spring 2018 were used to estimate program impact.

Analytical Approach

All qualitative data were analyzed using a grounded theory approach (Glaser & Strauss, 1967). Recorded data were transcribed, and handwritten observation notes were compiled, and uploaded into qualitative analysis software NVivo (QSR International). Qualitative data were organized by data source and analyzed using an iterative coding process. Within each data source, a structure of codes emerged from patterns in participant voices and/or team members' field notes. All codes were consistently reviewed for uniqueness and cohesion. Main findings are patterns that emerged strongly from multiple data sources.

Questionnaire data were analyzed descriptively using quantitative software (SPSS). Independent samples *t*-tests were used to compare questionnaire responses on items administered to both treatment and comparison students. Frequencies and descriptive statistics for the teacher questionnaire and student questionnaire are presented in Appendix E and Appendix F, respectively.

Student achievement data were analyzed using quantitative analysis software (STATA). For kindergarten students, achievement data were analyzed descriptively because the majority of kindergarten students were missing pretest data, precluding an analysis of program effects on math improvement. For student achievement analysis of NWEA MAP data from grades 1-5, and state assessment data for grades 4 and 5, we used two-level hierarchical linear modeling (HLM) to compare improvement in mathematics achievement between the springs of 2017 and 2018 for treatment and comparison students. Specifically, this analysis compared spring 2018 mathematics achievement for treatment and comparison students while controlling for the following variables: student's baseline reading and math performance, grade level, race/ethnicity, gender, and special education and ELL designations.[3] The model also, by default,

[3] Student socio-economic status or economic disadvantage was not available.

controlled for school fixed and random effects.

In addition, similar analyses were conducted to estimate the extent to which digital component usage related to improvement on students' mathematics scores. We examined associations between improvement in mathematics scores and total minutes spent participating in the digital component, total lessons completed, total weeks active in the digital component, average minutes per week using the digital component, average lessons completed per week, and proportion of on-grade level lessons completed, while also controlling for the covariates listed above. We also explored usage data descriptively for treatment students and calculated average usage and lessons completed per week by grade and school.

We also conducted a number of supplementary analyses to determine if the treatment effect was consistent across different student subgroups based on student demographics, prior student achievement, and grade level. Subgroup analyses were conducted by adding interaction terms between the treatment and subgroup indicators into the HLM models. Subgroups were largely predetermined (e.g., ELL designations, grade-levels, gender). The research team did create prior achievement subgroups for each analytic sample. For prior achievement subgroups of NWEA MAP data, we defined low prior achievement as scoring at or below the 25th percentile and high prior achievement as scoring at or above the 75th percentile, respectively, on the 2017 spring NWEA MAP mathematics assessment. For prior achievement subgroups of State Assessment data, we defined prior achievement in terms of performance levels (Level 1, 2, 3, or 4). Low prior achievement was defined as scoring at or below a Level 1 on the 2017 spring mathematics assessment and high prior achievement was defined as scoring a 3 or 4. Students in the higher levels of 3 and 4 were combined given the small overall proportion of students in these categories. The vast majority (88%) of students who scored at the lowest level on the state math assessment also scored at or below the 25th percentile on NWEA MAP math test, and 91% of students who scored at one of the two highest levels on the state math assessment also scored at or above the 75th percentile on the NWEA MAP math test. Thus, the prior achievement subgroups overlapped across the two assessments, but the NWEA MAP sample differed from the state sample due to grade levels included.

Results

The following sections includes findings from all data sources related to implementation, professional development, fidelity, digital component usage, student and teacher outcomes, student and teacher response, student achievement, and overall stakeholder perceptions.

Preparation and Implementation

Participants were asked to describe their experiences and impressions of curriculum implementation in the district at large and, for teachers and administrators, specifically at their school. Usage data provided by Zearn provide accurate descriptions of weekly usage duration by students and teachers, and questionnaire data and classroom observations shed light on how teachers have implemented the curriculum in their classrooms. Findings in this section include a description of implementation at each case study school, including stakeholder perceptions of preparation to implement and overall impressions of implementation, as well as impressions of implementation fidelity and student and teacher usage.

School and district context. Case study schools were three traditional K-5 schools and one K-8 school. All four schools had a relatively high representation of students who receive free or reduced lunch. Two case study schools were currently in receivership, and one had just emerged from receivership last year. All four schools served a relatively high proportion of African American and Hispanic students. Principals at two schools spoke about modest gains in student performance in English and Language Arts in recent years. One school was currently part of a school-wide mathematics intervention, which included a full-time on-site mathematics coach provided by the district. All schools in the current study are "schools of choice", meaning that students do not have to live in the neighborhood to attend the school.

Two case study schools had fully implemented Zearn Math in all classrooms in grades 1-5, whereas one school had implementation only in grades 3-5. The fourth school implemented in grades 3-5 and was described as "moving toward full implementation" by introducing the curriculum to individual teachers in younger grades. Administrative teams were asked to describe why they decided to implement Zearn Math instead of other curricula. Two administrative teams cited previous knowledge or recommendations from other school leaders while two schools cited the perceived fit of the curriculum with the mission and goals of their school. One school described Zearn Math as the best core mathematics curriculum option that fit with the 1:1 technology emphasis of the school; the other cited the schools' prior experience with EngageNY and a somewhat urgent need to match the gains the school recently saw in ELA. This administrative team described blindly accepting the opportunity to be a pilot school because they were eager to "try something new."

In interviews, Zearn personnel described their relationship with the district as typical of school districts where Zearn is providing targeted support. Many districts across the country are implementing Zearn Math but not all of them need or are open to the type of relationship Zearn maintains with the district. Personnel explained that the company often reaches out to districts where teachers have found Zearn Math on their own to offer support. This is exactly the case with the district; a critical mass of teachers began to use the curriculum and Zearn reached out, establishing contact with the district administrator who oversees mathematics curriculum in the district.

Both Zearn personnel interviewed described the district as one well positioned to be a Zearn Math success story because of the number of teachers opting in to the curriculum on their own, the history of using EngageNY mathematics curriculum, and the relative room for improvement in student achievement scores. Both also noted the unique role district administrators have played in the implementation and expansion of Zearn Math in the district and ultimately the selection of the district as the site for the program evaluation. Zearn personnel explained that Zearn Math tends to fare better in districts when teachers discover the product on their own and cultivate their own excitement. They explained that when teacher excitement is met with administrative support, the likelihood of positive outcomes increases dramatically. One of the personnel explained,

> *Administration has to be on board and intentional about building teachers' understanding of how digital lessons work and of small group instruction content. If administration is on board, they are going to make sure everything falls into place around that.*

Both Zearn personnel expressed optimism about the implementation in general and the future impact Zearn Math may have on student achievement in mathematics in the district.

Initial training. Zearn offered two days of onsite professional development in August 2017 for approximately 52 teachers and administrators from all 15 implementing schools. All schools had at least one administrator attend the training. Some schools sent mathematics coaches, lead teachers in mathematics, or grade-level leaders. Each session facilitated by Zearn lasted three hours and was designed to support school representatives to return to their buildings and facilitate PD for teachers implementing Zearn Math. Teachers also completed three hours of digital professional development prior to the school year, although district leaders described it as "impossible" to estimate how many teachers actually completed the digital PD.

According to questionnaire responses, teachers conveyed mixed perceptions regarding their preparation and support for implementation (see Figure 1). Just under half of respondents agreed that they felt prepared to implement Zearn Math in their classroom. A greater proportion (66.2%) agreed that they had sufficient access to technical support.

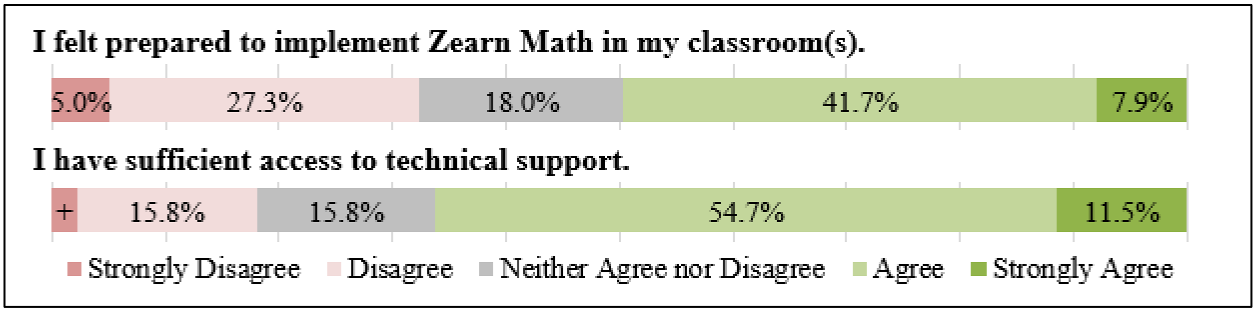

Figure 1. Teachers' questionnaire responses regarding preparation and support for implementation.
Note: + < 5.0%

In focus groups, teachers described professional development opportunities provided by Zearn as helpful but minimal. All four administrative teams interviewed also implied minimal up-front professional development and a general sense of unpreparedness to implement among teachers ("Teachers needed more support and information to ease into the new instructional approach."). In focus groups, teachers agreed they would have benefited from more professional development, but they also indicated that learning from each other was realistically the best way of learning new curriculum. Teachers referenced learning walks and informal conversations when prompted to describe how they learned the curriculum, if not through the professional development provided by Zearn. One teacher said, "We learned through trial and error of course, but that's normal. We did observe a classroom and that was really helpful." Administrative teams also qualified their perceptions of teacher preparedness with, "Teachers learn as they go." Both teachers and administrators implied that the Zearn Math implementation was business as usual. As one administrator noted, "We always need earlier PD and more time and more PD. Always."

Implementation fidelity. Administrative teams from case study schools described the expectations set for implementing teachers with varying emphasis on fidelity. Just two of the four case study schools described Zearn Math as the core curriculum in all classrooms, grades K-5, and that the expectation for teachers was to implement Zearn Math with fidelity during all mathematics blocks. Administrative teams from one school intended for the curriculum to be fully implemented with fidelity in grades K-5 but explained that, in general, tension and discord among teachers was one of the school's primary challenges; administrators described their hesitancy to "force the hand" of resisters and holdouts in a context where other organizational concerns were salient. The administrative team from the final school described full implementation in grades 3-5 with some usage of the curriculum by lower elementary teachers. Administrators in this school explained some K-2 classrooms were "using the games and stuff" to supplement core content but that no expectation was set for full implementation with fidelity in the lower grades.

In focus groups, teachers were asked to describe current Zearn Math implementation in their classrooms, including the degree to which implementation reflected practices prescribed by Zearn. Based on teachers' description of implementation, we infer that most teachers have implemented Zearn Math with moderate to high fidelity. The phrase "flexible implementation" captured teachers' most common description of a mostly high-fidelity implementation with minor adaptations of curriculum routines and features. For example, one teacher described occasionally doing whole group instruction when much of the class "isn't getting it." Other teachers described retaining some center-based instruction and activities from previous curricula or using various online teacher resources as supplemental to Zearn Math.

Descriptions of high-fidelity implementation was the second most frequent description, followed by descriptions of low fidelity. Teachers in high-fidelity classrooms were described as employing the half-class rotational model, with 30 minutes rotations of *Small Group Lessons* led by the primary teacher and *Independent Digital Lessons*. High-fidelity classroom teachers perform notebook checks, require *Exit Tickets*, check *Teacher Reports*, and have a visual representation of lesson completion displayed in their classroom. The majority of teachers in low-fidelity classrooms exclude or modify the use of printed materials. Some teachers did not use *Exit Tickets* or regularly monitor their completion, and other teachers described inconsistent workbook checks. One teacher said, "I don't always know what they aren't getting because I can't see their whole book because that's way too much to go over, their notes and the *Exit Tickets*, but I'm gauging what I see on their small group work."

In the teacher questionnaire, teachers using Zearn Math indicated what instructional approaches they employed during mathematics time each day. The majority of kindergarten teachers (71.5%) reported spending at least half of classroom time on small group instruction, while just over half (57.1%) spent at least half of the classroom time on whole group instruction. The majority of teachers in grades 1-5, though, reported spending at least half of the classroom time on small group instruction, then half on *Independent Digital Lessons*, with a minority of teachers indicating using at least half of class time for whole group instruction.

In terms of barriers to successful implementation, technical aspects appeared to be the most frequent issue for teachers (see Figure 2). Specifically, just under two thirds of teachers indicated that lack of headphones was a frequent or major barrier to implementation. Half of teachers indicated that a lack of devices was a frequent or major barrier. Lack of paper materials was viewed as not a barrier (63.1.%), and not enough time in the mathematics block was sometimes a barrier (40.0%).

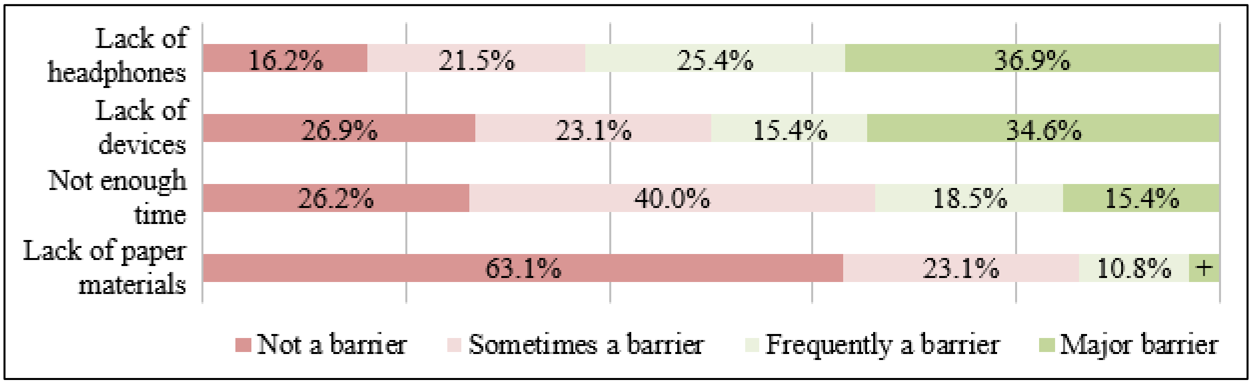

Figure 2. Degree to which teachers viewed factors as a barrier to implementation.
Note: + < 5.0%

Student usage. Tables 6-11 display descriptive information related to student usage. Student usage data were analyzed separately for kindergarten and grade 1-5 students because the curriculum is designed to be used differently by kindergarten students. Kindergarten students work through lessons much more quickly than students in grades 1-5. On average, kindergarten students spent roughly 236 total minutes and just under 10 total weeks active in the digital component, and completed an average of 46 total lessons. Kindergarten students averaged 20 minutes of usage and 4.5 lessons completed per week.

Table 6
Average usage of Zearn Math by kindergarten students (N = 517)

	Total Minutes	Total Weeks	Total Lessons	Minutes per Week	Lessons per Week	*N*
All	236.61	9.79	46.23	19.97	4.59	517

Average usage for all treatment schools, and descriptive usage information for each individual school, are displayed in Table 7. Students in grades 1-5 (*N* = 3,803) used Zearn Math, on average, almost 2,200 total minutes, 28 total weeks, and completed 53 total lessons. Students in grades 1-5 averaged 72 minutes per week, which is 48 minutes below the intended usage of 120 minutes per week, and 1.8 lessons per week, which is below the intended 4 lessons per week. An examination of the range of school-level averages sheds light on usage patterns across all treatment schools, which is quite high in variability. The school-level range of total minutes using the digital component is 1,376 minutes to 3,034 minutes; total weeks using the digital component ranges from 20 weeks to 34 weeks; the range of average total lessons completed is 36 to 77. Average minutes per week ranged from roughly 58 minutes to just over 89 minutes, and average lessons completed per week ranged from 1.39 to 2.26.

Table 7
School-level average usage of Zearn Math by students in grades 1-5

School	Total Minutes	Total Weeks	Total Lessons	Minutes per Week	Lessons per Week	*N*
School A	2529.09	29.44	68.27	84.98	2.26	209
School B	1633.01	24.83	41.19	59.32	1.60	156
School C	2526.27	27.44	54.31	89.48	2.08	210
School D	1375.72	20.58	36.05	57.56	1.52	355
School E	2167.08	27.54	47.55	74.45	1.69	238
School F	1743.54	26.46	37.76	61.54	1.39	226
School G	2383.72	32.24	61.88	73.14	1.90	216
School H	2148.61	28.65	48.64	70.50	1.67	204
School I	2104.77	27.95	56.37	70.11	2.02	589
School J	3034.21	34.55	77.06	87.20	2.23	287
School K	2334.88	24.65	49.21	77.81	1.72	235
School L	2427.71	29.76	57.09	74.70	1.85	293
School M	2135.93	24.97	51.04	65.61	1.87	223
School N	2017.56	29.72	48.64	64.00	1.55	213
School O	2848.42	30.33	62.67	83.07	2.06	149
All	2199.91	27.81	53.31	72.36	1.84	3,803

Table 8 displays usage information by grade level. Across treatment schools, students in the upper elementary grades used the digital component, on average, more than lower elementary grades. For grades 1 and 2, average total minutes using the digital component was 1,433.41 minutes and 1,448.53 minutes, respectively. For grades 3-5, the lowest average total minutes using the digital component was over 1,000 more minutes. The average total weeks active in the digital component for grades 3-5 was at least 29 weeks, and average minutes per week was at least 77. Grade 1 and 2, however, averaged only 21 and 24 weeks of total usage respectively, and 55 and 53 minutes per week, respectively. Interestingly, first grade students in the treatment group averaged higher lessons completed per week than all other grades. This grade group, however, completed the lowest proportion of on-grade lessons (65.2%) as compared with grade 2 (81.2%), grade 3 (86.1%), grade 4 (82.9%), and grade 5 (85.5%).

Table 8
Grade-level average usage of Zearn Math by students in grades 1-5

Grade	Total Minutes	Total Weeks	Total Lessons	Total On-grade Lessons	Minutes per Week	Lessons per Week	*N*
1	1433.41	21.41	63.15	41.19	55.56	2.73	620
2	1448.53	24.01	39.48	32.05	53.44	1.49	718
3	2474.41	31.05	57.28	49.32	77.59	1.79	875
4	2753.25	30.79	55.84	46.30	85.37	1.73	822
5	2616.14	29.63	51.05	43.66	83.72	1.61	768
All	2199.91	27.81	53.31	42.94	72.36	1.84	3,803

Table 9 displays usage data by grade-level within each treatment school. We observed important differences in usage *within schools.* Using School A as an example, students in grade 1 completed an average of 20.56 lessons total, which is over twice what students in grade 4 (at the same school) averaged. For School O, first grade students averaged just 1 total lesson completed while third grade students had one of the highest average total lessons completed in the whole treatment sample (23.81). These observable within-school differences of usage by grade shed further light on the variation in implementation that emerged from other data sources during site visits to case study schools.

Table 9
Average total lessons completed by grade within schools (treatment group only)

	Average Total Lessons Completed				
School	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
School A	92.27	52.28	63.42	33.07	47.17
School B	17.91	4.36	39.05	58.03	25.65
School C	54.91	30.25	42.58	61.41	36.17
School D	1.17	30.92	53.71	34.38	13.53
School E	67.35	42.76	25.79	35.54	21.02
School F	22.71	34.49	38.05	37.35	37.43
School G	102.41	29.26	48.37	37.94	33.96
School H	31.65	62.89	37.41	39	44.54
School I	38.21	22.28	51.63	45.36	40.88
School J	66.33	62.09	57.31	84.19	43.43
School K	3.35	0	57.15	43.69	57.1
School L	18.89	37.98	47.41	70.76	61.65
School M	1.19	0.51	51.97	38.37	77.63
School N	29.12	33.56	42.21	65.1	54.02
School O	0	20.74	82.68	27.68	77.47
All	41.19	32.05	49.32	46.30	43.66

Table 10 displays usage data related to average total remedial lessons completed by school and by grade across the treatment sample. On average, students completed just under 19 total remedial lessons. Across the entire sample, first grade students completed more remedial lessons than all other grades. Similar to other usage patterns, we observed quite a bit of variation between schools, ranging from 8.38 total remedial lessons to 36.49.

Table 10
Average total remedial lessons completed by school and grade (treatment group only)

	Average Total Remedial Lessons Completed	*N*
School A	13.58	209
School B	31.59	156
School C	17.87	210
School D	13.26	355
School E	19.67	238
School F	8.38	226
School G	19.62	216
School H	12.31	204
School I	24.38	589
School J	19.44	287
School K	21.76	235
School L	18.37	293
School M	36.49	223
School N	9.07	213
School O	16.51	149
All grade 1 students	31.04	620
All grade 2 students	20.29	718
All grade 3 students	15.38	875
All grade 4 students	15.68	822
All grade 5 students	15.91	768
All	18.99	3,803

Table 11 displays usage data by prior achievement subgroups. For the NWEA MAP sample, which included students in grades 1-5, students in the low-prior achievement subgroup used Zearn Math more, on average, than all other subgroups in the sample. Students in the low-prior achievement subgroup completed more total lessons and lessons per week, on average, than students in the mid- and high-prior achievement subgroups. For the NWEA MAP sample, mid-prior achievers used the digital component the least. For the State Assessment sample, which included students in grades 4 and 5, students in the high-prior achievement subgroup used Zearn Math more, on average, than all other subgroups in the sample. Students in the high-prior achievement subgroup completed more total lessons and lessons per week, on average, than students in the mid- and low-prior achievement subgroups.

Table 11
Average usage of Zearn Math by students in prior achievement subgroups

Subgroup	Total Minutes	Total Weeks	Total Lessons	Minutes per Week	Lessons per Week	*N*
NWEA MAP Low	2530.75	30.54	81.52	79.01	2.65	309
NWEA MAP Mid	2082.20	26.90	44.00	69.50	1.56	2,320
NWEA Map High	2345.42	28.88	64.29	76.26	2.17	1,174
State Low	2643.77	29.80	46.64	83.92	1.47	955
State Mid	2728.83	30.74	58.96	85.45	1.85	326
State High	2824.69	31.51	77.29	86.43	2.37	212

Table Notes:

(a) The NWEA MAP sample includes students in grades 1-5; the State Assessment sample includes students in grades 4 and 5.

Teacher usage. Based on analysis of usage data, the average teacher minutes per active week for the entire district (*n* = 323) was 47 minutes, which includes time spent reviewing *Teacher Reports*. The lowest average teacher usage occurred in kindergarten (19 minutes). The highest averages occurred in first grade (51 minutes) and fifth grade (58 minutes). Third and fourth grade teachers gradually increased in usage from 36 to 49 minutes. Teachers also conveyed their degree of use and reactions to various Zearn Math components through questionnaire items.

Student-facing components. As shown in Figure 3, teachers were most likely to report using *Independent Digital Lessons* (79.6%), student notes and workbooks (86.2%), small-group lessons (82.4%), and paper *Exit Tickets* (75.4%) frequently (weekly) or very frequently (daily). Less often did teachers report at least frequent use of the printed homework (27.7% never used), printed assessments (26.2%), printed problem sets (20.8%), and *Mission Overviews* (19.0%). Teachers most often indicated that they were not aware of the particular feature when prompted to select the reason why these features were never used. Teachers were able to select "Other" as a response option for questions related to why they were not using certain features of Zearn. When describing "other" reasons for not using printed homework, problem sets, and assessments, teachers most frequently said that they used other sources for materials, including from non-descript online sources and making their own. Several teachers described using EngageNY materials or materials "from the module" as homework. We infer this to mean that teachers are using the whole-group fluency warm-ups and word problems to build homework and/or assessments. It should be noted that some Zearn materials, including printed homework and printed problem sets, are not required elements of the curriculum.

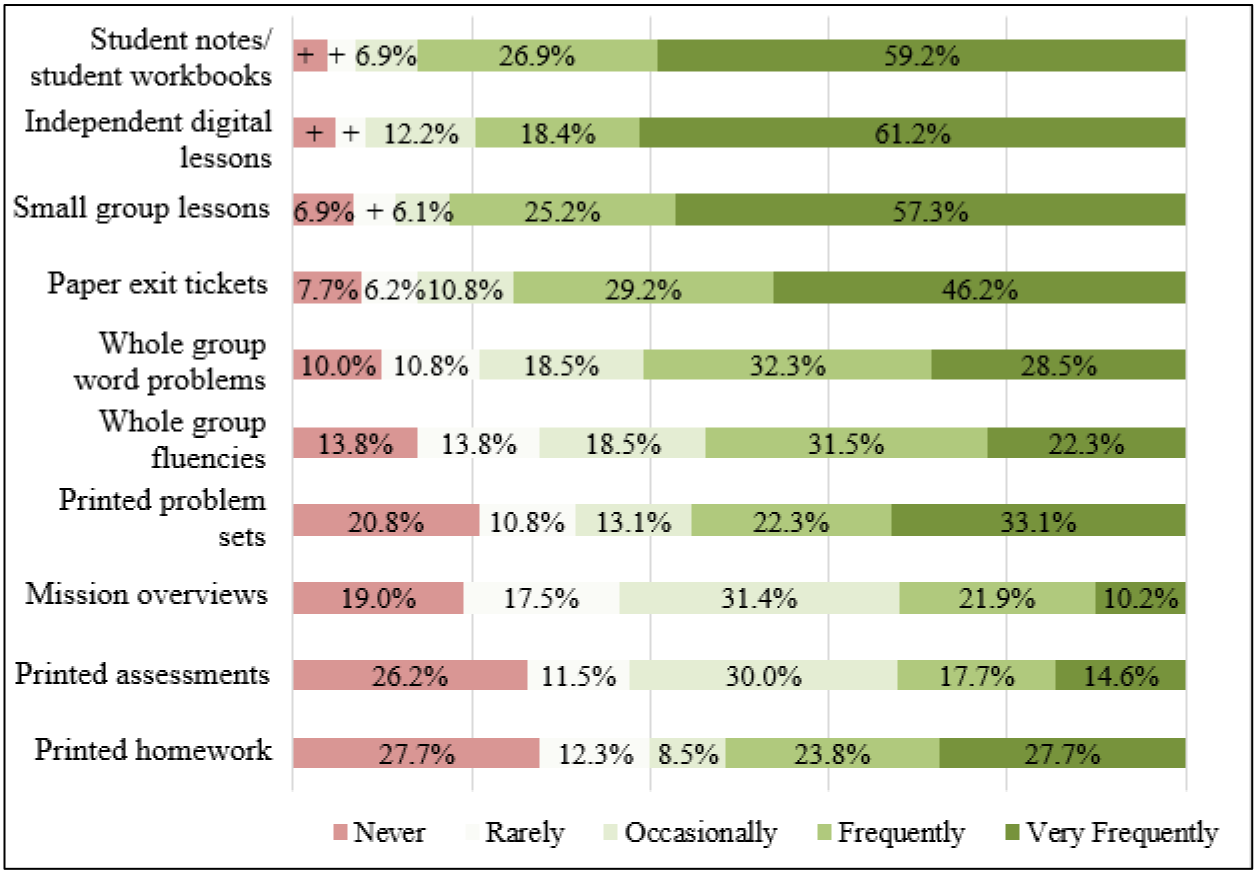

Figure 3. *Reported frequency of student-facing components on the teacher questionnaire.*
Note: + < 5.0%

The majority of teachers that used curriculum features rated them as at least somewhat effective. As displayed in Figure 4, teachers had highly positive perceptions of each feature, particularly *Small-Group Lessons* (92.2% at least somewhat effective), paper *Exit Tickets* (85.1%), *Independent Digital Lessons* (84.4%), and paper student notes or student workbooks (84.0%).

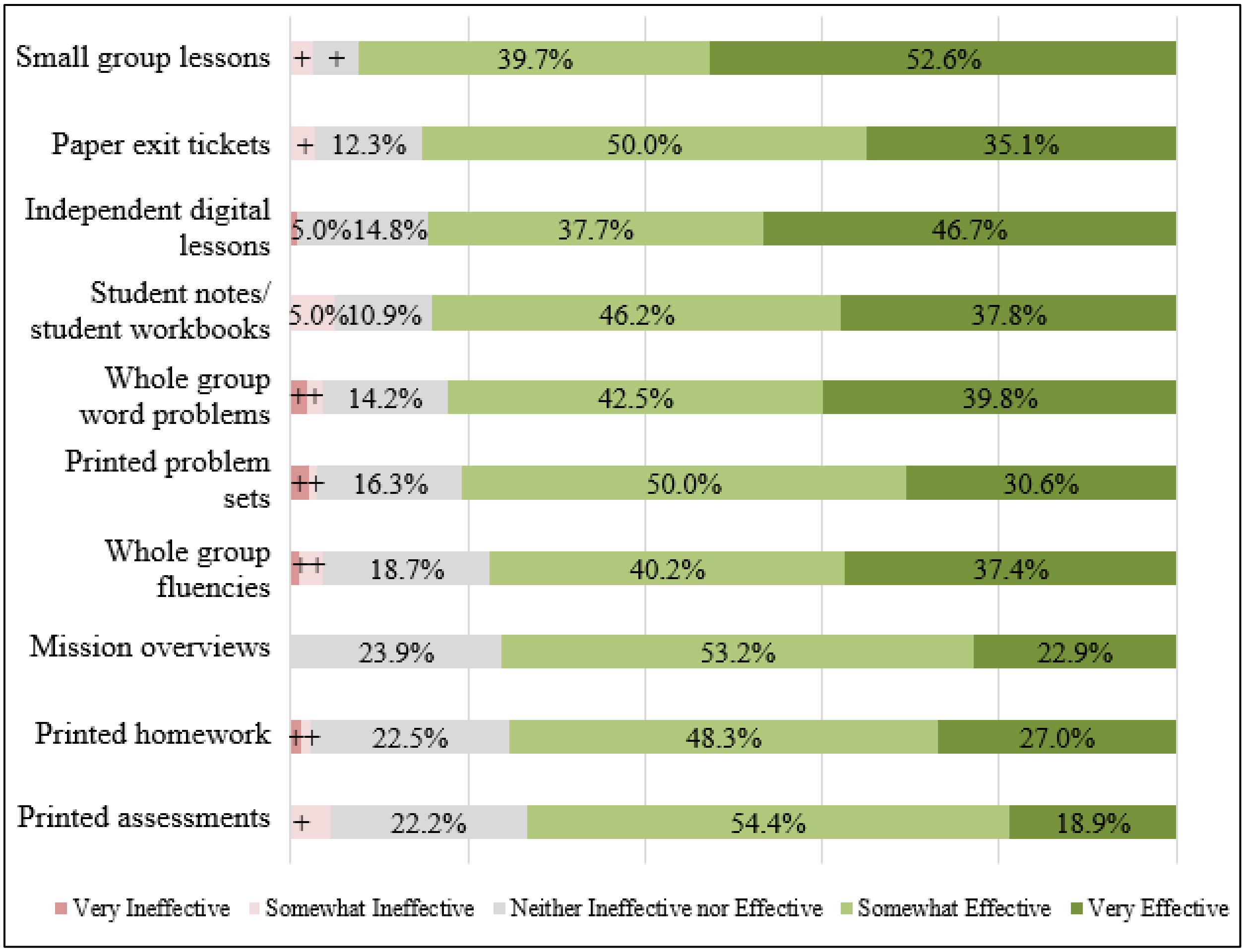

Figure 4. Teachers' perceptions of student-facing components.
Note: + < 5.0%

Teacher-facing components. Questionnaire results indicated that teachers generally used the teacher components less frequently than the student-facing components (see Figure 5). The most frequently used teacher-facing components were the *Teacher Reports*, including the *Pace Report* (74.0% at least frequently), followed by the *Tower Alerts Report* (57.5%) and *Progress Report* (52.2%).

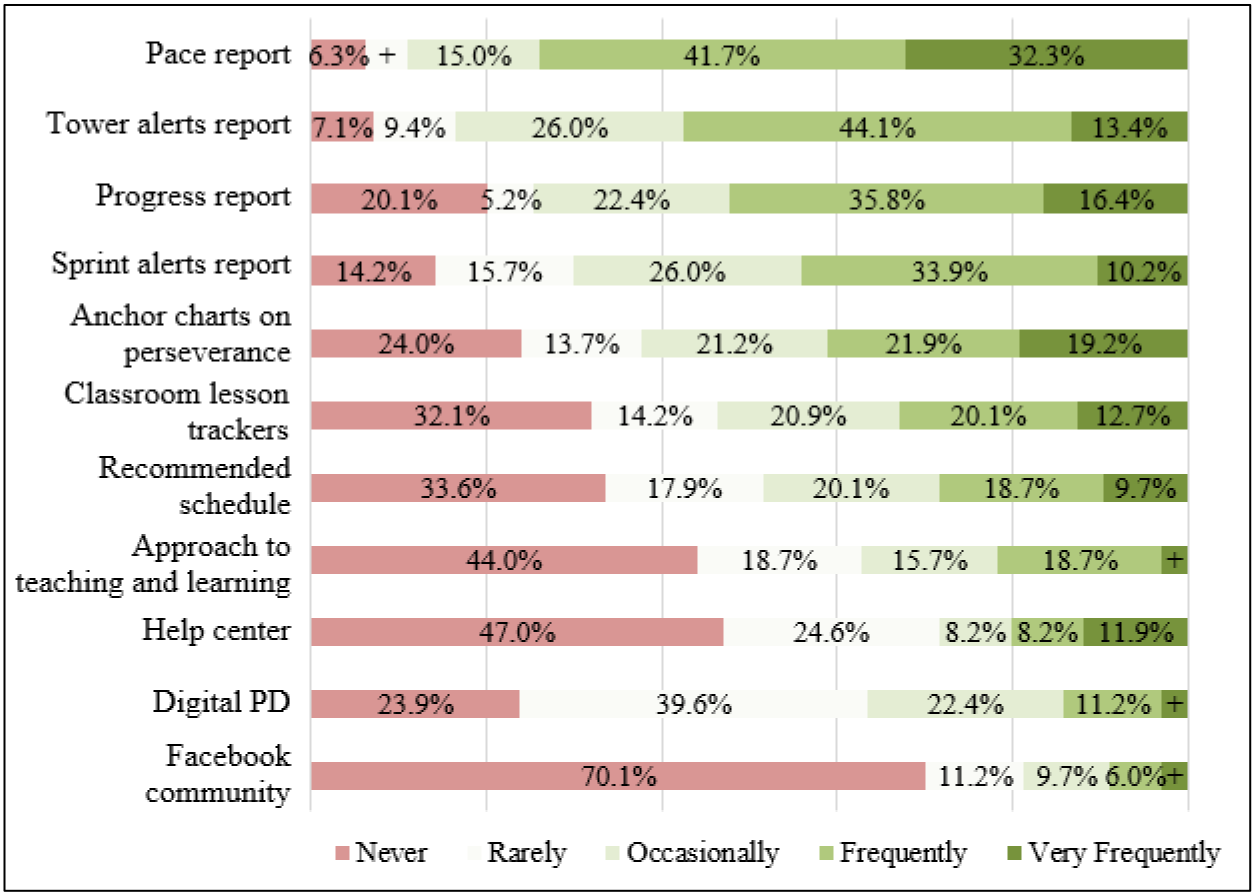

Figure 5. Reported frequency of use for teacher-facing components on the teacher questionnaire.
Note: + < 5.0%

The majority of teachers reported never accessing the Facebook community (70.1%) and never or rarely accessing the *Help Center* (47.0%), though their primary reasons included not having a Facebook account and not needing help. Other less frequently accessed features included Zearn's *Teaching and Learning Guide* (44.0%), the recommended schedule (33.6%), and classroom lesson trackers (32.1%). The majority of participants who indicated never using these features also indicated that the reason was that they were not aware the feature was available. Teachers also described (under "Other" reasons) using their own established methods to track and record student progress. Many viewed the recommended scheduling and pacing from Zearn as not feasible for their students or inconsistent with their classroom norms (e.g., "student-led pace" or "I teach special education and it's not valuable for my students").

As with student-facing components, the majority of teachers rated the teacher-facing components they employed as at least somewhat effective. As displayed in Figure 6, teachers had highly positive perceptions on each of the curriculum features, particularly the *Teacher Reports* (*Pace Report*: 84.6% at least somewhat effective, *Tower Alerts Report*: 79.5%, *Progress Report*: 74.0%), classroom lesson trackers (76.4%), and anchor charts on perseverance/growth mindset (74.5%).

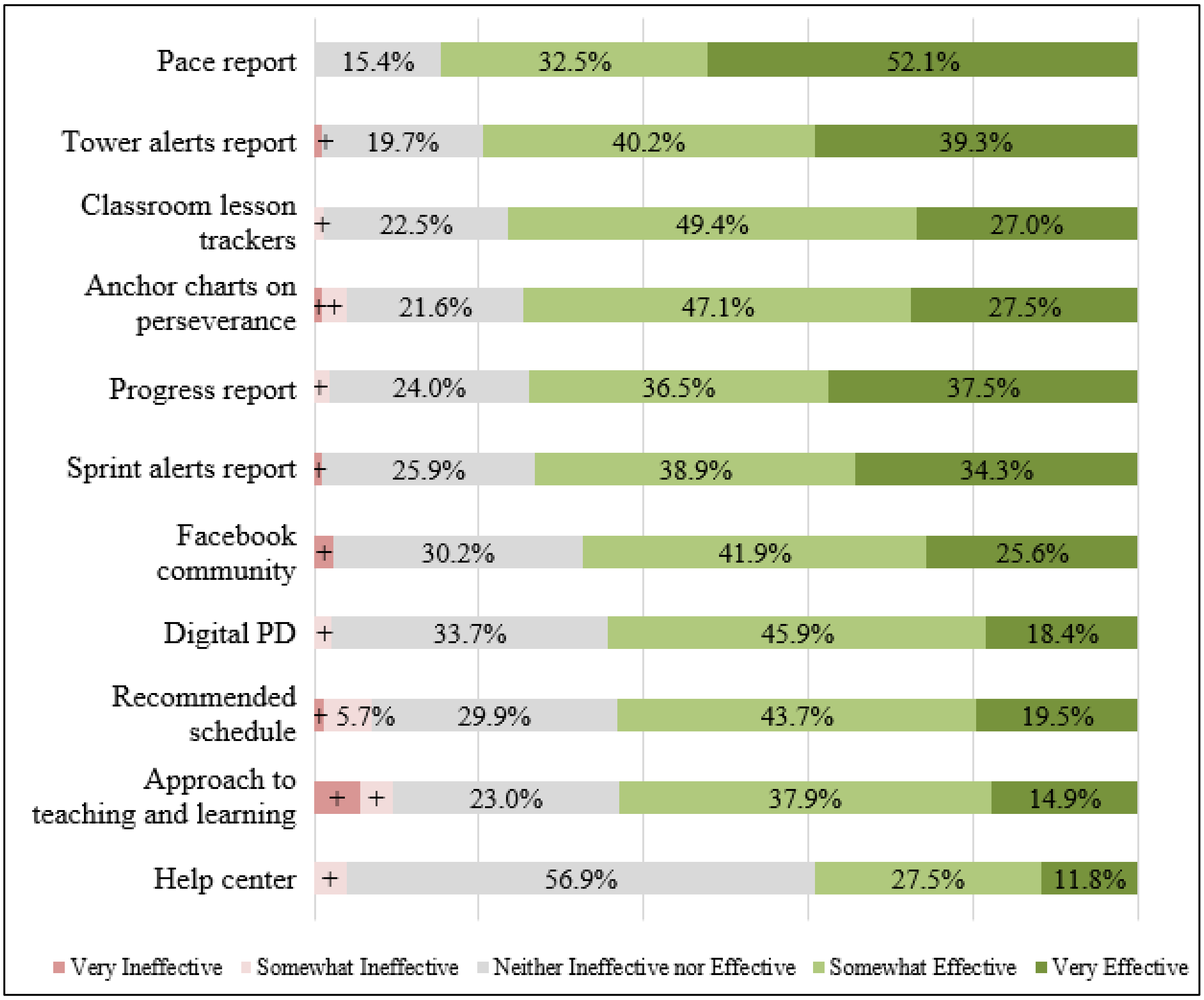

Figure 6. Teachers' perceptions of teacher-facing components.
Note: + < 5.0%

Summary. Onsite visits with case study schools revealed evidence of varied implementation practices and emphasis on fidelity. Not all case study schools had fully implemented Zearn Math in all classrooms, in all grades. In focus groups, teachers most frequently described a modified implementation where they generally followed the classroom structure provided by Zearn Math (i.e., the half-class rotational model) with some adaptations to routines and features. Teachers indicated on the questionnaire that they are using the student-facing features more consistently than teacher-facing tools. Among the student-facing tools, *Digital Lessons*, workbooks, and *Small Group Lessons* are used most frequently; *Mission Overviews* and printed materials (homework, problem sets, and assessments) are used least frequently. Notably, over 70% of teachers rated all of the student-facing tools as being at least somewhat effective. Teachers also generally agreed that teacher-facing components are at least somewhat effective. It appears that teachers generally think highly of curriculum features even if they are not incorporating all features into their daily routines.

Administrators and teachers indicated that up-front professional development was helpful but limited. Multiple data sources suggest that some teachers were likely under-prepared to implement the curriculum with fidelity and stand to benefit from more preparation up front. The most concrete data point related to this finding emerged from the teacher questionnaire: Just under half of teachers reported feeling prepared to implement Zearn Math. Further underscoring the limitations of the professional development provided, for both student-facing and teacher-facing curriculum features, teachers were most likely to never use a curriculum feature because they were not aware the feature existed. In focus groups, teachers expressed a desire for more professional development and implied that learning from each other (e.g., learning walks, time for collaboration) is not only most effective but also preferred. We discuss these findings in more detail in the last section of this report.

Student usage data suggest that students in the district generally did not meet expectations for weekly duration or lesson completion in Zearn Math. Data from the teacher questionnaire provide evidence that external factors, such as access to headphones or individual devices, are a major barrier to meeting usage and completion expectations for students in grades 1-5. We discuss student usage in more detail in later sections of this report. Also included in later sections are additional data sources, which provide more context for why student usage in the district is generally below expectations.

Student Outcomes: Attitudes, Engagement and Achievement

This section includes findings from the student questionnaire related to students' self-reported interest in mathematics education and mathematics self-efficacy. Findings in this section also include stakeholder perceptions about the impact of Zearn Math on students' mathematics achievement and engagement in mathematics education, and results from the analysis of students' achievement on grade level NWEA MAP mathematics assessments.

Student interest and self-efficacy. Students in both treatment and comparison schools responded to questionnaire items regarding their interest in mathematics. As shown in Figure 7, the vast majority of students in both groups indicated agreement that they like to learn new things related to mathematics (treatment: 86.5% at least agreed, comparison: 90.9%), want to get good grades in mathematics (treatment: 95.7%, comparison: 96%), and are interested in mathematics (treatment: 85.9%, comparison: 89.3%). For both interest in mathematics and learning new things related to mathematics, the comparison group was significantly more likely than the treatment group to agree, $p < .01$, though the differences were small.

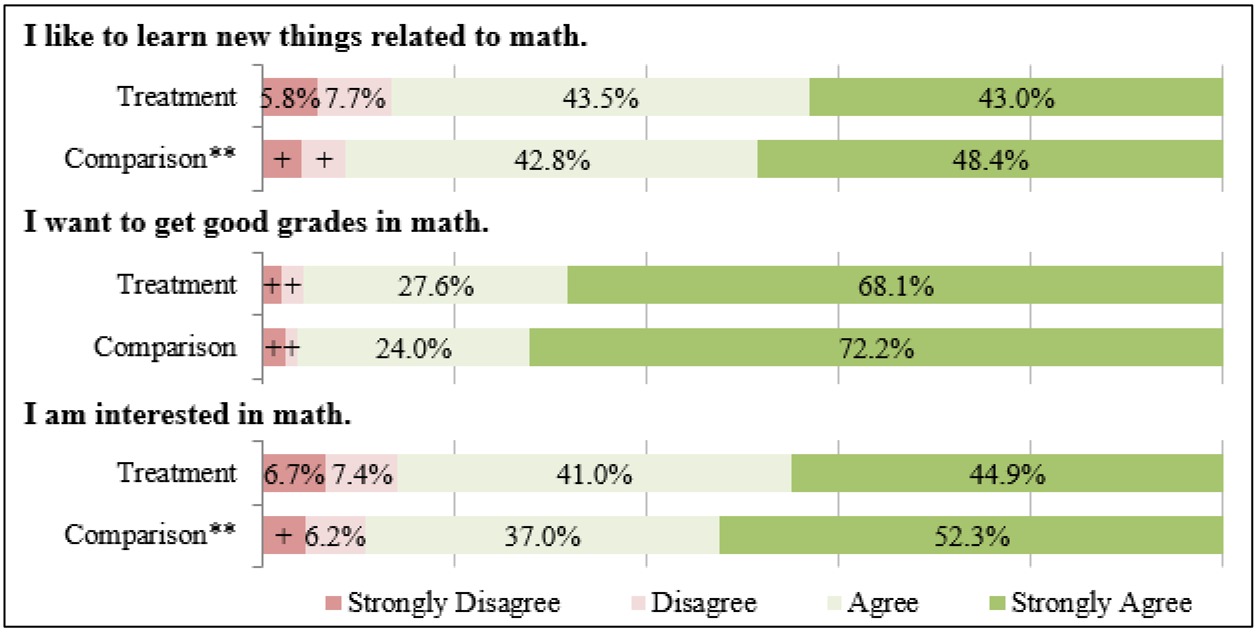

Figure 7. Students' questionnaire responses regarding interest in mathematics.
Note: + < 5.0%, ** $p < .01$

Students in both treatment and comparison schools responded to questionnaire items regarding their self-efficacy toward mathematics. As shown in Figure 8, most students in both groups of schools agreed that they are confident during mathematics tests (treatment: 79.9% at least agreed, comparison: 83%). Though the vast majority of students agreed that they could learn mathematics even when the work was hard (treatment: 91.1% at least agreed, comparison: 93.2%), students in comparison schools were significantly more likely to agree when compared with treatment students, $p < .001$. Additionally, though the vast majority of students also agreed that they could complete all their work in mathematics class if they were persistent (treatment: 88.4% at least agreed, comparison: 91.6%), students in comparison schools were significantly more likely to agree as compared with treatment students, $p < .01$.

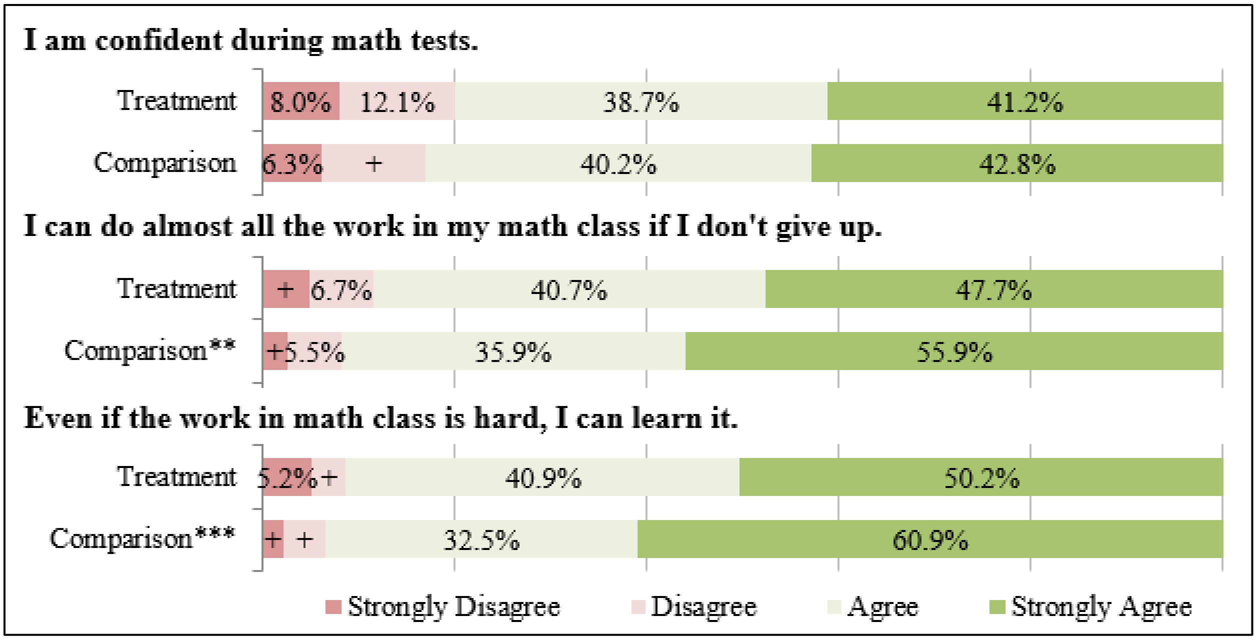

Figure 8. Students' questionnaire responses regarding mathematics self-efficacy.
Note: + < 5.0%, ***p* < .01, ****p* < .001

Teacher and administrator perceptions. At the time of focus group and interview data collection, teachers and administrators most often indicated that it was too early in the school year *and* the implementation process to speak definitively about student achievement (e.g., according to one teacher, "Until I see results then I really can't say."). Teachers generally expressed optimism, however, and emphasized the positive reactions from students to the curriculum. Additionally, while none of the four administrative teams made any definitive statements about student achievement, all expressed general satisfaction with the curriculum and with the effort of their teachers to implement the curriculum.

Teachers responded to questionnaire items regarding the perceived impact of Zearn Math on students in terms of engagement and learning (see Figure 9). The majority of teachers agreed that Zearn Math engages students in mathematics education (83.6% at least agreed). Fewer teachers, though still the majority, agreed that the digital lessons and small-group lessons promote higher-order thinking skills such as critical thinking or problem solving (70.3% and 74.3% respectively). Just under two-thirds of teachers agreed that the curriculum is effective for increasing student achievement over and above regular practices. A relatively large portion of teachers (30.2%) indicated a neutral position regarding increase student achievement, which echoes teachers' responses to related prompts during focus groups.

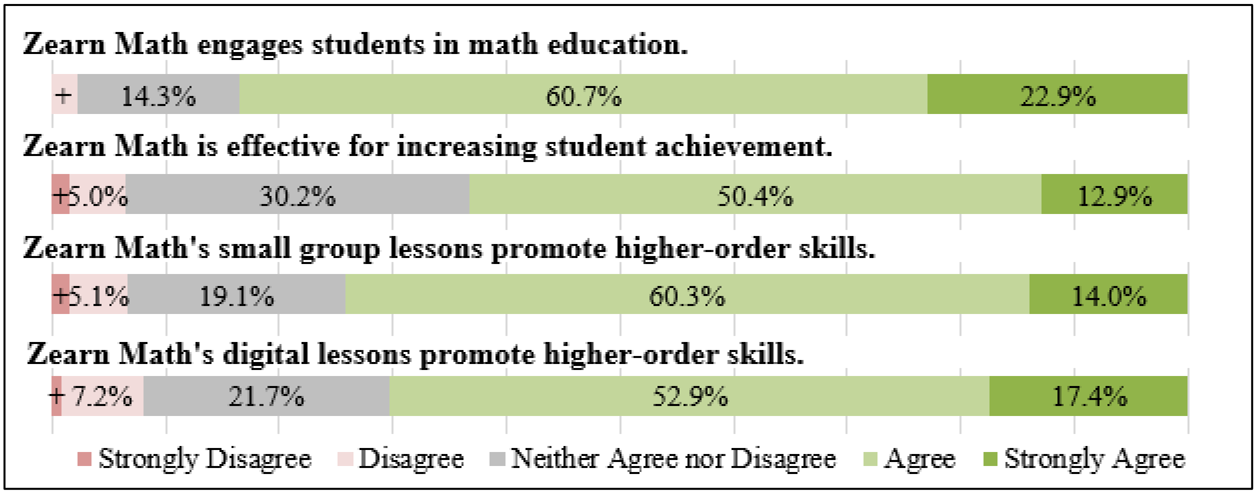

Figure 9. Teachers' questionnaire responses regarding the impact on student learning.
Note: + < 5.0%

Focus groups provided more context for teachers' perceptions of Zearn Math's impact on students. Regarding students' enjoyment of mathematics, teachers describe students as engaged, excited and motivated by the competition embedded in Zearn Math. One teacher said, "Kids LOVE Zearn, they love the competition and the gaming. The leveling up board is huge. Gotta have that in the classroom. The kids light up." Teachers described their students' excitement after completing lessons and emphasized that using a computer during mathematics instruction keeps student engaged in content that is often one of the most difficult subjects during the school day.

Administrative teams also consistently perceived students as being engaged in the curriculum and particularly excited to use technology during learning ("Obviously, kids love to get on the computer to use technology."). Notable here are administrators' perceptions that students remain engaged regardless of skill level. One administrator said, "The biggest thing for me that even the kids who are normally struggling are still trying. They are still on that computer."

Student achievement. The school district employs two standardized mathematics assessments: the NWEA MAP and the State Assessment. The current analysis provides evidence for the following primary research questions using data from both of these mathematics assessments:

- Does the adoption of Zearn Math relate to improved student performance on standardized mathematics assessments?
- How do usage levels of Zearn Math relate to improved student performance on standardized mathematics assessments?

To address the first question, we employed two-level hierarchical linear modeling to compare improvement in mathematics achievement on both exams for treatment and comparison students while controlling for differences in baseline achievement in mathematics and reading, race/ethnicity, gender, special education status, English language learner (ELL) status, grade level, and school effects. Additional analyses were used to estimate the extent to which digital component usage related to improvement in mathematics achievement. Our presentation of results is organized by these two research questions and their respective analytic approaches. The overall treatment effects are presented first, and findings for specific student subgroups follow.

Overall treatment effect. After controlling for the variables specified above, a statistically significant difference in improvement in mathematics achievement was not found between treatment and comparison students. Table 12 displays regression coefficients related to the analysis of NWEA MAP data. The regression-adjusted NWEA MAP mathematics scores were 187.88 for comparison students and 188.37 for treatment students. This difference reflected an effect size of +0.024, which slightly but not significantly favored the treatment group.

Table 12
NWEA MAP mathematics achievement for treatment and comparison students in grades 1-5

	Estimate	Standard error	P-value
Zearn Math	0.490	0.540	0.360
Constant	187.880***	0.350	0.000
Student N	8,718		
School N	35		

Table Notes:
(a) ***p<.001
(b) The statistical model controlled for baseline achievement in both mathematics and reading, race/ethnicity, gender, special education status, ELL status, grade level, and fixed and random school effects

Table 13 displays regression coefficients related to the overall treatment effect according to State Assessment data. Again, state mathematics assessment scores were standardized to accommodate the difference in score scales between 2016-2017 and 2017-2018. Similar to results above, a statistically significant difference in improvement in mathematics achievement on the State Assessment was not found between treatment and comparison students, controlling for other variables. The regression-adjusted assessment scores were 0.032 for comparison students and -0.021 for treatment students. This difference reflected an effect size of -0.05 standard deviations. Again, these results are not statistically significant.

Table 13
State assessment achievement for treatment and comparison students in grades 4 and 5

	Estimate	Standard error	P-value
Zearn Math	-0.053	0.052	0.304
Constant	0.032	0.035	0.355
Student N	3,218		
School N	34		

Table Notes:
(a) The statistical model controlled for baseline achievement in mathematics, race/ethnicity, gender, special education status, ELL status, grade level, and fixed and random school effects.

To summarize, no significant findings emerged from analyses related to overall treatment effects. We conclude that no treatment effect was observed in either of the analytic samples in the current analysis.

Usage analyses. Analyses were also conducted to determine the extent to which time spent using Zearn Math related to improvement in mathematics achievement, relative to comparison students, controlling for student prior achievement and characteristics. Using the NWEA MAP sample in grades 1–5, results for usage variables were positive and statistically significant (see Table 14). On average, more time spent using Zearn Math—in terms of total minutes, total lessons completed, total weeks, and average minutes and lessons per week—was associated with achievement gains on the NWEA MAP mathematics exam.

- Total minutes using Zearn Math: On average, each additional minute using the digital component was associated with an increase in mathematics MAP score of 0.0006 points for participating students relative to comparison students ($p < .001$); each additional hour using the digital component was associated with an increase of 0.0375 points; each additional 5 hours was associated with an increase of 0.1873 points.
    - Students would need to spend roughly 26.5 additional total hours using the digital component in order to increase their mathematics MAP score by 1 point.
- Total lessons completed: On average, each additional Zearn Math lesson completed was associated with an increase in mathematics MAP score of 0.045 points for participating students relative to comparison students ($p < .001$)
    - Students would need to complete approximately 22 additional lessons in order to increase their MAP score by 1 point.
- Total weeks using Zearn Math: On average, each additional week using the digital component was associated with an increase in mathematics MAP score of 0.041 points for participating students relative to comparison students ($p < .01$)
    - Students would need to use Zearn Math for 24 additional weeks in order to increase their MAP score by 1 point.

Table 14
NWEA MAP mathematics improvement and usage for treatment students relative to comparison students

	Estimate	Standard error	P-value
Total minutes using Zearn Math	0.0006***	0.000	0.000
Constant	187.49***	0.303	0.000
Total lessons completed	0.045***	0.004	0.000
Constant	187.06***	0.330	0.000
Total weeks using Zearn Math	0.041**	0.014	0.004
Constant	187.59***	0.324	0.000
Average minutes per week	0.022 ***	0.004	0.000
Constant	187.39***	0.324	0.000
Average lessons completed per week	1.241***	0.121	0.000
Constant	187.88***	0.350	0.000
N	8,718		

Table Notes:
(a) These models were estimated for students in grades 1-5 only.
(b) $^{**}p<.01$, $^{***}p<.001$
(c) All regression models control for variables previously indicated.

In the state assessment sample in grades 4 and 5 only, results for all but one usage variable were positive and statistically significant (see Table 15). The magnitude of the regression estimates described below may appear much smaller than those found in the NWEA MAP analysis because this analysis uses standardized scores instead of scale scores. On average, more time spent using Zearn Math was associated with the following achievement gains on the state mathematics assessment.

- Total minutes using Zearn Math: On average, each additional minute using the digital component was associated with an increase in state mathematics assessment score of 0.0001 standardized scale points for participating students relative to comparison students ($p < .001$); each additional hour using the digital component was associated with an increase of 0.006 standardized scale points; each additional 5 hours was associated with an increase of 0.03 standardized scale points.
  - Students would need to spend roughly 167 additional total hours using the digital component in order to increase their state mathematics score by 1 standardized scale point.
- Total lessons completed: On average, each additional Zearn Math lesson completed was associated with an increase in state mathematics assessment score of 0.004 standardized scale points for participating students relative to

comparison students (p < .001).

- Students would need to complete approximately 250 additional lessons in order to increase their assessment score by 1 standardized scale point.

We found no association between the number of weeks using Zearn Math and improved math performance on the state mathematics test.

Table 15

State assessment improvement and usage for treatment students relative to comparison students

	Estimate	Standard error	P-value
Total minutes using Zearn Math	0.0001***	0.000	0.000
Constant	-0.055	0.034	0.108
Total lessons completed	0.004***	0.000	0.000
Constant	-0.084*	0.036	0.019
Total weeks using Zearn Math	0.002	0.002	0.277
Constant	-0.015	0.035	0.672
Average minutes per week	0.002 ***	0.000	0.000
Constant	-0.061	0.035	0.087
Average lessons completed per week	0.134***	0.015	0.000
Constant	-0.091*	0.037	0.014
N	3,218		

Table Notes:

(a) These models were estimated for students in grades 4 and 5 only.
(b) *p<.05, ***p<.001
(c) All regression models control for variables previously indicated.

In addition to examining the relationship between usage and improvement in mathematics, we also examined the relationship between the number of remedial lessons and improvement in mathematics scores. Remedial lessons were defined as the number of lessons completed below one's grade level divided by total lessons completed. In contrast to the positive relationship between total lessons completed and improvement in mathematics scores, the percentage of remedial lessons completed was negatively associated with improvement on the NWEA MAP mathematics assessment ($p < .001$; see Table 16) and the state mathematics assessment ($p < .01$; see Table 17). These findings indicate, however, that a student completing remedial lessons would still positively benefit from completing lessons, as long as some lessons were not remedial.

Table 16
NWEA MAP mathematics improvement and usage for treatment students relative to comparison students in grades 1-5, controlling for proportion of remedial lessons

	Estimate	Standard error	P-value
Total lessons completed	0.047***	0.004	0.000
Percent remedial lessons	-0.048***	0.005	0.000
Constant	187.41***		
N	8,718		

Table Notes:
(a) ***$p<.001$
(b) All regression models control for variables previously indicated.

Table 17
State mathematics assessment improvement and usage for treatment students relative to comparison students in grade 4 and 5, controlling for proportion of remedial lessons

	Estimate	Standard error	P-value
Total lessons completed	0.004***	0.000	0.000
Percent remedial lessons	-0.002**	0.001	0.004
Constant	-0.074*		
N	3,218		

Table Notes:
(a) *$p<.05$, **$p<.01$, ***$p<.001$
(b) All regression models control for variables previously indicated.

Subgroup analysis. A number of analyses were completed to understand if program effects varied for specific student subgroups. Some subgroup analyses are displayed in Tables 18 and 19. Complete regression tables related to all subgroup analyses of both analytic samples are found in Appendix I. For the NWEA MAP sample, no statistically significant subgroup findings emerged for students of different race, gender, or ELL status, with one exception discussed below (see Table 18). Statistically significant findings did emerge from prior achievement subgroups and grade-level subgroups.

Table 18
NWEA MAP mathematics improvement: Selected subgroup analyses

	Estimate	P-value
Prior Achievement		
Low Prior Achievement	1.112	0.062
Mid Prior Achievement	0.103	0.852
High Prior Achievement	1.325	0.104
Grade		
Grade 1	-0.959	0.163
Grade 2	1.286	0.071
Grade 3	1.571*	0.024
Grade 4	-0.655	0.330
Grade 5	1.405*	0.039

Table Notes:
(a) Results were estimated for students in grades 1-5 only.
(b) $*p<.05$
(c) The treatment effect for each subgroup was calculated by adding the overall treatment effect and the treatment interaction term for the subgroup. The p-values reported in this table show whether the curriculum appeared to have a positive effect for the subgroup relative to similar comparison students.

For prior achievement subgroups analyses of NWEA MAP data, we defined low prior achievement as scoring at or below the 25th percentile and high prior achievement as scoring at or above the 75th percentile, respectively, on the 2017 spring NWEA MAP mathematics assessment. Findings revealed that the low prior achievement subgroup had higher average gains in mathematics relative to their counterparts in the comparison sample subgroup (advantage = 1.112 points), but this difference only approached statistical significance at $p = .06$. There were no observed differences in achievement gains for treatment and comparison students with mid or high prior achievement. These findings suggest that Zearn Math may have particularly positive effects for students in the lowest 25th percentiles of achievement. Exploratory analyses of usage also revealed that low-prior achievement students averaged the most minutes using Zearn, followed by high- and then mid-achieving students. Therefore, the positive program effect for students with low prior achievement may be partially explained by greater program usage for these students. These findings should be interpreted cautiously, however, given the marginally significant effect for low-prior achievers.

Subgroup analyses of NWEA MAP assessment data also revealed positive program effects for students in the third and fifth grades. Treatment students outperformed their comparison peers by an average of 1.57 points ($p < .05$) in grade 3, and 1.41 points ($p < .05$) in grade 5. These findings are partially explained by grade-level differences in Zearn Math usage. Students in grades 3–5 used Zearn Math to a greater extent, on average, than students in grades 1–2. However, while grades 3 and

5 had significant program effects and higher average usage, this pattern did not fit grade 4, which had non-significant program effects and high usage. Grade 2 had lower usage and the program effect was not statistically significant at $p = .07$. No statistically significant subgroup findings emerged for students of different gender or ELL status using the State Assessment sample.

Statistically significant findings also emerged from students' prior achievement subgroups and grade-level subgroups using the State Assessment sample (see Table 19). Students' prior achievement was defined in terms of performance levels (Level 1, 2, 3, or 4). Students in the prior achievement levels 3 and 4 were combined given the small overall proportion of students in these categories. For students with high prior math achievement (e.g., those who scored at a level 3 or 4 on the prior year's math assessment), treatment students outperformed comparison students by 0.151 standard deviations (or standardized units) on the state mathematics assessment, and this was statistically significant ($p < .05$). This finding should be interpreted with caution, however, given that treatment students had substantially higher achievement at baseline than comparison students. Despite this limitation, this finding is corroborated by the directionally (but not statistically significantly) positive program effect identified for high-achieving treatment students on the NWEA MAP mathematics assessment.

Table 19
State mathematics assessment improvement: Treatment effects for selected subgroups

	Estimate	P-value
Prior Achievement		
Low Prior Achievement	-0.102	0.058
Mid Prior Achievement	-0.013	0.839
High Prior Achievement	0.151*	0.036
Grade		
Grade 4	0.026	0.642
Grade 5	-0.137*	0.016

Table Notes:
(a) Results were estimated for students in grades 4 and 5 only.
(b) $*p<.05$, $**p<.01$, $***p<.001$
(c) The treatment effect for each subgroup was calculated by adding the overall treatment effect and the treatment interaction term for the subgroup. The p-values reported in this table show whether the curriculum appeared to have a positive effect for the subgroup relative to similar comparison students.

For students with low prior math achievement, there was a negative program effect (-0.102 standardized points), that approached statistical significance ($p = .06$). This finding is also deemed inconclusive, considering the opposite finding that emerged from the NWEA MAP sample. In the NWEA MAP sample, the low prior achievement subgroup had higher average gains in mathematics relative to the low prior achievement subgroup in the comparison group (p =.06), yet the samples also differed on the basis of grade levels included. Treatment students with high prior achievement outperformed similar comparison peers by 0.151 standardized points ($p < .05$); this finding should be interpreted with caution, given the unacceptably large differences between treatment and comparison students in baseline achievement for this subgroup. Similar to findings from NWEA MAP assessment data, there was no difference in math improvement for treatment and comparison students who were mid-achieving at baseline.

Using the State Assessment sample, treatment students in grade 4 had similar learning gains as their grade-level comparison counterparts, and there was no statistically significant difference. Treatment students in grade 5 under-performed their grade-level comparison counterparts by -0.137 standardized points, on average ($p < .05$). It is important to note, however, that treatment students in grade 5 used Zearn Math less than treatment students in grade 4, in terms of total minutes and average lessons completed. Thus, it is possible that these subgroup findings reflect differences in usage by grade level. Taken together with the NWEA MAP subgroup analysis of grade level difference, these findings indicate that there may be grade-level program effects, but these effects were not entirely consistent across different assessments and were dependent on usage.

To conclude our subgroup analysis, we explored differential treatment effects across different schools. Our exploratory analyses indicated there were few cases in which individual treatment schools appeared to diverge from the overall non-significant program effect. A positive treatment effect was found in one treatment school using the NWEA MAP sample, and negative effects were found in two treatment schools using the State Assessment sample. Differences in usage (in terms of total minutes using Zearn Math) did not account for these differential effects. Moreover, relationships between the treatment indicator and school-level characteristics (e.g., demographics, prior achievement) were also examined, and no patterns were identified. Therefore, these school-level findings related to student achievement suggest an observation of outliers as opposed to larger patterns in which some schools reaped greater benefit of Zearn Math implementation than others.

Summary. Students in all schools, treatment and comparison, demonstrated positive attitudes toward mathematics education and a relatively high degree of mathematics-related self-efficacy. Significant differences, however, emerged between treatment and comparison schools related to attitudes toward mathematics in general. Students from comparison schools were significantly more likely to be interested in mathematics and enjoy learning new things related mathematics. Students from comparison schools were also significantly more likely to feel like they can learn mathematics even when it is hard and that, with persistence, they can complete all of their mathematics work. We discuss these findings in more detail in a later section of this report.

Teachers and administrators shared perceptions of increased student enjoyment of and engagement in mathematics education due to Zearn Math. Administrators in particular were impressed by the degree to which students with varying abilities remained engaged with content in digital lessons. The majority of teachers agree that both major components of Zearn Math—the *Small Group Lessons* and *Individual Digital Lessons*—promote higher order thinking skills but were less enthusiastic about the impact of Zearn Math on student achievement. In focus groups, teachers tended to deflect conversations about achievement, focusing instead on increased engagement. Administrators also expressed cautious optimism when discussing Zearn Math's impact on student achievement. At the time of data collection (March 2018), most explained it is simply too early to tell.

Our analysis of student achievement data on the NWEA MAP mathematics assessment and the state mathematics assessment indicated that no overall difference was observed between improvement in achievement of treatment and comparison schools. All digital component usage variables were positively and significantly associated with improvement in achievement of treatment students relative to comparison students on the NWEA MAP assessment; all but one digital component usage variable (total weeks using the digital component) were positively and significantly associated with improvement in achievement of treatment students relative to comparison students on the state mathematics assessment. We also found some marginally significant and statistically significant findings from subgroup analyses related to prior achievement and grade level. These findings were not consistent or replicated across samples and are difficult to interpret with confidence considering the confounding influence of usage patterns and baseline achievement not being met in all prior achievement subgroups. No statistically significant or conclusive findings emerged from subgroup analysis related to race, gender, or ELL status of students in either of the analytic samples. We discuss implications of these findings in a later section of this report.

Teacher Outcomes: Attitudes and Instructional Practice

Administrators and teachers were prompted to describe the impact of Zearn Math on teachers and teacher practices. Findings in this section include stakeholder perceptions of and questionnaire responses related to the impact of Zearn Math on teachers' instructional practice and mathematics teaching efficacy.

Mathematics teaching efficacy. Teachers in both treatment and comparison schools responded to questionnaire items regarding their mathematics teaching self-efficacy. The reader should exercise caution in interpreting these results, however, due to a low response rate by comparison group teachers. As shown in Figure 10, the majority of teachers in both groups of schools agreed that they know what to do to increase student interest in mathematics and are confident they can both teach mathematics effectively and answer students' questions pertaining to mathematics. They tended to agree that they are continually improving their mathematics teaching practices. There were, however, statistically significant differences in two of the questionnaire items. First, teachers in treatment schools were significantly more likely to agree that they know what to do to increase student interest in mathematics as compared with teachers in comparison schools (treatment: 83.3% at least agreed, comparison: 72.2%), $p < .05$. Second, treatment teachers were also significantly more likely to agree that they are confident they can answer students' mathematics questions (treatment: 98.6%, comparison: 85.2%), $p < .05$.

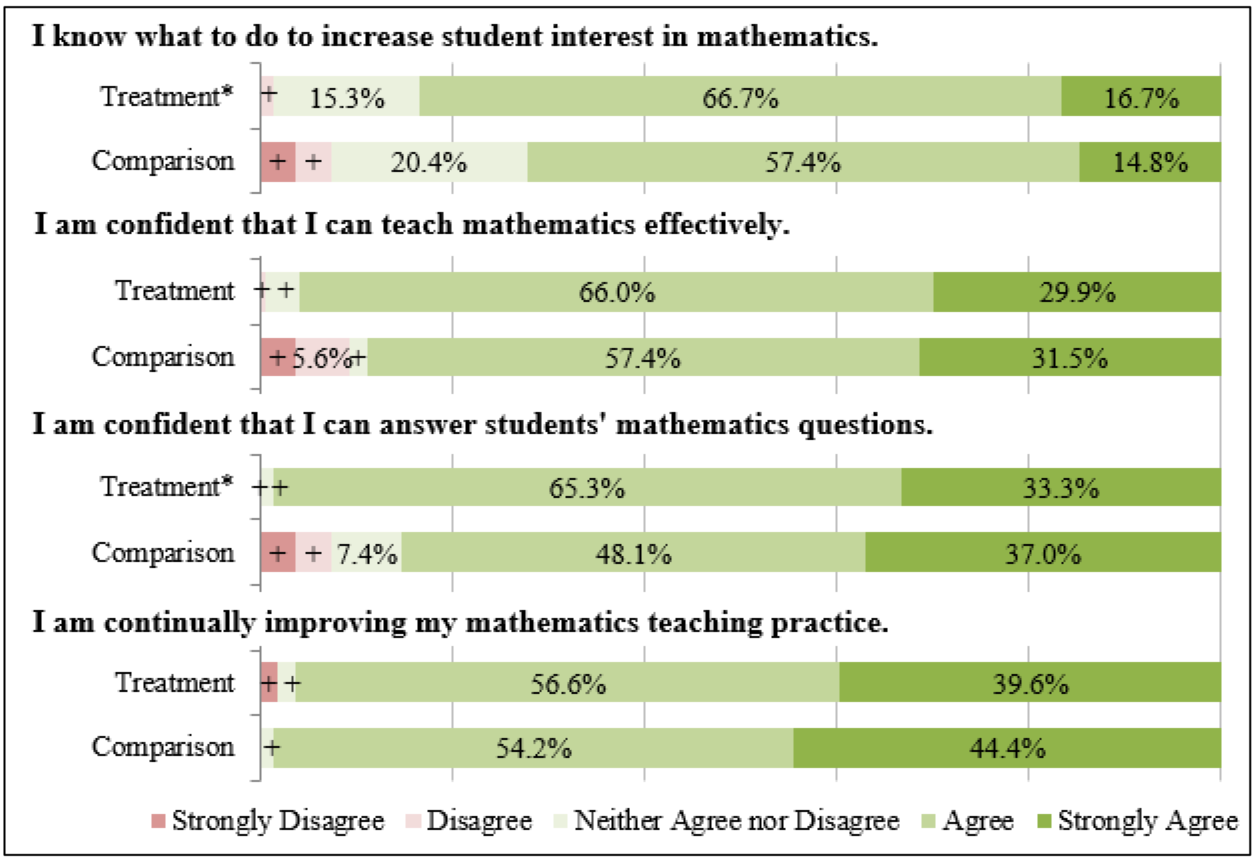

Figure 10. Teachers' questionnaire responses regarding mathematics teaching efficacy.
Note: + < 5.0%

Instructional practice. In the teacher questionnaire, the majority of teachers (79.7%) indicated agreement in the teacher questionnaire that Zearn Math allows them to differentiate instruction based on the needs of their students. Focus group data shed light on the specific features teachers appreciate the most during instructional planning, including the tools that facilitate differentiation. Teachers were consistently enthusiastic about the use of *Teacher Reports* in lesson planning and classroom management. Teachers described using data from *Teacher Reports* to group students, determine how to use "flex" days, and keep up with student progress. Teachers explained that this type of user-friendly information has never been available from prior curriculums and that the impact on their teaching practice is palpable. Further data from the teacher questionnaire affirmed teachers' descriptions of the centrality of *Teacher Reports* to teachers' positive experience with Zearn Math (see Figures 5 and 6 in a previous section of this report). Teachers indicated that *Teacher Reports* were the most frequently used teacher-facing features of Zearn Math and rated them among the most effective features for increasing student achievement.

Administrative teams were also prompted to describe how the curriculum has impacted teacher practices. They most frequently mentioned aspects of teacher practice that extend from the half-class rotational model of Zearn Math. Administrative teams were unanimously complimentary of this classroom format even as they acknowledged that letting students work independently has involved a major adjustment for teachers. Administrators emphasized how the format has facilitated a shift in teacher practice to cultivate independent learners and built in application of knowledge into each mathematics block. According to one principal:

> *In the past we saw a lot of great teaching but a lot of hand holding with kids. Like, there would be a multi-step problem and students would answer the first part, then the teacher would say, 'Ok what now?' And they'd say, 'Ok, I have another step to do and I have another step to do.' They may not have been giving answers to kids, but they were holding their hand. Now we're asking teachers to work with kids but then send them on their own and not help them at all and let's see how they do.*

Summary. The teacher questionnaire revealed generally high level of mathematics teaching efficacy among teachers although there were some significant differences between the treatment and comparison groups. Teachers from the treatment group were significantly more likely to agree that they know what to do to increase student interest in mathematics and that they are confident they can answer students' mathematics questions. We discuss these findings in further detail later in the report.

Multiple data sources suggest that teacher practices were positively impacted by Zearn Math. Teacher described incorporating *Teacher Reports* into their routines, and implied that doing so has improved their instructional practices. Specifically, teachers indicated that *Teacher Reports* are central to instructional planning and differentiating instruction based on student needs. In focus groups, teachers described using *Teacher Reports* to make data-driven decisions with ease and with real-time, accurate information. Administrators described a shift in teacher practice, extending from the half-class rotational format of Zearn Math, toward independent learning and application of knowledge during each mathematics block.

Zearn Math Perceptions

All stakeholders were asked to provide overall impressions of the curriculum, including their initial response to the curriculum and overall strengths and weaknesses. In this section, we report on teacher and administrators' perceived strengths and weaknesses of Zearn Math and on teachers' general reactions to the curriculum. This section also includes student questionnaire data related to students' general impressions of the curriculum and their favorite and least favorite components of individual digital lessons.

Administration perceptions. Administrative teams were asked to provide a description of teacher reactions to Zearn Math. They described generally positive responses to the curriculum from teachers. Administrators specifically mentioned their teachers' familiarity with the curriculum content and the availability of student usage data to explain teacher satisfaction. Their most frequent feedback, however, was that teachers were initially skeptical of the half-class rotational model and that implementation involved a period of adjustment. One administrator said, "The 50/50 model is a new structure for all teachers. They have to adjust to that. They were used to whole group instruction." One principal said of his teachers, "Some teachers are struggling with 50/50 especially the larger classes, larger than 20 kids. It is hard to monitor them. But there are ways that teachers can adapt." Yet another explained, "Half class has been hard. It's a concern, 'How do I approach this half class'." Administrative teams elaborated on the supports in place to facilitate the adjustment to the half-class rotational model, including establishing weekly meetings with mathematics coaches to share strategies and providing human capital (e.g., coaches, general staff, and AmeriCorps members) to certain classrooms during mathematics blocks.

Administrative teams were asked to describe the strengths and weaknesses of the curriculum. In terms of strengths, they most frequently cited the availability of usage data for teachers and the degree to which usage data helps classroom teachers be better mathematics teachers ("The reports help teachers be better lesson planners. They learn more about what their students need."). Administrative teams also considered the repetition and consistency of constructs throughout the curriculum, including across grade levels, to be a major strength of the curriculum. One principal said, "The repetition and seeing it in more than one way is a big deal. The teacher does the small group and then they see it again in the digital lessons." Finally, administrative teams felt that Zearn Math allows students to work independently from wherever they personally are with mathematics content ("We're able to differentiate much more than we've had in the past."). Administrators appreciated that students of all ability levels have an opportunity to be self-directive in Zearn Math. One administrator explained, "Another appealing piece is that teachers might not have the time to devote individual instruction to every student, but they [students] can use Zearn to be on their own."

The most frequent criticism of Zearn Math by administrative teams is the inability to customize individual digital lessons. One mathematics coach summarized,

> *As a district we spent time digging through the curriculum and finding lessons that aren't at grade level, based on our standards, or that we saw as overkill. As a group of teachers and coaches and administration, we work through all of that, we say like, 'Lesson 13 module 5 can be taken out.' But you can't do that. If we choose to skip a lesson, if they choose to skip teaching that, they can, but kids still have to do it.*

Another coach said, "I want more control over Zearn Stuff and My Stuff. More flexibility to determine what kids can access what." This coach explained that there are students who need targeted interventions, which mathematics coaches design to bring students up to speed by focusing on specific skills and gaps in knowledge. For these exceptional cases, flexibility would allow mathematics coaches and intervention specialists to integrate Zearn Math into individualized plans in a way that is more meaningful. Administrators also frequently expressed concerns about the inability of the curriculum to meet the needs of very low-level students. They expressed frustration on behalf of their teachers that, while Zearn Math is an exceptional curriculum for average students and high-flyers, some students who struggle to engage with the very beginning of grade-level content remained effectively "stuck" in Zearn.

Teacher perceptions. At the conclusion of each focus group, we asked teachers explicitly, "Would you recommend Zearn Math to other teachers." The majority of teachers in all focus groups indicated they would recommend Zearn Math to other teachers. Teachers' questionnaire responses (see Figure 11) also indicated that they had overall very positive perceptions of Zearn Math. Specifically, the vast majority agreed that they would like to use Zearn Math in the future (87.1%) and that they would recommend the curriculum to other teachers (83.5%). Importantly, over one-third of teachers indicated strong agreement to these two questionnaire items.

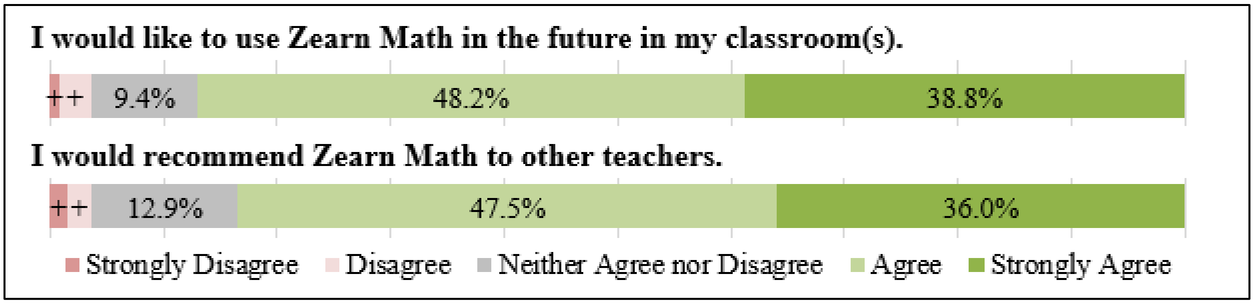

Figure 11. Teachers' questionnaire responses regarding perceptions of Zearn Math.
Note: + < 5.0%

Throughout focus groups, teachers provided specific positive and negative feedback related to the curriculum. As mentioned previously, teachers were particularly complementary of Teacher Reports. Teachers cited the ease of access to the type of information provided by the reports, and the utility of that information for instructional planning. Positive feedback also included that Zearn Math provides reinforcement of content and that *Individual Digital Lessons*, when used correctly, is akin to having an additional teacher in the classroom. Teachers also appreciate the independent learning encouraged by digital lessons. One teacher said, "I like that if the student is on Zearn, not only are they taken care of for 30 minutes, but they are doing their own thing. That's student accountability." Teachers complimented the extent to which Zearn has made teaching mathematics easier. They described the curriculum as streamlined and praised the problem sets provided for small-group lessons. One teacher explained,

> *I'd work so hard with centers and making games, but if they're not playing it right, they're not learning. I like that Zearn is right there teaching what they need to know. I like that it is less work. At the beginning of the year when kids were not on Zearn I'd try to find things for them to work on for 30 minutes. Now Zearn takes away that so I can really concentrate on the small group lesson. I feel like I can focus more on the lesson in my planning.*

This teacher captures a sentiment shared by others: Teachers believe that they are better teachers in small groups, and that small-group teaching is more effective for mathematics especially. Teachers described the consistency of small group interaction as a major strength of the curriculum.

The most frequent concern from teachers was the perception that Zearn Math is particularly unable to meet the needs of all students in their classes and the inability of Zearn Math to facilitate targeted remedial work to address specific gaps in knowledge. This concern was pervasive across all focus groups. One teacher asked, "How do I manage a kid where mission one is over their head?" In general, teachers were quite expressive regarding the current district's requirement to keep students at grade level in *Independent Digital Lessons*. Teachers also are not totally convinced of the utility of moving ahead in *Small Group Lessons* even as students remain "behind schedule" in the digital component.

Teachers also described their discomfort with the half-class rotational model of Zearn Math at length. Many were not immediately comfortable letting students learn on their own and felt additionally burdened by students' inability to manage themselves for that length of time. During the focus groups, some teachers described different strategies to mitigate disruptions from students working on *Independent Digital Lessons*. Several teachers felt that Zearn Math would be a more effective tool if they had a second teacher to work with struggling students and assist with classroom management.

Finally, teachers complained about district issues related to the use of Zearn Math such as the (lack of) availability of headphones, poor Wi-Fi functionality, aging technology, and shortened mathematics block times. These themes echo findings from the teacher questionnaire, where just under two thirds of teachers indicated that lack of headphones was a frequent or major barrier to implementation; 50% indicated a lack of devices and 40% indicated shortened mathematics blocks were frequent or major barriers to implementation.

Student perceptions. Students' questionnaire responses indicated that they overall had positive perceptions regarding Zearn Math (see Figure 12). Specifically, the vast majority of students indicated agreement that Zearn Math is a good program for learning mathematics (89.2%) and that they look forward to using the program (81.3%). Most students also agreed that Zearn Math makes learning mathematics fun (79.4%) and that they would like to use the program in the future (74.4%).

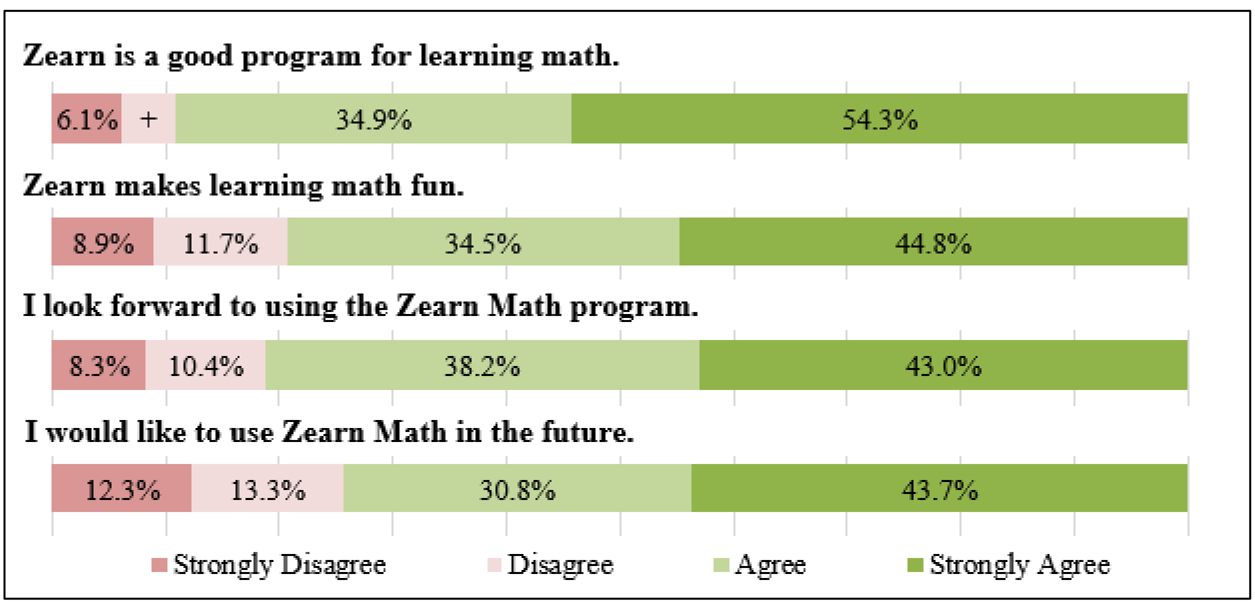

Figure 12. Students' questionnaire responses regarding Zearn Math.
Note: + < 5.0%

Students were prompted at the end of the student questionnaire to describe their favorite and least favorite component of Zearn Math. These prompts were open-ended and optional (*n* = 1,304). The most frequently mentioned favorite component of Zearn Math overall is "learning" (*n* = 229). Responses related to learning included, "I just like it because it lets me learn math and math is my favorite," and "Zearn teaches me new things I didn't know before." The most frequently mentioned favorite specific curriculum feature is the *Tower of Power* (*n* = 225), followed by *Wild Cards* (*n* = 111) and *Math Chat* (*n* = 104). Students also mentioned the fact that Zearn makes mathematics fun (*n* = 107; "You can be a math star while you have fun."). Other relatively frequent mentions were related to feeling a sense of achievement or accomplishment when completing lessons (*n* = 91).

Students were prompted to describe their least favorite component of Zearn Math. Responses such as "nothing," "I like everything about Zearn," "I can't think of anything," represented the largest portion of responses (*n* = 285). The Tower of Power is the most frequently mentioned least favorite component of Zearn Math (*n* = 265). Just over half of all mentions of the *Tower of Power* (*n* = 148) were dual-coded to represent negative emotional states experienced relative to the *Tower of Power*. Students used phrases such as, "when you get one thing wrong in the Tower and they bring you down," or "I dislike the Towers because it's hard to not be stuck," to describe experiences extending from that section of the curriculum. The other (roughly) half of mentions only identified the curriculum feature, with no additional context.

Students also frequently mentioned that Zearn Math is hard (*n* = 180). Students said, "It's very hard for me, I can't complete it," and, "It gets too hard." Students also described frustration and that they need more help in the digital component (*n* = 107), and that they disliked starting over in features or losing their points in the *Tower of Power* (*n* = 95). A full analysis of all open-ended student responses is in Appendix G.

Summary. Data from the teacher questionnaire and focus groups suggest teachers had an overwhelmingly positive attitude toward Zearn Math and would recommend the curriculum to other teachers. Data from focus groups provided more details about what exactly teachers like and dislike about the curriculum. Teachers described their adjustment to the half-class rotational model as uncomfortable, but they *also* consistently cite the positive contribution of the format to student learning. Teachers appreciated that digital lessons provide students with independent work time while *Small Group Lessons* give teachers the opportunity to provide meaningful and effective direct instruction. Teachers expressed concerns about the ability for Zearn Math to meet the needs of struggling students. While teachers may have been able to "float" in previous formats, the half-class rotational model completely removes the core teacher from part of the classroom while they are teaching small group, leaving students without assistance.

Administrators echoed these strengths and weaknesses provided by teachers with near symmetry. Administrators were especially enthusiastic about the differentiation and independent learning facilitated by the half-class rotational model. Administrators also considered the availability of usage data to teachers and the consistently of constructs throughout grade levels as major strengths of the curriculum. Some of the administrative teams we spoke with included mathematics coaches and/or vice-principals who oversaw mathematics and science curriculum at their school. These administrators in particular shared a desire for more flexibility in Zearn Math to meet the needs of students with targeted intervention plans *and* to be integrated more seamlessly with the mission/focus of mathematics education at their particular school. Teachers and administrators also shared concerned related to the availability of functional computers and headphones and a need for more mathematics time in general and more teachers to provide support for core teachers during mathematics blocks. These concerns, while notably frequent, are not addressable by Zearn.

The results of the student questionnaire suggest that students have overall positive feelings toward Zearn Math. Open-ended questions revealed some specific information about what students like and dislike about the curriculum. When naming their favorite component of Zearn Math, students most frequently described that the curriculum helps them learn mathematics or get better at certain math-related skills. Students also named specific curriculum features such as the *Tower of Power*, *Math Chat*, and *Wild Cards*. When discussing their least favorite component, students most frequently said, "nothing, I love Zearn." Students tended to mention negative feelings provoked by the curriculum, especially related to the *Tower of Power* and feeling like Zearn Math is difficult. We discuss these finding more and provide recommendations in the next section of this report.

Discussion

The purpose of the present study was to gather summative and formative data related to the implementation of Zearn Math in a large urban school district. The current report includes findings from all data sources, including classroom observations, focus groups with teachers, student and teacher questionnaires, interviews with school administrators and Zearn Math personnel, student usage data, and student achievement data. In the present section, we discuss main findings and their implications.

Preparation and Implementation

Our findings suggest that implementation of Zearn Math in the district is going well overall and that stakeholders hold largely positive views of the curriculum and their experiences with Zearn. Based on site visits and interviews with Zearn personnel, the curriculum appears to be a good fit for the district. Zearn Math is currently quite popular among teachers in the district, and teachers and administrators appear to be mutually invested in its success. Zearn will likely benefit from maintaining a highly supportive relationship with the district because of what can be learned about its product through an in-depth look at district implementation and use over time.

In terms of implementation, multiple data sources indicate that while a majority of the district teachers implemented Zearn Math with moderate to high fidelity, expectations for implementation were inconsistent across treatment schools. We infer that district administrators expected full implementation in all grades, in all treatment schools, but that school-based administrators were hesitant to make over-arching demands of teachers. One of the reasons Zearn Math stands to be successful in the district is the relatively high degree of organic teacher buy-in. Administrators are likely acutely aware of the importance of teacher buy-in to the success of new curriculums and may rightly continue to allow gradual implementation in their school. It is important for Zearn to support school-level administrators by providing resources and strategies to engage resisters.

Usage data also indicated that students in treatment schools are generally not meeting expectations for weekly duration and average lessons completed. Questionnaire and focus group data suggest teachers are concerned about shortened mathematics blocks and a lack of headphones and devices for their students, as well as students being off-task. Teachers may need more support to establish transition routines to help students maximize the amount of time they are in the curriculum.

We infer from multiple data sources that the overwhelming majority of teachers are excited about Zearn Math but that some district teachers were better prepared to implement than were others. Just under 50% of teachers felt prepared to implement Zearn Math in their classroom; however, teachers and administrators tended to qualify their own perceptions of teacher unpreparedness, describing minimal PD and learning on the job as business-as-usual. Deficits in teacher preparation likely explain why teachers most frequently reported not using certain curriculum features because they were unaware it existed, and may help explain the discomfort teachers described during the adjustment to the half-class rotational model. In interviews and focus groups, teachers, administrators and Zearn personnel consistently qualified descriptions of the professional development provided up front as not ideal, due to a lack of time between when teachers received PD and the start of the school year.

Student Outcomes

Comparison group students demonstrated significantly higher agreement with prompts related to interest in mathematics knowledge (i.e., I am interested in math; I am interested in learning new things related to math) and significantly higher mathematics self-efficacy (i.e., I can learn math even when it is hard; I can complete all my work in math). A potential explanation for these findings is that Zearn Math (treatment) students may have found Zearn Math material more challenging than previously experienced, which may have affected their feelings toward mathematics in general. Students in treatment schools generally did not meet expectations for weekly duration or lesson completion. While external factors in the district, such as too little time during mathematics block or lack of headphones, are cited as barriers to successful implementation, it is also the case that students in the district generally struggle to master grade-level content—just 8% of students were proficient in mathematics in 2017. Student engagement and efficacy are preceded by feelings of mastery (Bandura, 1997; Gecas, 1989; Pekun & Linnenbrink-Garcia, 2012; Salomon, 2008). To the extent that students are unable to complete lessons and achieve goals in Zearn Math, they may form more negative views about mathematics in general. Zearn should further explore students' sense of mastery in the curriculum and consider how to ensure that all students, regardless of ability, have tangible opportunities within the curriculum to feel successful.

When prompted during site visits, teachers and administrative teams were cautiously optimistic about student achievement outcomes. Teachers and administrators, though positive overall, were perhaps rightly hesitant to make predictions before end-of-year testing was completed. However, multiple data sources indicated that teachers and administrators saw definitive increases in student engagement in mathematics in general, and described Zearn Math as widely enjoyed by their students. Both groups described the majority of students was motivated to learn how to use technology and was positively impacted by the competition and "leveling up" embedded in *Individual Digital Lessons*. We infer from site visits that teachers and administrators believe student engagement and enjoyment precedes achievement in a meaningful way. This hypothesis is supported by prior research (Baroody, Rimm-Kaufman, Larsen, & Curby, 2016; Duncan et al., 2007; Hao, Yunhoo, & Wenye, 2018; Roorda, Jox, Zee, Oort, & Kroomen, 2017).

Our analysis of student achievement data did not indicate statistically significant differences between the improvement in achievement of students in the overall treatment and comparison samples. Within the treatment group, nearly all student usage variables, except total remedial lessons completed, were positively and significantly associated with achievement. While the regression model calculated controlled for student-level differences (i.e., demographics and baseline achievement), the current analysis could not rule out the possibility of unobservable differences—such as teacher quality or student motivation—between treatment and comparison students, which may confound the results.

Coupled with data from site visits, student usage data shed light on the school-level effects at play in the district. One of the main findings from site visits was that implementation between and within the four case study schools varied. Only two case study schools were implementing in all grades; the other two schools implemented in some grades and not others or partially implemented in certain classrooms. Descriptive statistics of student usage data from all treatment schools affirm that inconsistent implementation between and within schools was a pervasive component of the current implementation. Within several schools, usage trends were quite high in some grades and low in others. Across all schools, average usage varied widely. The snapshot provided by the case study schools, and our discussion of findings related to implementation, likely serve as a fairly accurate proxy for understanding current student achievement outcomes in the district.

Although there was no overall program effect on increased performance on standardized mathematics assessments, the relationship between digital component usage by students and achievement was positive and significant across multiple usage measures and in both analytic samples. While this represents the most robust finding to emerge from our analysis of student achievement data, it’s important to keep in mind that usage would have to increase dramatically to achieve the predicted improvements. There also were positive program effects for certain student subgroups and grade levels, although the supplementary analyses should be interpreted cautiously given smaller sample sizes, the inconsistent findings between analytic samples, and the risk that multiple analyses might yield isolated chance effects. Grade-level effects that emerged in the NWEA MAP sample were partly explained by usage, where positive program effects were identified in grades with higher average usage (grades 3 and 5). In the State Assessment analysis, grade 5 students under-performed their counterparts in the comparison group, while grade 4 students performed comparable to comparison students. Again, these findings may be partially explained by usage data; Grade 5 students demonstrated lower usage of Zearn Math, on average, than students in grade 4. Still, these findings potentially identify situations where the program was differentially effective during the year, perhaps due to implementation factors that in future years could be addressed adaptively.

Findings from the State Assessment sample related to prior achievement subgroups provide some evidence that providing differentiated (adaptive) instruction through small-group and independent digital lessons can accommodate content and pacing needs of students of student who otherwise might be slowed down (high-prior achievement group) by uniform lessons. Less convincing evidence emerged for the pacing needs of low-achieving students. While marginally significant positive effects were found for low-prior achieving students in the NWEA MAP sample, the finding was not replicated in the State Assessment sample. In fact, the opposite finding was observed, also with marginal significance. More research is needed to confirm that there are differential effects for different pretest subgroups beyond a positive relationship between usage and outcomes.

While lacking conclusive statistical support in this first-year study, the achievement results seem encouraging overall given their positive direction, significant impacts in some supplementary analyses, and perhaps especially, the early implementation efforts examined. In the latter regard, it is noteworthy that digital component usage levels (e.g., lessons completed) in most schools failed to meet expectations. Determining whether these early effects increase as implementation matures over time is important for evaluating program success. The CRRE team is addressing this question in a Year 2 study that will examine achievement outcomes for Zearn Math in the district over the 2018-19 school year.

Teacher Outcomes

Multiple data sources suggest that *Teacher Reports* are among the most celebrated features of Zearn Math. Teachers described using these features as intended and want more training on how to continue to use the reports effectively. Accessibility to data at the student, classroom, and school level that is easy to understand and relevant to the goals of teachers and administrators will likely remain central to teacher and administrator perceptions of technology as additive to curriculum and instruction (Straub, 2009).

Teachers from the treatment group were significantly more likely to agree they know what to do to increase student interest in mathematics and that they are confident they can answer students' mathematics questions than comparison group teachers were. Paired with other findings related to professional development and teacher preparedness, we are cautious to attribute these differences to treatment group teachers' exposure to professional development related to Zearn Math. We are more convinced by findings related to teachers' generally positive feelings toward Zearn Math and the resounding perceptions of increased student engagement in mathematics. Treatment group teachers may express more confidence in knowing what to do to increase student interest in mathematics because they have seen student interest in mathematics increase in their classroom. They may feel more confident about their ability to answer students' mathematics questions because of the history most teachers have with EngageNY content. It is likely that consistently teaching in a small group and the availability of *Teacher Reports* facilitate more confidence about where individual students are with content and the gaps in knowledge that underlie questions.

Curriculum Perceptions

Our findings suggest that stakeholders have, overall, positive perceptions of Zearn Math. Among the most robust findings from the current study is that teachers and administrators want to use Zearn Math in the future and would recommend the curriculum to their peers. Teachers generally perceived student- and teacher-facing curriculum features as effective, with few exceptions. Students also expressed overwhelmingly positive feelings about the curriculum —almost 90% of students agree it is a good curriculum for learning mathematics and roughly 80% of students look forward to using the curriculum each day. These findings should be convincing to district administrators and provide support for maintaining and expanding the use of Zearn Math throughout the district.

The greatest strength of the curriculum, according to teachers and administrators, appears to stem from the half-class rotational model of Zearn Math. Specifically, administrators and teachers both described the format as the foundation of better mathematics teaching and more engaging, personalized learning experiences. Both groups also appreciated that the format allows students to encounter content in multiple ways. Students receive small group instruction—which is preferred among teachers for teaching mathematics—*and* have opportunity to be independent learners with access to differentiated instruction and practice, with little to no additional burden on teachers. Interestingly, teachers and administrators spoke at length about the difficult transition to half-class rotational model while unanimously citing its benefits. Both groups emphasized a need for support during implementation and the importance of establishing classroom routines to mitigate student disruptions during independent work time. In interviews and focus groups, teachers and administrative teams cited the benefit of sharing effective classroom management techniques among teachers, and that the period of adjustment, among teachers and students, was finite. These findings provide evidence for a need for more training and support that anticipates the discomfort teachers may feel during implementation.

Considering all data sources from all stakeholders, the biggest concern about Zearn Math is the perceived inability of the curriculum to meet the needs of all students. To be sure, teachers and administrators are enthusiastic about the ease at which differentiated instruction is available to students and consistently praise the extent to which students are able to work independently from where they personally are with content knowledge. We infer that teachers and administrators are concerned about the handful of students who perform below grade-level, who may or should receive targeted intervention services. Teachers and administrators are emphatic about the extent to which the very format of Zearn Math—the demand for individual learning, the integration of listening, reading, typing and writing—is fundamentally incompatible with the needs of a small, but important, population of students. This a fruitful avenue for Zearn to explore in future iterations of the curriculum. We believe that additional tools and flexibility in the curriculum that allow teachers to integrate Zearn Math into individual intervention plans will mitigate this most pervasive concern about the curriculum.

Recommendations

Based on the findings on this evaluation, we offer the following suggestions for Zearn Math and the district to improve the curriculum and implementation:

Provide district and school administrators with targeted support to lead implementation of Zearn Math.

District and school administrators will benefit from additional professional development and support materials that are geared toward the role of administration during implementation. Professional development for administrators should offer strategies for supporting teachers during the transition to the half-class rotational model, engaging resisters, and should address the importance of fidelity in achieving positive student outcomes.

Provide teachers with resources that are easily accessible and support understanding and usage of features of Zearn Math.

Among the most concerning findings from the current study is the lack of knowledge teachers have about certain Zearn Math features, pervasive perceptions of teacher unpreparedness and the less than favorable feedback related to the professional development teachers did receive prior to implementation. Circumstances may be unique within the district, but Zearn may consider strategies to increase awareness of features and access to as-needed professional development. Zearn may also consider updating features currently perceived as ineffective and centralizing teacher resources that are currently spread across multiple platforms (e.g., YouTube channel, Zearn Facebook page, Zearn website).

Provide teachers and administrators with strategies for supporting students during independent learning.

Multiple data sources suggest students experience both positive and negative emotions extending from the gamified components of Zearn Math and that teachers experience some difficulty with classroom management in the half-class rotational format. Zearn may consider further research into the negative emotions experienced by students, and how to translate negative emotional states into positive learning experiences. Zearn may also consider including specific guidance about classroom routines and teacher practices that help students persevere during *Independent Digital Lessons*.

Consider providing more flexibility for school-based adults to modify digital components of the curriculum to fit individual school and student needs.

We believe that mathematics teachers and coaches will continue to emphasize the needs of a small and important group of students who need targeted skills interventions. Zearn Math may consider how to increase flexibility within the curriculum to fit unique needs while maintaining cohesiveness of lessons and emphasis on implementation fidelity.

References

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman & Company.

Baroody, A. E., Rimm-Kaufman, S. E., Larsen, R. A., & Curby, T. W. (2016). A multi-method approach for describing the contributions of student engagement on fifth grade students' social competence and achievement in mathematics. *Learning and Individual Differences, 48*, 54-60. doi: https://doi.org/10.1016/j.lindif.2016.02.012

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstain, L., Engel, M. Brooks-Gun, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428-1446. Doi:10.1037/0012-1649.43.6.1428

Gecas, V. (1989). The social psychology of self-efficacy. *Annual Review of Sociology, 15,* 291-316.

Glaser, B. and Strauss, A. (1967). *The discovery of grounded theory: strategies for qualitative research.* London: Weidenfeld and Nicolson.

Hao, L., Yunhuo, C., & Wenye, Z. (2018). Relationships between students' engagement and academic achievement: A meta-analysis. *Social Behavior & Personality: An International Journal, 46*(3), 517-528.

Pekrun R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. Christensen, A. Reschly, & C. Wylie (Eds), *Handbook of research on student engagement*. Boston: Springer.

Roorda, D. D., Jak, S., Zee, M., Oort, F. J., & Koomen, H. Y. (2017). Affective teacher-student relationships and students' engagement and achievement: A meta-analytic update and test of the mediating role of engagement. *School Psychology Review, 46*(3), 239-261.

Salomon, G. (1983). The differential investment of mental effort in learning from different sources. *Educational Psychologist, 18*(1), 42-50.

What Works Clearing House. (2017). *Standards Handbook.* Retrieved from: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Appendix A: Classroom Observation Guide

Comment on the following aspects of the classroom:

1. **Environment and Setting:**
   What activities are you observing (e.g. independent work, peer teaching, problem sets) and how are students grouped (e.g. whole group, small group, independently)?
   How are children arranged in the classroom relative to 1) other students, 2) personal technology devices, 3) the teacher and 4) shared technology devices?
   What classroom systems / anchor charts do you observe?

2. **Curriculum and Instruction:**
   How and to what extent technology is integrated into instruction and activities?
   How and to what extent are digital citizenship norms integrated into learning and/or play?

3. **Teacher Activities**
   Describe the instructional practices of the teacher.
   Describe how the teacher is facilitating learning.

4. **Student Activities and Engagement:**
   Describe what students are doing in the classroom.
   How and to what extent do students appear to be engaged in learning through the program?

Appendix B: Teacher Focus Group Protocol

1. How is Zearn implemented in your classrooms?

    a. To what extent/how did you use the Independent Digital Lessons, Small Group and Whole Group Materials?
    b. To what extent and how did you use the Teacher Reports?

2. To what degree are school administrators or others involved in the implementation of Zearn Math?

3. Were you adequately prepared to implement Zearn Math? Why or why not?

4. What, if any, changes have you made to your instructional practices as a result of using Zearn Math?

5. To what extent has Zearn Math impacted student achievement?

6. To what extent has Zearn Math affected student attitudes towards math, such as their mindset and motivation?

7. How does Zearn Math compare to other math curricula?

8. What do you see as the strengths of Zearn Math?

9. What suggestions would you have to improve the program or its implementation in your school?

10. Would you recommend Zearn Math to other educators? Why or why not?

11. What routines and systems did you establish in your classroom to implement Zearn Math?

    - Routines/Systems for digital time and small group time?
    - Routines/Systems to help students persevere through challenge?
    - Routines/Systems for providing feedback on Student Notes and Exit Tickets?
    - Routines/Systems for tracking and celebrating progress?

12. Do you feel Zearn Math meets the needs of most of your students? Why or why not?

Appendix C: Administrator Interview Protocol

Let's begin by learning more about your school and your history there.

- Please briefly describe the school with regard to size, types of students, the community, and student outcomes.
- How long have you been principal there?
- What math curriculum did you use before Zearn Math?

General Topic: Implementation of Zearn Math

- Why did you decide to implement Zearn Math in your school?
- How prevalent is Zearn Math in your school? Are all teachers and all grades using the program?
- What guidelines / expectations did you establish in how classrooms should implement Zearn Math?
- What professional development was in place to support the implementation of Zearn Math?
- Were your teachers adequately prepared to implement Zearn Math?
    - If yes, what went well?
    - If not, what should have gone differently?

General Topic: Impact of Zearn Math on Students and Teachers

- How do you perceive teachers are responding to Zearn Math?
- To what degree has the program had a positive impact on student achievement (e.g. grades, test scores)?
- Do you feel Zearn Math meets the needs of most of your students? Why or why not?
- Do you feel Zearn Math meets the needs of most of your teachers? Why or why not?
- Would you recommend this program to other school administrators? Why or why not?

General Topic: Overall Perceptions

- To what degree do you believe Zearn Math benefits your school overall?
- What do you see as the strengths of Zearn Math?
- What suggestions would you have to improve the existing program?
- Is there anything else you would like to add?

Appendix D: Zearn Math Personnel Interview Protocol

Let's begin by learning more about your role at Zearn and how you were involved with Zearn in the school district.

- Please briefly describe your role at Zearn. How long have you worked for Zearn?
- What is your role in regards to the district?

General Topic: Implementation in the district

- How did the relationship between Zearn and the district begin?
- To what degree are you and other Zearn personnel involved in the day-to-day implementation of Zearn Math in the district?
    - How does your role change over time?
- In your opinion, were the teachers in the district adequately prepared to implement Zearn Math?
    - If yes, what went well?
    - If not, what should have gone differently?
- Is there anything specific Zearn can do to better support teachers and administrators during implementation?

General Topic: Impact of Zearn Math on Students and Teachers

- To what degree have the district teachers responded well to the program?
- In your experience, what factors tend to influence whether or not teachers and students respond well to the program?
- To what degree do you believe Zearn Math is a good fit for the district?

General Topic: Overall Perceptions of Implementation Fidelity

- In your opinion, to what degree have teachers implemented the Zearn Math program effectively?
- What factors generally influence implementation fidelity?
- What factors are at play in the district, in terms of the fidelity of implementation in classrooms there?
- Is there anything Zearn Math should be doing differently to ensure teachers are implementing Zearn effectively?

Appendix E: Teacher Reaction Questionnaire

Rate your level of agreement to the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	N	*M*	*SD*
I am continually improving my mathematics teaching practice								
Treatment	0.0%	0.0%	1.4%	54.2%	44.4%	144	4.43	0.52
Comparison	1.9%	0.0%	1.9%	56.6%	39.6%	53	4.32	0.70
I am confident that I can teach mathematics effectively.								
Treatment	0.0%	0.0%	1.4%	65.3%	33.3%	144	4.25	0.55
Comparison	3.7%	3.7%	7.4%	48.1%	37.0%	54	4.07	0.95
I am confident that I can answer students' mathematics questions.								
Treatment	0.0%	0.7%	3.5%	66.0%	29.9%	144	4.32[a]	0.50
Comparison	3.7%	5.6%	1.9%	57.4%	31.5%	54	4.11	0.96
I know what to do to increase student interest in mathematics								
Treatment	0.0%	1.4%	15.3%	66.7%	16.7%	144	3.99[a]	0.61
Comparison	3.7%	3.7%	20.4%	57.4%	14.8%	54	3.76	0.89

[a]Treatment teachers were significantly more likely to agree than comparison teachers, $p < .05$

Rate your level of agreement to the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	N	*M*	*SD*
I felt prepared to implement Zearn Math in my classroom(s)								
Treatment	5.0%	27.3%	18.0%	41.7%	7.9%	137	3.20	1.09
I have sufficient access to technical support								
Treatment	2.2%	15.8%	15.8%	54.7%	11.5%	137	3.57	0.97
Zearn Math engages students in math education.								
Treatment	0.0%	2.1%	14.3%	60.7%	22.9%	138	4.04	0.68
Zearn Math is effective for increasing student achievement over and above regular practices.								
Treatment	1.4%	5.0%	30.2%	50.4%	12.9%	137	3.68	0.82
Zearn Math's small group lessons promote higher-order skills such as critical thinking or problem solving.								
Treatment	1.5%	5.1%	19.1%	60.3%	14.0%	134	3.81	0.80
Zearn Math's digital lessons promote higher-order skills such as critical thinking or problem solving.								
Treatment	0.7%	7.2%	21.7%	52.9%	17.4%	136	3.79	0.85
Zearn Math allows me to differentiate my instruction based on student needs.								
Treatment	1.4%	6.5%	12.3%	59.4%	20.3%	136	3.90	0.85
I would like to use Zearn Math in the future in my classroom(s).								
Treatment	0.7%	2.9%	9.4%	48.2%	38.8%	137	4.20	0.78

I would recommend Zearn Math to other teachers.								
Treatment	1.4%	2.2%	12.9%	47.5%	36.0%	137	4.13	0.83

Use the scale to indicate how frequently you or your classroom(s) use different features of the Zearn Math program.

	Never Used	Rarely Used (once or twice a semester)	Occasionally Used (once or twice a month)	Frequently Used (weekly)	Very Frequently Used (daily)	N	*M*	*SD*
Mission Overviews	19.0%	17.5%	31.4%	21.9%	10.2%	137	2.87	1.25
Independent Digital Lessons	4.8%	3.4%	12.2%	18.4%	61.2%	147	4.28	1.11
Anchor Charts on perseverance/growth mindset	24.0%	13.7%	21.2%	21.9%	19.2%	146	2.99	1.45
Teacher Report: Progress Report	20.1%	5.2%	22.4%	35.8%	16.4%	134	3.23	1.35
Facebook community	70.1%	11.2%	9.7%	6.0%	3.0%	134	1.60	1.08
Digital Professional Development	23.9%	39.6%	22.4%	11.2%	3.0%	134	2.30	1.05
Classroom Lesson Trackers	32.1%	14.2%	20.9%	20.1%	12.7%	134	2.67	1.43
Recommended Schedule	33.6%	17.9%	20.1%	18.7%	9.7%	134	2.53	1.38
Approach to Teaching and Learning	44.0%	18.7%	15.7%	18.7%	3.0%	134	2.18	1.26
Help Center	47.0%	24.6%	8.2%	8.2%	11.9%	134	2.13	1.40
All but Kindergarten:								
Student Notes or Student Workbooks	3.8%	3.1%	6.9%	26.9%	59.2%	130	4.35	1.01
Paper Exit Tickets	7.7%	6.2%	10.8%	29.2%	46.2%	130	4.00	1.23
Whole Group Fluencies	13.8%	13.8%	18.5%	31.5%	22.3%	130	3.35	1.34
Whole Group Word Problems	10.0%	10.8%	18.5%	32.3%	28.5%	130	3.58	1.28
Small Group Lessons	6.9%	4.6%	6.1%	25.2%	57.3%	131	4.21	1.18
Printed Homework	27.7%	12.3%	8.5%	23.8%	27.7%	130	3.12	1.61
Printed Problem Sets	20.8%	10.8%	13.1%	22.3%	33.1%	130	3.36	1.54
Printed Assessments	26.2%	11.5%	30.0%	17.7%	14.6%	130	2.83	1.38
Teacher Report: Pace Report	6.3%	4.7%	15.0%	41.7%	32.3%	127	3.89	1.11
Teacher Report: Tower Alerts	7.1%	9.4%	26.0%	44.1%	13.4%	127	3.47	1.07
Teacher Report: Sprint Alerts	14.2%	15.7%	26.0%	33.9%	10.2%	127	3.10	1.21

If you indicated you never use a feature, please indicate why this feature was not used.

	Not enough time	Materials or tools are lacking	Not valuable to instruction	Don't know how to use	Wasn't aware of this component	Other:	N
Mission Overviews	18.2%	0.0%	4.5%	13.6%	59.1%	4.5%	22
Independent Digital Lessons	20.0%	0.0%	10.0%	0.0%	70.0%	0.0%	10
Anchor Charts on perseverance/growth mindset	13.3%	0.0%	0.0%	0.0%	76.7%	10.0%	30
Teacher Report: Progress Report	18.2%	1.0%	3.0%	15.2%	29.3%	28.3%	99
Digital Professional Development	15.6%	0.0%	0.0%	12.5%	62.5%	9.4%	32
Classroom Lesson Trackers	11.6%	0.0%	2.3%	27.9%	48.8%	9.3%	43
Approach to Teaching and Learning	11.8%	0.0%	0.0%	15.7%	66.7%	5.9%	51
All but Kindergarten:							
Student Notes or Student Workbooks	20.0%	40.0%	0.0%	20.0%	0.0%	20.0%	5
Paper Exit Tickets	10.0%	0.0%	10.0%	10.0%	10.0%	60.0%	10
Whole Group Fluencies	5.6%	0.0%	5.6%	16.7%	61.1%	11.1%	18
Whole Group Word Problems	7.7%	0.0%	7.7%	15.4%	69.2%	0.0%	13
Small Group Lessons	22.2%	0.0%	0.0%	22.2%	55.6%	0.0%	9
Printed Homework	8.3%	0.0%	2.8%	5.6%	41.7%	41.7%	36
Printed Problem Sets	7.4%	0.0%	0.0%	3.7%	51.9%	37.0%	27
Printed Assessments	8.8%	0.0%	2.9%	2.9%	52.9%	32.4%	34
Teacher Report: Pace Report	0.0%	0.0%	0.0%	37.5%	50.0%	12.5%	8
Teacher Report: Tower Alerts	11.1%	0.0%	0.0%	33.3%	44.4%	11.1%	9
Teacher Report: Sprint Alerts	27.8%	0.0%	11.1%	22.2%	33.3%	5.6%	18

	Not enough time	Materials or tools are lacking	Not valuable to instruction	Don't know how to use	Wasn't aware of this component	Other	Schedule moves too slow/fast	N
Recommended classroom schedule	18.2%	0.0%	4.5%	13.6%	59.1%	4.5%	0.0%	45
	Not enough time	Materials or tools are lacking	Not valuable to instruction	Don't know how to use	Wasn't aware of this component	Other	I never needed help	

Help center	7.2%	0.0%	2.9%	5.8%	18.8%	11.6%	53.6%	69
	Not enough time	Materials or tools are lacking	Not valuable to instruction	Don't know how to use	Wasn't aware of this component	Other	I do not have a Facebook account	
Facebook community	0.0%	0.0%	0.0%	0.0%	28.6%	42.9%	28.6%	7

Use the scale to indicate how effective you consider different features of the Zearn Math program to be for increasing students' knowledge and achievement related to math.

	Very Ineffective	Somewhat Ineffective	Neither Ineffective or Effective	Somewhat Effective	Very Effective	N	*M*	*SD*
Mission Overviews	0.0%	0.0%	23.9%	53.2%	22.9%	109	3.99	0.69
Independent Digital Lessons	0.8%	0.0%	14.8%	37.7%	46.7%	122	4.30	0.78
Anchor charts on perseverance/ growth mindset	1.0%	2.9%	21.6%	47.1%	27.5%	102	3.97	0.84
Teacher Report: Progress Report	0.0%	1.9%	24.0%	36.5%	37.5%	104	4.10	0.83
Facebook community	2.3%	0.0%	30.2%	41.9%	25.6%	43	3.88	0.88
Digital Professional Development	0.0%	2.0%	33.7%	45.9%	18.4%	98	3.81	0.76
Classroom Lesson Trackers	0.0%	1.1%	22.5%	49.4%	27.0%	89	4.02	0.74
Recommended Schedule	1.1%	5.7%	29.9%	43.7%	19.5%	87	3.75	0.88
Approach to Teaching and Learning	4.6%	3.4%	23.0%	37.9%	14.9%	87	4.34	1.86
Printed Homework	1.1%	1.1%	22.5%	48.3%	27.0%	89	3.99	0.81
Printed Problem Sets	2.0%	1.0%	16.3%	50.0%	30.6%	98	4.06	0.84
Printed Assessments	0.0%	4.4%	22.2%	54.4%	18.9%	90	3.88	0.76
Paper Student Notes or Student Workbooks	0.0%	5.0%	10.9%	46.2%	37.8%	119	4.17	0.82

Paper Exit Tickets	0.0%	2.6%	12.3%	50.0%	35.1%	114	4.18	0.74
Whole Group Fluencies	0.9%	2.8%	18.7%	40.2%	37.4%	107	4.10	0.87
Whole Group Word Problems	1.8%	1.8%	14.2%	42.5%	39.8%	113	4.17	0.87
Small Group Lessons	0.0%	2.6%	5.2%	39.7%	52.6%	116	4.42	0.71
Teacher Report: Pace Report	0.0%	0.0%	15.4%	32.5%	52.1%	117	4.37	0.74
Teacher Report: Tower Alerts	0.9%	0.0%	19.7%	40.2%	39.3%	117	4.17	0.80
Teacher Report: Sprint Alerts	0.9%	0.0%	25.9%	38.9%	34.3%	108	4.06	0.83
Help Center	0.0%	3.9%	56.9%	27.5%	11.8%	51	3.47	0.76

How much time per day is spent on math instruction in your classroom?

30 minutes or less	30-60 minutes	60-90 minutes	90-120 minutes	120 minutes or more	N	*M*	*SD*
2.3%	12.3%	70.8%	13.8%	0.8%	130	95	2.98

Of the time spent on math instruction, how much time per day is spent on…

	Very little (<25%)	About a quarter of the time	About half of the time	About three quarters of the time	Almost all of the time (>75%)	N	*M*	*SD*
Whole group instruction	50.8%	35.4%	12.3%	0.8%	0.8%	130	1.65	0.79
Small group instruction	3.8%	16.2%	43.1%	12.3%	24.6%	130	3.38	1.14
Individual digital lessons	8.5%	19.2%	63.8%	3.1%	5.4%	130	2.78	0.86

Please indicate the degree to which you consider the following factors to be a barrier to the successful implementation of Zearn Math in your classroom.

	Not a barrier	Sometimes a barrier	Frequently a barrier	Major barrier	N	*M*	*SD*
Lack of headphones	16.2	21.5	25.4	36.9	130	2.83	1.10
Lack of devices	26.9	23.1	15.4	34.6	130	2.58	1.22
Lack of paper materials	63.1	23.1	10.8	3.1	130	1.54	0.81
Not enough time in math block	26.2	40.0	18.5	15.4	130	2.23	1.01

Appendix F: Student Reaction Questionnaire

Rate your level of agreement to the following statements.

	Strongly disagree	Disagree	Agree	Strongly agree			
	%	%	%	%	N	*M*	*SD*
I am interested in math.							
Treatment	6.7	7.4	41.0	44.9	1185	3.24	0.86
Comparison	4.5	6.2	37.0	52.3	470	3.37[b]	0.79
I want to get good grades in math.							
Treatment	2.1	2.2	27.6	68.1	1282	3.62	0.64
Comparison	2.4	1.4	24.0	72.2	508	3.66	0.63
I like to learn new things related to math.							
Treatment	5.8	7.7	43.5	43.0	1195	3.24	0.83
Comparison	4.2	4.6	42.8	48.4	477	3.35[b]	0.76
I am confident during math tests.							
Treatment	8.0	12.1	38.7	41.2	6.3	3.13	0.92
Comparison	6.3	10.7	40.2	42.8	460	3.20	0.87
I can do almost all the work in my math class if I don't give up.							
Treatment	4.9	6.7	40.7	47.7	1192	3.31	0.80
Comparison	2.7	5.5	35.9	55.9	476	3.45[b]	0.72
Even if the work in math class is hard, I can learn it.							
Treatment	5.2	3.6	40.9	50.2	1209	3.36	0.79
Comparison	2.3	4.3	32.5	60.9	486	3.52[c]	0.69

[b] Statistically significant difference between treatment and comparison schools, $p < .01$. [c] Statistically significant difference between treatment and comparison schools, $p < .001$.

Rate your level of agreement to the following statements.

	Strongly disagree	Disagree	Agree	Strongly agree			
	%	%	%	%	N	*M*	*SD*
Zearn is a good program for learning math.							
Treatment	6.1	4.7	34.9	54.3	1227	3.37	0.83
Zearn makes learning math fun.							
Treatment	8.9	11.7	34.5	44.8	1211	3.15	0.95
I look forward to using the Zearn math program.							
Treatment	8.3	10.4	38.2	43.0	1187	3.16	0.92
I would like to use Zearn Math in the future.							
Treatment	12.3	13.3	30.8	43.7	1138	3.06	1.03

Appendix G: Students' Responses to Open-Ended Section of the Student Reaction Questionnaire

Table 1.
Students' favorite components of Zearn Math (overall).

Code	Frequency
Learning	229
Tower of Power	225
Wild Cards	111
Makes math fun, easy	107
Math Chat	104
Substantive content	95
Achievement	91
Nothing, I do not like Zearn	81
Songs	75
Sprints	41
Jokes or Math Laughs	41
"Math games"	40
Uninterpretable	40
Everything, I love Zearn	35
Number Gym	32
Misc. features	26
Using technology	17
Boost	17
Zearn Squad	16
Learning Lab	15
Paper Materials	8

Table 2.
Students' favorite component of Zearn Math: Misc. features.

	Frequency
Miscellaneous features	**26**
Multiply Mania	*12*
Sum Snacks	*5*
Discovery Canyon	*5*
Blast	*4*

Table 3.
Favorite component of Zearn Math: Code descriptions.

Code	Description
Learning	Students describes learning or increased math skills
Tower of Power	Student mentions the Tower of Power in their response
Wild Cards	Student mentions Wild Cards in their response

Makes math fun, easy	Student explains Zearn Math makes learning fun or easy
Math Chat	Student mentions Math Chat in their response
Substantive content	Student describes substantive mathematics topics in their response
Achievement	Student references completing lessons, moving through the program, feelings of accomplishment
Nothing, I do not like Zearn	Student response includes "nothing" or overall dissatisfaction with the program
Songs	Student describes songs or Z Beats in the response
Sprints	Student mentions Sprints in their response
Jokes or Math Laughs	Student mentions Math Laughs or indicates Zearn Math is funny
"Math games"	Student describes a gamified feature, or mentions "math games"
Uninterpretable	Researchers are unable to understand the student's comments
Everything, I love Zearn	Student response includes "everything" or overall satisfaction with the program
Number Gym	Student mentions Number Gym in their response
Misc. features	Student mentions a miscellaneous feature in their response (see Table 2)
Using technology	Student mentions the use of technology or a preference for working on the computer
Boost	Student mentions Boost in their response
Zearn Squad	Student mentions Zearn Squad in their response
Learning Lab	Student mentions Learning Lab in their response
Print materials	Student references Exit Tickets, their workbook, or other paper-based materials in their response

Table 4.
Favorite component of Zearn Math: Examples of Codes.

Code	Example
Learning	"I like Zearn because it make me smart and good at math."
Tower of Power	"I really like tower of power after I complete a video."
Wild Cards	"I like when they have wild cards."
Makes math fun, easy	"My favorite is that they have fun stuff for math so it does not make it boring."
Math Chat	"I like the math chat because it helps me learn the stuff I need for tower of power."
Substantive content	"My favorite is learning about area."
Achievement	"My favorite thing is completing lessons and passing levels."
Nothing, I do not like Zearn	"I don't like anything about Zearn."
Songs	"I like the songs and music."
Sprints	"My favorite thing is the sprints."
Jokes or Math Laughs	"I like the funny jokes."
"Math games"	"My favorite part is all of the math games."
Uninterpretable	"Hi zearn."
Everything, I love Zearn	"I don't have a favorite thing. Zearn is just GREAT!!!"
Number Gym	"I like the games that comes before all the work."
Misc. features	
Using technology	"My favorite thing about Zearn is that it is on the computers."
Boost	"My favorite thing about Zearn is that they give you the boost in tower of power."
Zearn Squad	"My favorite part about Zearn is the people."

Learning Lab	"Learning Lab is my favorite."
Paper materials	"I like to do the exit ticket and turn it in."

Table 5.

Students' least favorite components of Zearn Math (overall).

	Frequency
Nothing; I love Zearn	285
Tower of Power (total)	265
"It's too hard."	180
Math Chat	109
Getting stuck; Frustration	107
Starting over; Losing points	95
I hate everything about Zearn	74
Takes too long; Zearn is boring	66
Misc. features	64
Uninterpretable	50
Getting things wrong	48
Sprints	44
Paper or print materials	36
Substantive content	34
Not enough chances	32
Zearn Squad	19
Technology-related concerns	18
Boost	17

Table 6.

Least favorite component of Zearn Math: Tower of Power.

	Frequency
Tower of Power (no context)	**117**
Tower of Power (with context)	**148**
Starting over; Losing points	*54*
It's hard	*36*
Getting stuck; Frustration	*32*
Not enough chances	*17*
Takes too long; Zearn is boring	*5*
Getting things wrong	*4*

Table 7.

Least favorite component of Zearn Math: Misc. features.

	Frequency
Miscellaneous features	**64**
Laughs or jokes	*12*
Number Gym	*10*
Blast	*7*

Learning Lab	*7*
Songs	*6*
More Wild Cards	*4*
Multiply Mania	*4*
Discovery Canyon	*3*
Sum Snacks	*2*
Totally Times	*2*
Mix and Match	*1*
Shape Shifter	*1*

Table 8.
Least favorite component of Zearn Math: Code descriptions.

Code	Description
Nothing	Student says, "nothing" or uses space to describe why they like Zearn Math
Tower of Power (total)	Student mentions the Tower of Power in their response
Tower of Power (no context)	Student provides only the name of the feature
Tower of Power (with context)	Student provides some context for why the Tower of Power is their least favorite
It's hard	Student describes Zearn Math as hard or difficult
Math Chat	Student mentions Math Chat or "videos" in their response
Getting stuck; Frustration	Student describes feeling stuck in the program, or feeling frustrated by their experiences
Starting over; Losing points	Student describes losing points or having to "go back"
I hate everything about Zearn	Student response includes "everything" or overall dissatisfaction with the program
Takes too long; Zearn is boring	Student describes Zearn Math as boring, or describes lessons as taking too long to complete
Misc. features	Student mentions a specific program feature (see Table 7)
Uninterpretable	Researchers are unable to make sense of the student's response
Getting things wrong	Student describes their least favorite component as making a mistake or getting something incorrect in the program
Sprints	Student mentions Sprints in their response
Paper or print materials	Student references Exit Tickets, their workbook, or other paper-based materials in their response
Substantive content	Student describes substantive mathematics topics in their response
Not enough chances	Student describes feeling the program does not allow them enough chances to demonstrate learning
Zearn Squad	Student mentions Zearn Squad in their response
Technology-related concerns	Student references technology or hardware concerns in their response
Boost	Student mentions Boost in their response

Table 9.
Least favorite component of Zearn: Examples of codes.

Code	Example
Nothing	"There is nothing bad about Zearn."
Tower of Power (total)	
Tower of Power (no context)	"Tower of Power."
Tower of Power (with context)	"The Tower of Power is hard for me."
It's hard	"They make it so hard."
Math Chat	"I don't like all the talking and long explanations in the videos."
Getting stuck; Frustration	"When I need help on Zearn and I just give up."
Starting over; Losing points	"When on the Tower of Power you get an answer wrong and the bar goes down."
I hate everything about Zearn	"I hate Zearn so much." or "I hate math."
Takes too long; Zearn is boring	"The lessons are very long. It's boring to me."
Misc. features	
Uninterpretable	"I don't like the chain."
Getting things wrong	"I dislike when I get my answers wrong."
Sprints	"Sprints because it goes too fast."
Paper or print materials	"Doing the sheets that go with the lesson."
Substantive content	"I don't like to divide."
Not enough chances	"They only give a few chances and then you have to go over it again."
Zearn Squad	"I do not like the Z Squad."
Technology-related concerns	"It's buggy sometimes and too slow."
Boost	"I dislike when I need a Boost."

Appendix H: Digital Component Usage Data

Table 1.
Student usage data for students in analytic sample (N = 3,803).

Class Grade	Active Students	Average Lessons On-Grade	Average Total Lessons	Average Student Minutes per Active Week	Average Student Active Weeks
K	517	46.23	46.23	19.97	9.79
1	620	41.19	63.15	55.56	21.41
2	718	32.05	39.48	53.44	24.01
3	875	49.32	57.28	77.59	31.05
4	822	46.3	51.05	85.37	30.79
5	768	43.66	53.31	83.72	29.63

Table 2.
Teacher usage data (N = 323).

Class Grade	Active Teachers	Average Teacher Minutes per Active Week	Average Teacher Active Weeks
K	30	19.0	6
1	54	51.0	15
2	55	36.0	15
3	67	39.0	18
4	68	49.0	21
5	63	58.0	18
6	9	16.0	9

Table 3.
Administrator usage data.

Class Grade	Active School Admin	Average Admin Minutes per Active Week	Average Admin Active Weeks
All	58	43	8.8

Appendix I: Subgroup Analyses Regression Models

Table 1
Regression results with prior achievement interactions: NWEA MAP sample

	Estimate	Standard Error	P-value
Zearn Math	0.103	0.553	0.852
Zearn*High Achievement	1.221	0.701	0.082
Zearn*Low Achievement	1.009 *	0.399	0.012
Constant	187.867 ***	0.350	0.000
Student N	8718		
School N	35		

Table Notes:
(a) *p<.05, ***p<.001

Table 2
Regression results with gender interaction: NWEA MAP sample

	Estimate	Standard Error	P-value
Zearn Math	0.498	0.575	0.387
Zearn*Female	-0.014	0.431	0.974
Constant	187.877 ***	0.352	0.000
Student N	8718		
School N	35		

Table Notes:
(a) ***p<.001

Table 3
Regression results with race/ethnicity interactions: NWEA MAP sample

	Estimate	Standard Error	P-value
Zearn Math	-0.261	0.802	0.745
Zearn*Black	0.614	0.700	0.381
Zearn*Latino	1.406	0.768	0.067
Constant	187.878 ***	0.352	0.000
Student N	8718		
School N	35		

Table Notes:
(a) ***p<.001

Table 4
Regression results with prior achievement interactions: State Assessment sample

	Estimate		Standard Error	P-value
Zearn Math	-0.013		0.062	0.839
Zearn*High Achievement	0.163	**	0.059	0.006
Zearn*Low Achievement	-0.089	*	0.045	0.046
Constant	0.029		0.035	0.403
Student N	3218			
School N	34			

Table Notes:
(a) *$p<.05$, ** $p<.01$, ***$p<.001$

Table 5
Regression results with gender interaction: State Assessment sample

	Estimate	Standard Error	P-value
Zearn Math	-0.029	0.056	0.605
Zearn*Female	-0.052	0.045	0.255
Constant	0.033	0.035	0.350
Student N	3218		
School N	34		

Table 6
Regression results with race/ethnicity interactions: State Assessment sample

	Estimate	Standard Error	P-value
Zearn Math	-0.017	0.083	0.844
Zearn*Black	-0.012	-0.012	-0.012
Zearn*Latino	-0.108	-0.108	-0.108
Constant	0.032	0.035	0.357
Student N	3218		
School N	34		