Evaluation Study of the Istation Early Reading Program in Idaho

Rebecca Wolf, PhD
Steven Ross, PhD
Jane Eisinger, MS
Alan Reid, PhD
Clayton Armstrong, BA BS

March 2020

JOHNS HOPKINS
SCHOOL *of* EDUCATION

Center for Research and
Reform in Education

About the Center for Research and Reform in Education (CRRE) at Johns Hopkins University

Established in 2004, the Center for Research and Reform in Education (CRRE) works to improve the quality of education for children in grades pre–K through 12 through its program effectiveness research. CRRE's work focuses on evaluating approaches to teaching and school organization, with studies completed on teacher professional development, comprehensive school reform, educational technology, resource allocation, bilingual education, reading and math curricula and assessments, and after-school and summer learning programs.

Affiliated with Johns Hopkins University's School of Education, the CRRE research team includes Johns Hopkins University professors, research scientists, and research staff with backgrounds in quantitative, qualitative, and evaluative research. The research team has published over 200 research reports, and within the past five years, conducted over 45 program evaluations.

Contents

EXECUTIVE SUMMARY:
Evaluation Study of the Istation Early Reading Program in Idaho

The Center for Research and Reform in Education (CRRE) is a research center affiliated with the School of Education at Johns Hopkins University (JHU) that specializes in education program evaluations in K–12. Istation contracted with the CRRE at JHU to conduct a study of the effectiveness of Istation's Early Reading (ER) program ([Link to iStation's website](#)) in the state of Idaho.

Istation is a digital-based instructional intervention tool for various content areas aimed at pre-K through 8th grade learners. The Istation Reading program consists of formative assessments, named Istation Indicators of Progress (ISIP™), which are computer-adaptive and diagnostic literacy assessments designed to track student growth over time. Istation Reading also includes an adaptive, online curriculum, which generates personalized learner data profiles that teachers can use to make data-driven instructional decisions and assign custom learning interventions. Istation's Early Reading (ER) program was designed specifically for students in grades K–3, focusing on the critical areas of early reading, including phonemic awareness, phonics, vocabulary, reading comprehension, and fluency.

The ISIP became the state of Idaho's early literacy assessment in the 2018–19 school year. The state-mandated early literacy assessment for students in grades K–3 is referred to as the Idaho Reading Indicator (IRI). The state of Idaho used a different assessment in previous years and switched to using the ISIP for the IRI in the 2018–19 school year. The IRI is administered to all K–3 public school students in the state, and is intended to serve as an early reading diagnostic and screener. A sample of public elementary schools in Idaho piloted the ISIP and the Istation curriculum during the 2017–18 school year. In the 2018–19 school year, all public elementary schools in Idaho administered the ISIP. Schools were required to administer the ISIP in both the fall and spring, and had the option to administer it more frequently for yearly progress monitoring. Schools also had the option of purchasing Istation's related curricular resources.

This study examines effects of the Istation ER program on student reading achievement in the state of Idaho and the validity of the ISIP for predicting student performance on the Idaho Standards Achievement Test (ISAT). It also highlights implementation successes and challenges experienced by educators who piloted Istation's ER program during the 2017–18 school year. The main findings of the study are summarized below.

Findings

In the 2018–19 statewide sample of schools, there were positive associations between increased Istation usage and improved student performance on the spring 2019 IRI (ISIP) and ISAT, relative to schools with the lowest levels of Istation usage.

Higher average number of sessions, number of weeks, and total minutes of use for either progress monitoring and/or curriculum purposes were each positively associated with improved student performance on the spring 2019 IRI (ISIP) and ISAT, relative to schools with the lowest quartile of Istation usage. The percentage of students in the school who used Istation (for progress monitoring and/or curriculum) was also related to improved student performance on the spring 2019 IRI (ISIP), relative to schools with the lowest level of usage. Students in schools where the majority used Istation (testing only or testing and curriculum) outperformed students in schools where a smaller percentage of students used Istation.

Students in schools that used Istation curriculum had higher average student performance compared with students in schools that did not use Istation curriculum. This finding implies that in this first year of the statewide ISIP usage as the IRI, there may have been some benefit to student performance when schools also implemented the Istation curriculum. The average time students spent using the curriculum was not related to improved student performance, however.

Student scores on the ISIP in the second and third grades were good predictors of student scores on the ISAT ELA, which is administered to students starting in the third grade.

The approximate correlation between second and third grade students' ISIP scores and students' scores on the ISAT ELA was around .70. In addition, second and third grade student scores from the fall, winter, and spring administrations of the ISIP were all good predictors of student scores on the ISAT ELA.

Pilot school students outperformed similar comparison students on the spring 2018 IRI when the vast majority of students in the school used Istation, for either progress monitoring and/or curriculum purposes.

While higher levels of usage of Istation were generally unrelated to improved student performance in reading relative to students in comparison schools on both the IRI and ISAT ELA administered in the springs of 2018 and 2019, we found a positive association between the percentage of students in the school who had used Istation and improved student performance on the spring 2018 IRI, on average. This finding suggests that a schoolwide implementation of Istation may yield greater benefits than a more piecemeal approach. Yet this finding should also be interpreted with caution, given that this study cannot rule out systematic differences among schools with different levels of Istation usage.

Pilot school students outperformed similar comparison students on a few IRI subtests.

Students in pilot schools outperformed similar comparison peers on some of the IRI and ISAT subtests, on average. Students in pilot schools had higher average growth on the spring 2018 IRI letter sound fluency subtest and on the vocabulary and text fluency subtests of the spring 2019 IRI (ISIP) relative to comparison students. While this study cannot rule out that these results occurred due to chance alone as a result of the large number of statistical analyses conducted, students in pilot schools did not underperform students in comparison schools on any subtest.

Educators generally expressed positive opinions of Istation during its initial implementation in the 2017–18 school year.

Educators interviewed in the state of Idaho generally expressed positive opinions of Istation during its initial implementation in the 2017–18 school year. What teachers liked best was the immediate feedback and rich data on student performance, which allowed them to adapt their instruction and reduced guesswork in how best to target interventions to individual students. Teachers also reported that students generally liked using Istation, citing its game-like attributes and personalized approach to learning.

Discussion

This study examined the relationship between the implementation of Istation, for either progress monitoring and/or use of the online curriculum, and improved student performance in reading in the state of Idaho in the 2017–18 and 2018–19 school years. The first research question concerned the relationship between Istation usage and student performance for all schools statewide during the 2018–19 school year. Results consistently showed a positive association between increased Istation usage and improved student performance, relative to schools with the lowest levels of Istation usage. Moreover, the sample size was large and included all public schools serving grades K–3 in Idaho. Therefore, this study supports the conclusion that moderate amounts of Istation usage related to improved student performance in reading, relative to schools with the lowest levels of Istation usage. Findings should be interpreted with caution, however, because this study cannot rule out systematic differences between schools with different Istation usage levels. As such, this study provides "promising" as opposed to causal evidence of the efficacy of Istation in improving student performance in reading per the Every Student Succeeds Act (ESSA).

In addressing the second research question, findings showed that the ISIP also was a good predictor of students' scores on the ISAT ELA, which is the state accountability assessment for students starting in grade 3. The correlation between students' ISIP and ISAT ELA scores was relatively high at around .70. Therefore, educators can look to second and third grade students' ISIP scores to forecast how

their students will fare on the ISAT ELA, which is administered in the spring of students' third grade year.

The study also addressed a third, exploratory question comparing student outcomes for a group of schools that piloted the Istation assessment and curriculum components in 2017-18 and similar, comparison schools. Relative to students in comparison schools, students in pilot schools did not perform significantly higher on either ISIP or ISAT achievement tests. However, they did outperform comparison students on some IRI subtests in the spring of 2018 and 2019. Additionally, students in pilot schools outperformed similar comparison peers on the spring 2018 IRI when the vast majority of students in the school participated in Istation. Usage data confirmed that the typical pilot school made tangible efforts to implement Istation components. However, in this initial year, usage was limited in both absolute and relative terms compared with time devoted to regular reading and language arts instruction. For example, the typical pilot student received only 2–3.5 hours of total exposure to Istation during the school year.

A fourth research question concerned the perceptions of educators who piloted Istation in 2017-18. Results revealed reports of mostly positive experiences. Educators believed that the ISIP provided them with more nuanced information about their students' performance than did the state's previous IRI, which enabled them to adapt their instruction and target specific gaps in learning.

Conclusion

In conclusion, this study yielded a number of generally favorable correlational results of the benefit of Istation products on student performance in reading. However, the study design was limited for examining the efficacy of the Istation performance monitoring and curriculum due to what appears to have been limited or at best modest implementation at many schools. For example, in the typical school in the statewide sample, Istation was used (progress monitoring and/or curriculum) with only two-thirds of students in the school[1] and curriculum was used for only 9% of students. Therefore, while these results depict statewide patterns for usage and associated performance outcomes, they may understate effects for schools that used the program at optimal levels. Another limiting factor was that many teachers were implementing Istation for the first time and therefore still learning how to employ it effectively.

More research is needed on the efficacy of Istation, whereby student performance in schools opting to implement the progress monitoring and curriculum with high fidelity is compared over multiple years to the performance of students in

---

[1] The usage metric was determined at the school level because student-level usage data for progress monitoring were not available; therefore, a greater percentage of students in grades K-3 participated in the progress monitoring and mandatory testing than all students schoolwide.

schools that do not participate in Istation or participate only in the progress monitoring. Having reliable student and classroom level usage metrics would further allow for rigorous analyses examining the relationship between usage of Istation and improved student performance.

Evaluation Study of the Istation Early Reading Program in Idaho

The Center for Research and Reform in Education (CRRE) is a research center affiliated with the School of Education at Johns Hopkins University (JHU) that specializes in education program evaluations in K–12. Istation contracted with the CRRE at JHU to conduct a study of the effectiveness of Istation's Early Reading (ER) program (Link to Istation's website) in the state of Idaho.

Istation is a digital-based instructional intervention tool for various content areas aimed at pre-K through 8th grade learners. The Istation Reading program consists of formative assessments, named Istation Indicators of Progress (ISIP™), which are computer-adaptive and diagnostic literacy assessments designed to track student growth over time. Istation Reading also includes an adaptive, online curriculum, which generates personalized student data profiles that teachers can use to make data-driven instructional decisions and assign custom learning interventions. Istation's Early Reading (ER) program was designed specifically for students in grades K–3, focusing on the critical areas of early reading, including phonemic awareness, phonics, vocabulary, reading comprehension, and fluency.

The ISIP became the state of Idaho's early literacy assessment in the 2018–19 school year. The state-mandated early literacy assessment for students in grades K–3 is referred to as the Idaho Reading Indicator (IRI). The state of Idaho used a different assessment in previous years and switched to using the ISIP for the IRI in the 2018–19 school year. The IRI is administered to all K–3 public school students in the state, and is intended to serve as an early reading diagnostic and screener. A sample of public elementary schools in Idaho piloted the ISIP during the 2017–18 school year. In the 2018–19 school year, all public elementary schools in Idaho administered the ISIP. Schools were required to administer the ISIP in both the fall and spring, and had the option to administer it more frequently for yearly progress monitoring. Schools also had the option of purchasing Istation's related curricular resources.

This study examines the efficacy of the Istation ER program in improving student reading achievement in the state of Idaho. The study also highlights implementation successes and challenges experienced by educators who piloted Istation's ER program during the 2017–18 school year.

Program Description

The Istation ER program centers on four areas: assessment, instruction, reporting, and teacher tools.[2] The assessment is administered via the Istation Indicators

---

[2] Putman, R.S. (2017). Technology versus teachers in the early literacy classroom: an investigation of the effectiveness of the Istation integrated learning system. *Educational Technology Research & Development, 65*, 1153-1174.

of Progress (ISIP™), which is a computer adaptive testing system designed to assess students' early reading skills. Students complete an online assessment consisting of multiple choice and fill-in-the-blank questions, lasting approximately 40 minutes. These assessments may be used as benchmarks and as continuous progress monitoring tools, with the option of monthly assessments assigned to students by Istation based on their skill levels.

The instructional component of Istation provides interactive, computer-based instruction to students in the critical areas of early reading literacy (e.g., phonemic awareness, phonics, vocabulary, reading comprehension, and fluency) that is adaptive to student performance on the assessment. Detailed data profiles are created for each student based on their interactions and performance within Istation. Data include an ability index (a calculated estimate of a student's reading ability), national norms, tiers for response to intervention strategies, grade-level equivalencies, and Lexile reading scores. Data analytics on individual student performance are made available to teachers through various data reporting tools such as skill growth reports, classroom and student summary reports, usage trend reports, and executive summary reports, which provide key insights into student performance. These insights facilitate pinpointing gaps in students' skills and differentiated instruction.

In addition to assessment and instructional tools for students, Istation provides over one thousand lesson plans and other instructional resources to help teachers personalize student learning.[3] The online lesson libraries host teacher-directed lessons, which include handouts and scripts for teachers, as well as on-demand assessments that allow teachers to assign independent work in targeted areas to students.

Research Questions

The purpose of this study was two-fold. First, the study investigated the relationships between the extent of Istation usage and student learning gains in reading. Second, the study investigated the validity of the Istation ISIP in predicting students' scores on the ISAT. Third, in an exploratory study component, we compared the reading achievement for students attending schools that piloted the Istation ER program during the 2017–18 school year and students attending comparison schools that did not participate in the pilot. Finally, the study documented educator perceptions and attitudes towards the Istation ER program. The research questions included the following:

1. Was greater usage of the Istation ER program related to higher learning gains in reading and ELA?

---

[3] Istation Reading. Istation. Retrieved from (Link to Istation's website)

2. How predictive of third grade student performance on the ISAT ELA test were second and third grade students' ISIP scores?

3. How did students in schools that piloted the Istation ER program in 2017-18 compare in performance in reading and English language arts (ELA) to students in similar schools that did not participate in the pilot?

4. What were implementation successes and challenges faced by educators in schools that piloted the ISIP during the 2017–18 school year?

Methods

Evaluation Design

Istation usage and learning gains. For all schools in the 2018–19 school year, we analyzed whether increased Istation usage was related to improved reading achievement relative to schools that used Istation to a lesser extent. In other words, for Idaho schools with more or less Istation usage, we analyzed the difference in student performance on the Idaho Reading Indicator (IRI) and ISAT in English language arts (ELA) in the spring of 2019. These analyses inform potential associations between Istation usage and student learning gains in reading, and if findings are statistically significant, would yield ESSA Tier 3 evidence.

Istation versus comparison. In this component of the study, we compared student performance for schools that piloted the Istation ER program and schools that did not participate in the program during the 2017–18 school year. We employed a quasi-experimental study design by identifying comparison schools that were similar to the pilot schools in terms of prior student achievement, student demographics, and school and district characteristics—see Appendix A for more details. For pilot and comparison schools, we analyzed the difference in student performance on the IRI and ISAT in ELA in the springs of 2018 and 2019. Although all schools participated in the ISIP in the 2018–19 school year, we examined whether students in pilot schools that had been implementing the ISIP for two consecutive years outperformed students in comparison schools that had implemented the ISIP for only one school year as of spring 2019.

Given that the pilot schools were in their first year of using Istation, and appeared to be mostly autonomous in their implementation strategies and usage decisions, we viewed this quasi-experimental design (QED) as exploratory for identifying any statistically significant outcomes or suggestive patterns that might occur. In addition, although we conducted rigorous analyses, this type of study does not support causal inferences about the impact of Istation. There may have been unobserved factors, such as principal or teacher quality or student characteristics, that were unrelated to Istation implementation and yet may have influenced the findings.

Mixed methods. Finally, the study incorporated a qualitative component to provide more context about Istation implementation in the schools that piloted the ISIP during the 2017–18 school year. Site visits and phone interviews were conducted with educators in eight schools.

Data

Student achievement data. Student achievement data were collected by CRRE from the Idaho State Department of Education (SDE). The SDE also provided data on student demographics and school and district characteristics. Student achievement was analyzed on the following assessments:

- The previous IRI, which was administered bi-annually to students in grades K–3 in the 2017–18 school year and before implementation of ISIP as the new IRI. The IRI composite scores were limited in capturing variation in student performance in that they ranged from 1 to 3 and were integer values only. Sub-tests included letter naming fluency, letter sound fluency, and a reading curriculum-based measure, and each of these had a different range of values.
- The ISIP, which replaced the previous IRI in 2018–19, was administered bi-annually to students in grades K–3 in the 2018–19 school year. Student ISIP scores for K–3 students in this sample ranged from approximately 60 to 375. The sub-tests include phonemic awareness, letter knowledge, alphabetic decoding, spelling, vocabulary, listening comprehension, and text fluency.
- The ISAT in ELA, which is administered to students in grades 3–8 annually in the spring. Student scores on this measure range from 2000 to 3000.[4] The sub-tests include reading, writing, listening, and research.

In addition to collecting fall and spring ISIP scores from the SDE for all Idaho students in grades K–3 in 2018–19, CRRE also collected 2017–18 ISIP scores for students in Istation pilot schools only, as well as 2018–19 winter ISIP scores for students in schools that participated in a winter administration of the ISIP.[5]

Istation usage data. In the 2017–18 school year, some schools piloted the ISIP, and in the 2018–19 school year, all schools were required to administer the ISIP as the new IRI. Schools had the option of using Istation for monthly progress monitoring, which was generally implemented schoolwide. Schools also had the option to purchase additional Istation curricular and online resources, which may have been implemented in specific classes or grades or with a certain group of students, as opposed to schoolwide. Istation usage indicators were determined at the school level,

---

[4] See Link to Idaho State Department of Education for more information.
[5] The spring ISIP score defaulted to the May score. The fall ISIP score defaulted to the (1) September, (2) October, or (3) August score, in that order. The winter ISIP score (if non-missing) defaulted to the (1) January, (2) February, or (3) December score, in that order.

and some indicators captured usage of all components of Istation (e.g., progress monitoring and curriculum), while other indicators captured use of curriculum components only. The following usage indicators captured usage of both progress monitoring and curriculum tools:

- Average sessions: The number of times students logged into their Istation accounts, averaged at the school or school-by-grade level.
- Average weeks: The number of weeks students used Istation, averaged at the school or school-by-grade level.
- Average total minutes: The number of minutes spent using Istation, averaged at the school or school-by-grade level.
- Percentage Istation users: The percentage of students in the school who used Istation, determined by the number of Istation users in the school divided by the school's enrollment and multiplied by 100.[6]

In addition, the following indicators specifically captured use of Istation curriculum tools only:

- Average curriculum minutes: The number of minutes spent using Istation curriculum, averaged at the school or school-by-grade level.
- Percentage Istation curriculum users: The percentage of students in the school who used Istation curriculum, determined by the number of Istation curriculum users in the school divided by the school's enrollment and multiplied by 100.
- Percentage Istation curriculum threshold users: The percentage of students in the school who used Istation curriculum at the recommended levels, determined by the number of Istation curriculum users in the school who used Istation at the recommended levels divided by the school's enrollment and multiplied by 100.

Sample

Statewide sample. Correlational analyses examined the relationship between extent of Istation usage and improved student performance in reading and ELA in the 2018–19 school year, when all Idaho schools serving grades K–3 implemented the ISIP. All students in grades K–3 with non-missing achievement data were included in the analyses. Table 2 provides the demographic characteristics of the statewide sample of students included in the 2018–19 analyses of Istation usage. As shown in Table 1, the majority of students were White and about half were economically disadvantaged. Special education and English learner students each comprised about 10% of the student sample.

---

[6] In some cases, the number of Istation users exceeded the school's enrollment; in these cases, the percentage of Istation users in the school was recoded to 100%. Therefore, these usage indicators are good proxies for Istation usage, but they may not represent exact values.

Table 1: Statewide sample in grades K–3 in 2018–19

	Istation
Special education	9.8%
White	75.3%
Latino	18.3%
Other race	6.4%
Economically disadvantaged	49.7%
English learner	10.1%
Student N	85,747
School N	407

Pilot versus comparison schools. To compare student performance in reading and ELA for pilot and comparison students, a sample of pilot and comparable schools was first selected (See Appendix A for more details). Student performance on state assessments was then compared for pilot and comparison students within these schools. The sample included students in grades K–3 in the 2017–18 school year, and students in grades K–4 in the 2018–19 school year. As shown in Table 2, this selected sample was similar to the overall state sample in terms of student demographics. Moreover, the demographic characteristics for selected Istation and comparison students were similar. Economically disadvantaged was the only student characteristic in which there was a statistically significant difference between pilot and comparison students. However, the difference in percentage economically disadvantaged was relatively small at 5 percentage points or less, and the analyses controlled for this student characteristic, among others.

Table 2: Sample characteristics for pilot and comparison students

	Pilot	**Comparison**	**Difference**
2017-18 outcomes analyses			
Special education	9.7%	9.6%	0.1%
White	76.3%	75.8%	0.5%
Latino	18.2%	18.8%	-0.5%
Other race	5.4%	5.4%	0.0%
Economically disadvantaged	56.3%	53.4%	2.9%***
English learner	9.8%	9.7%	0.1%
2018-19 outcomes analyses			
Special education	11.2%	11.4%	-0.1%
White	76.8%	76.0%	0.9%
Latino	17.9%	18.6%	-0.7%
Other race	5.3%	5.5%	-0.2%
Economically disadvantaged	55.4%	50.0%	5.4%***
English learner	10.1%	9.8%	0.4%

The student and school sample sizes differed across the outcomes analyses, depending on outcome measure and school year, as shown in Table 3. Note that the ISAT was administered only to students in grade 3, hence the smaller sample sizes. Table 3 also provides the difference in baseline achievement (in terms of standard deviations) for each of the analyses. Baseline achievement for all analyses was determined using students' fall 2017 IRI score, which was collected prior to or near the time when some schools began piloting Istation. Pilot and comparison students were very similar in baseline achievement, and baseline achievement was controlled for in the analyses.[7]

Table 3: Baseline equivalence for pilot and comparison students

	Student N		School N		Difference in baseline achievement (Fall 2017 IRI)
	P	**C**	**P**	**C**	
2017-18 outcomes analysis					
IRI sample	16,155	13,198	70	70	0.03
ISAT sample	4,005	3,336	68	68	0.01
2018-19 outcomes analysis					
ISIP sample	11,523	9,172	70	69	0.04
ISAT sample	7,082	6,090	68	68	0.03

NOTES—1. P=pilot and C=comparison. 2. The difference in baseline student achievement is in terms of standard deviations. 3. One school was included in the 2017–18 comparison sample but dropped in the 2018–19 comparison sample because the school identification number changed between the 2017–18 and 2018–19 school year. 4. Four schools were included in the IRI and ISIP samples that were not included in the ISAT samples because the data included students in grades K–2 only.

Pilot schools were those that used Istation progress monitoring in the 2017–18 school year, although nearly all schools also used Istation curriculum with at least some of their students.[8] Because Istation progress monitoring was predominantly a schoolwide intervention for students in grades K–3, pilot status for students was determined by the school where they attended. For example, if a student attended a school piloting Istation in the 2017–18 school year and then transferred to a comparison school in the 2018–19 school year, the student would be included in the pilot sample for 2017–18 analyses but in the comparison sample for 2018–19 analyses. Table 3 above indicates that baseline achievement equivalence was met for the pilot and comparison student samples, despite some students switching schools between the 2017–18 and 2018–19 school years.

---

[7] The differences in baseline achievement satisfy the requirement in the What Works Clearinghouse (WWC) (2017) Standards Handbook that samples differ by no more than 0.25 standard deviations.
[8] Of the selected Istation school sample, only 6 out of the 70 schools did not also use Istation curriculum products in the 2017–18 school year.

Analytic Approach

Statewide analyses. For the statewide analyses in the 2018–19 school year, we used hierarchical linear modeling to compare student performance in reading and ELA for students in schools with more or less usage of Istation. We examined the effects of Istation on the following outcome measures:

- ISIP in spring 2019
- ISAT in spring 2019 (for third and fourth grade students only)

In the analyses, students were clustered within their schools. The models also controlled for:

- At the student level: grade level, English learner status[9], race/ethnicity[10], low-income, and baseline achievement in terms of students' fall 2018 ISIP score.
- At the school level: categorical Istation usage indicator, district or LEA enrollment, the percent of teachers at the school with 6+ years of experience, charter school status, percent English learner, percent special education, percent low-income, percent at each grade level, and average baseline achievement.

Other available variables were either redundant with these student and school covariates or unimportant in explaining variation in any of the analyses.

Because there was no comparison group, the models included categorical Istation usage indicators at either the school or school-by-grade levels. These analyses inform the improvement in student performance associated with an increase in Istation usage, relative to schools with less Istation usage. These analyses examined the association between the extent of Istation usage to student performance on ISIP and ISAT in spring 2019, while also controlling for baseline achievement and other student and school characteristics. For these analyses, baseline achievement was student scores on the fall 2018 ISIP.

We also examined whether there appeared to be differential effects of Istation for students with different demographic characteristics or prior achievement, and for schools with specific characteristics. To do so, we added interaction terms between the Istation usage indicator and student- and school-level covariates of interest.

---

[9] English learner status included three variables: English learner and within that subgroup, new or continuing English learner.

[10] Given the student demographics in Idaho, students were grouped on race/ethnicity into three categories: White, Latino, or other.

Predictive validity of ISIP. Next, we examined how well second and third grade students' fall, winter, and spring ISIP scores predicted their performance on the ISAT ELA in the third grade. To do so, we used the same hierarchical linear modeling approach described above with students clustered within their schools, but only students' ISIP scores were included in the model as a covariate. In other words, the model did not include student demographics or school characteristics, as policymakers would most likely be interested in the overall predictive validity of students' ISIP scores. We also descriptively compared the percentages of students who scored at various ISIP tiers with the percentages of students who scored at various ISAT levels.

Pilot versus comparison schools. We used hierarchical linear models similar to those described above to compare student performance in reading and ELA for students in pilot and comparison schools. One difference was that students' fall 2017 IRI score in reading was used as the baseline achievement measure. We compared these groups on the following:

- Previous IRI in spring 2018
- ISAT in spring 2018 (for third grade students only)
- ISIP in spring 2019
- ISAT in spring 2019 (for third and fourth grade students only)

In these models, Istation usage was also determined at the school level as was denoted by an Istation indicator (yes or no). Finally, to examine relationships between extent of Istation usage and improved student performance, the Istation indicator (yes or no) was replaced with categorical Istation usage indicators at either the school or school-by-grade level. These usage analyses conducted for pilot and comparison schools inform the improvement in student performance associated with an increase in Istation usage (defined in multiple ways), relative to schools with no Istation usage.

Findings

To address Research Question 1 (RQ1), we first present findings on the associations between the extent of Istation usage and student learning gains in reading in the statewide sample of Idaho schools serving grades K–3. For RQ2, we document the degree to which students' ISIP scores predicted their performance on the ISAT in English language arts (ELA). Third, for RQ3, we report the differences in student performance in reading and ELA for pilot schools relative to comparison schools. Finally, for RQ4, we discuss educator feedback and perceptions regarding the successes and challenges of implementing Istation in the 2017–18 school year.

Statewide Implementation: Istation usage and learning gains

Istation usage statewide. We explored to what extent all elementary schools in Idaho used Istation during the 2018–19 school year. All schools serving students in

grades K–3 were required to implement the ISIP in both the fall and spring as the state-mandated IRI, and some schools opted to also use Istation's progress monitoring and curriculum tools. Because there may be diminishing returns on usage[11], we first categorized usage by quartiles as follows:

- Low usage: Usage less than or equal to the 25th percentile
- Mid usage: Usage greater than the 25th percentile and less than or equal to the 50th percentile
- Mid-high usage: Usage greater than the 50th percentile and less than or equal to the 75th percentile
- High usage: Usage greater than the 75th percentile

In addition, the usage indicators that specifically captured usage of Istation curricular tools were estimated only for schools with non-zero curriculum usage.

In the typical school in the 2018–19 school year and statewide school sample, about 64% of students used Istation for either progress monitoring and/or curriculum. It is still possible that schools administered the ISIP to all of their students in grades K–3, as schools served students outside of these grades (i.e., higher grades may have accounted for all or most of the 36% of non-using students). In the typical school that also used Istation's curriculum, 9% of students in the school used the curriculum, and a percentage close to 0% used the curriculum at the recommended levels. Figure 1 shows the medians and interquartile ranges for the percentages of students enrolled in each school who used Istation, either for progress monitoring or curriculum purposes.

Figure 1: Medians and interquartile ranges for the school percentages of Istation users in the 2018–19 school year

---

[11] We explored potential diminishing returns on usage. From descriptive analyses, it appeared that diminishing returns to usage tended to occur at usage levels around or above the 75th percentile. There did not appear to be diminishing returns to usage in all cases, however, or on the percentages of Istation users in the school, as this number was bounded between 0 and 100%.
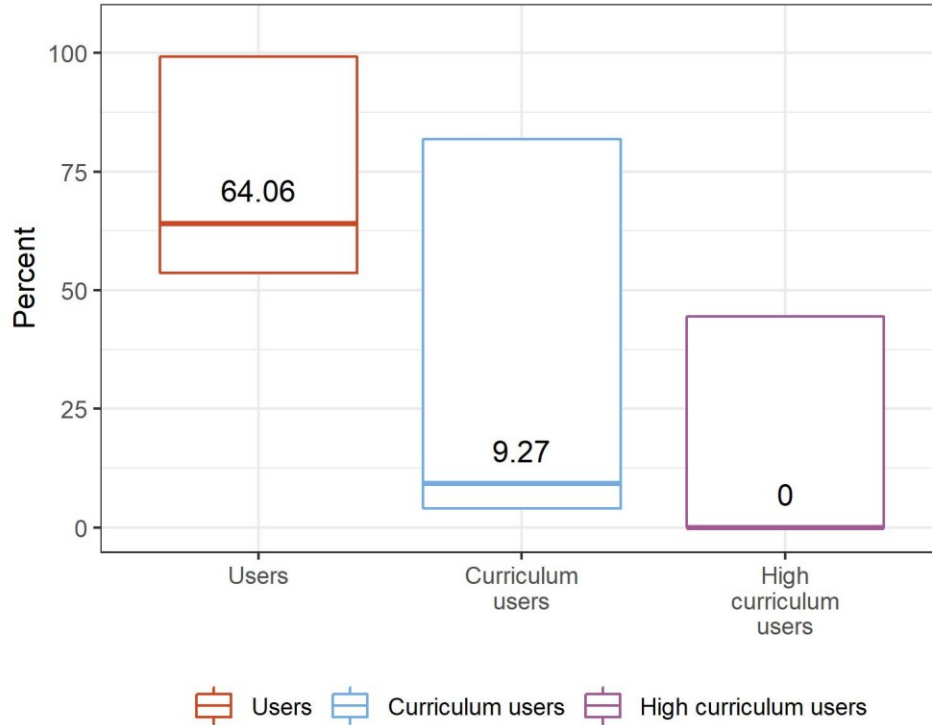
NOTE—Only schools with some Istation curriculum users were included in the estimates of the percentages of students in the school who used the curriculum.

Students in the typical school in 2018–19 used Istation for an average of 9–11 sessions over the course of about 9 weeks, depending on grade level. Istation usage was slightly higher on average for students in the first grade compared with other grades. Figure 2 shows the medians and interquartile ranges of the number of sessions and weeks of Istation usage by grade.

Figure 2: Medians and interquartile ranges for the number of sessions and weeks by grade in the 2018–19 school year
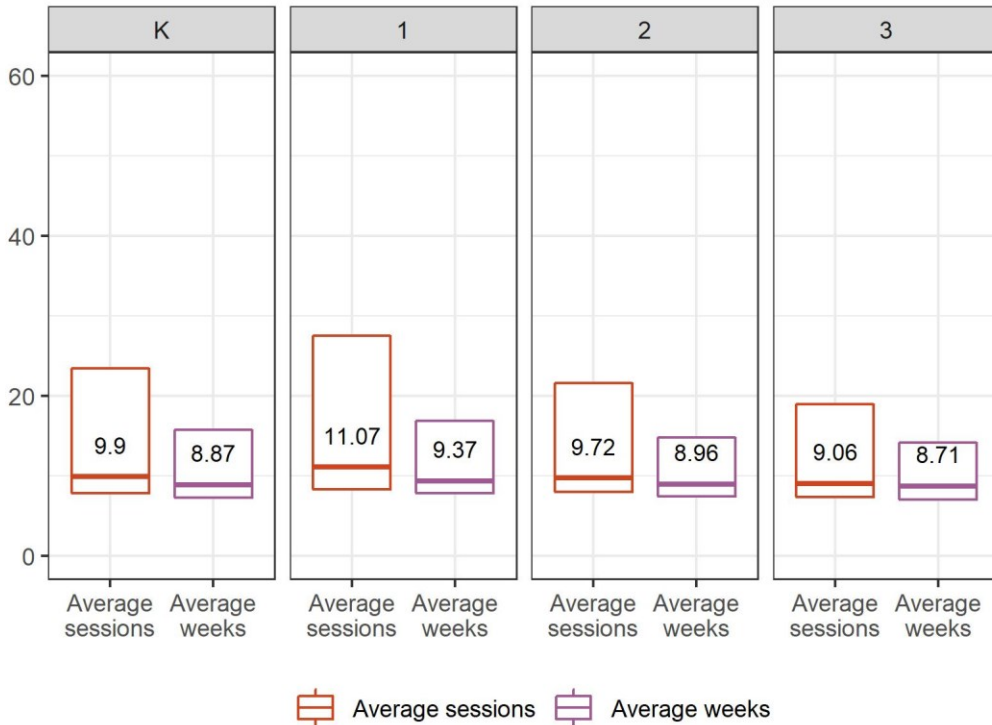
In terms of total time spent using Istation, students in the typical school in the 2018–19 school year used Istation for at least 2–3.5 hours total, with the most time spent using Istation on average occurring for students in the first grade. When schools also used the curriculum, the additional time spent with the curriculum averaged 1–2.5 hours, depending on grade. Students in kindergarten spent the least amount of time and students in the first grade spent the most amount of time using the curriculum, on average. Figure 3 shows the medians and interquartile ranges for total and curriculum minutes of Istation usage by grade.

Figure 3: Medians and interquartile ranges for the total and curriculum minutes of Istation usage by grade in the 2018–19 school year
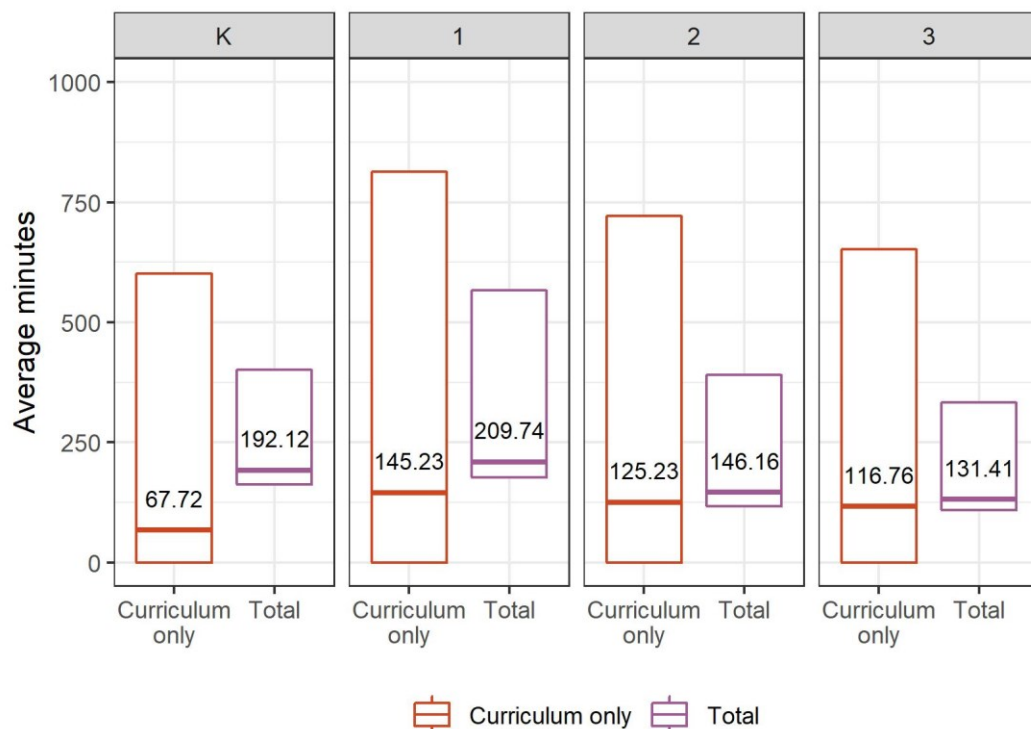
NOTE—Only schools with non-zero curricular usage were included in the estimates of time spent on Istation curriculum only.

Associations between Istation usage and student learning gains. We determined the associations between increased Istation usage and improved student performance in the statewide school sample, relative to schools with lower levels of Istation usage. We used the indicators of Istation usage previously described. Istation usage was first categorized into quartiles of usage at the school or school-by-grade levels using the statewide school sample in 2018–19.

When comparing schools that had the lowest levels of Istation usage to schools having higher usage, we found some positive associations between usage and student change in performance from fall 2018 to spring 2019. Higher average number of sessions, number of weeks, total minutes of use for either progress monitoring and/or curriculum purposes were each positively associated with improved student performance. This analysis also controlled for student and school characteristics and students' prior achievement on the 2018 fall ISIP. Full regression results are available in Appendix B.

As shown in Table 4, students in schools with the non-lowest Istation usage levels outperformed students in schools with the lowest quartile of Istation usage on the spring 2019 IRI (ISIP). More specifically, students in schools with an average of 9+

Istation sessions outperformed similar peers in schools with less Istation usage by an average of 2–3 points (p<.001). Similarly, students whose schools participated in Istation for at least 8 weeks outperformed similar peers in schools that used Istation less by an average of 2–3 points (p<.001). Finally, students had higher performance by an average of 2–3 points (p<.001) in schools where the typical student spent at least 2.3 hours (142 minutes) using Istation relative to schools where students spent less time using Istation. These gains of 2–3 points represent about 9–14% of the average student's annual gain on the ISIP. However, as previously noted, average gains may be larger or smaller for students depending on their grade level and prior achievement, and thus, this interpretation should be used with caution.

Table 4: Associations between schoolwide Istation usage and improved student performance on spring 2019 IRI (ISIP)

Usage indicator	Usage amount	Average IRI (ISIP) score
Progress monitoring + curriculum		*No comparison group*
Average sessions		
Low	<=8	231
Mid	9-10	233**
Mid-high	11-23	234***
High	>23	232***
Average weeks		
Low	<=7	230
Mid	8-9	233***
Mid-high	10-14	233***
High	>14	233***
Average min		
Low	<=141	231
Mid	142-173	233***
Mid-high	174-475	233***
High	>475	232***
Percent users		
Low	<=54	232
Mid	55-64	232
Mid-high	65-99	233*
High	>99	232
Average curriculum min		
No curriculum usage	0	232
Low	near 0	232
Mid	1-68	233
Mid-high	69-551	233
High	>551	233

Usage indicator	Usage amount	Average IRI (ISIP) score
Percent curriculum users		
No curriculum users	0%	230
Low	<=4%	232*
Mid	5-9%	233***
Mid-high	10-82%	233***
High	>82%	233***

NOTES—1. *$p<.05$, **$p<.01$, ***$p<001$. 2. The statistical significance refers to the difference in average student scores relative to schools with the lowest quartile of Istation usage. 3. The model also controlled for student grade level, English learner status, special education status, low-income status, race/ethnicity, and prior achievement (fall 2018 ISIP), as well as LEA enrollment, percentage of teachers at the school with 6+ years of experience, charter school status, school mean prior achievement, percentages of students in the school at various grade levels, and percentages of students in the school who were English learners, low-income, or special education.

The percentages of students in the school who used Istation (for progress monitoring and/or curriculum) were also related to improved student performance on the spring 2019 IRI (ISIP), relative to schools with the lowest level of usage. Students in schools where between 65–99% of students used Istation for progress monitoring and/or curriculum purposes outperformed students in schools where a smaller percentage of students used Istation by an average of 1 point ($p<.05$), about 5% of an annual learning gain for a typical student. Similarly, students in schools that used Istation curriculum had higher average student performance by 2–3 points ($p<.05$) compared with students in schools that did not use Istation curriculum. On the other hand, the average time students spent using Istation curriculum was not related with improved student performance. It is possible, however, that average minutes of curriculum use at the school level reflect usage for a relative small percentage of students in the school. Therefore, the percentage of students in the school who used the Istation curriculum may be a better indicator of schoolwide curriculum usage.

Results in Table 4 also indicate that there may have been diminishing returns on Istation usage after a certain point. Students scored higher on average on the spring 2019 IRI (ISIP) when in schools with the non-lowest level of Istation usage. However, there did not appear to be meaningful differences between average student performance among schools with greater but different amounts of Istation usage. It is also important to note that many of the usage indicators—average number of sessions, weeks, and minutes—were highly correlated with one another, and therefore, can be used somewhat interchangeably as general indicators of Istation usage.

Turning to findings regarding student performance on the spring 2019 ISAT ELA, we also found positive associations between Istation usage in the third grade and average third grade student performance (see Table 5). Students in schools with an average of 8–10 Istation sessions for third graders outperformed similar peers in schools that used Istation less by an average of 8 ($p<.01$) points. As an indicator of

magnitude, for students at the 50th percentile, this difference would equate to a gain of approximately 4-5 percentile points on the ISAT.[12] Additionally, students in schools where third graders averaged about 2 hours using Istation outperformed similar peers in schools that used Istation less by an average of 6 points (p<.05). Table 5 also reveals that more time spent using Istation curriculum was not associated with improved performance in reading on the ISAT. In these analyses, Istation usage was calculated for third grade students in each school only, so these usage indicators may have more accurately captured usage for the sample of students in the outcomes analyses, compared with analyses employing schoolwide usage metrics. On the other hand, it is also possible that these usage indicators were based on the usage of relatively few students in the school and were therefore not good indicators of grade-level usage.

Table 5: Associations between third grade Istation usage and improved student performance on spring 2019 ISAT ELA

Usage indicator	Usage amount	Average ISAT score
Progress monitoring + curriculum		*No comparison group*
Average sessions		
Low	<=7	2426
Mid	8-9	2434**
Mid-high	10-19	2426
High	>19	2429
Average min		
Low	<=109	2426
Mid	110-131	2432*
Mid-high	132-340	2428
High	>340	2428
Average curriculum min		
No curriculum usage	0	2429
Low	near 0	2427
Mid	1-117	2431
Mid-high	118-652	2428
High	>652	2428

NOTES—1. *p<.05, **p<.01. 2. The statistical significance refers to the difference in average student scores relative to schools with the lowest quartile of Istation usage. 3. The model also controlled for student grade level, English learner status, special education status, low-income status, race/ethnicity, and prior achievement (fall 2018 ISIP), as well as LEA enrollment, percentage of teachers at the school with 6+ years of experience, charter school status, school mean prior achievement, percentages of

---

[12] See Link to 2019 ISAT percentile ranks

students in the school at various grade levels, and percentages of students in the school who were English learners, low-income, or special education.

Taken together, the positive associations between indicators of Istation usage and improved student performance on two state assessments in spring 2019 suggest benefit to students when schools use Istation for more than the minimal levels. There also appeared to be diminishing returns to Istation usage, where students in schools with the highest levels of Istation usage did not perform substantially better (on average) than students in schools with moderate amounts of Istation usage.

Note that even the highest levels of usage on the various measures still reflect relatively small amounts of Istation testing and curriculum exposure a year relative to overall instructional time assumed to be allocated to ELA. Therefore, the higher usage schools (those beyond the lowest users) may be integrating Istation with regular instruction in generally beneficial ways. In addition, given the available data, the present study cannot rule out alternate explanations that may account for diminishing returns on usage. For example, it could be the case that schools in the different usage categories differed in systematic ways, or that school usage metrics were calculated based on high usage of few students, and any apparent diminishing returns on usage may be better explained by differences in school or student characteristics across usage categories. However, given the relatively large school sample included in the 2018–19 statewide analyses, the positive association between moderate Istation usage and improved student performance is suggestive of benefits for reading achievement.

Summary

- For schools in the statewide sample in the 2018–19 school year, about 64% of students in the typical school were exposed to Istation, for either progress monitoring and/or curriculum.[13] For schools where at least some students used the Istation curriculum, only 9% of the students in the typical school used the curriculum. Depending on grade level, the typical school used Istation for 9–11 sessions over the course of approximately 9 weeks with slightly higher average usage for students in the first grade compared with other grades.
- There were positive associations between increased Istation usage and improved student performance on the spring 2019 IRI (ISIP) and ISAT, relative to schools with the lowest levels of Istation usage. Higher average number of sessions, number of weeks, and total minutes of use for either progress monitoring and/or curriculum purposes were each positively associated with improved student performance, relative to schools with the lowest quartile of Istation usage.
- The percentage of students in the school who used Istation (for progress monitoring and/or curriculum) was also related to improved student performance

---

[13] This usage metric was determined at the school level because student-level usage data for progress monitoring were not available; therefore, usage may be underestimated for students in grades K-3 only.

on the spring 2019 IRI (ISIP), relative to schools with the lowest level of usage. Students in schools where the majority used Istation outperformed students in schools where a smaller percentage of students used Istation.

- Students in schools that used Istation curriculum had higher average student performance compared with students in schools that did not use Istation curriculum. This finding implies that in this first year of the statewide ISIP usage as the IRI, there may have been some benefit to student performance when schools also implemented the Istation curriculum. The average time students spent using the curriculum was not related to improved student performance, however.

Predictive validity of ISIP

A key question in employing progress monitoring or formative assessments in schools is the degree to which the resultant scores are predictive of performance on summative (i.e., state-mandated) assessments for which schools and school districts are held accountable. In this regard, our findings indicate that students' ISIP scores in grades 2 and 3 were good predictors of third grade students' scores on the ISAT ELA assessment. Table 6 shows the standardized regression coefficients that proximate the correlation between students' ISIP scores and students' scores on the ISAT ELA in the third grade.[14] Students' ISIP scores in both the second or third grade were related to students' ISAT ELA scores in the third grade by a pseudo correlation of at least .70. Moreover, this was true for ISIP scores from the fall, winter, and spring administrations.

Table 6: Validity of ISIP in predicting students' ISAT ELA scores

ISIP test	ISAT ELA sample	Standardized regression coefficient	Student N	School N
Fall 2017 in 2nd grade	Spring 2018 in 3rd grade	0.70	3,262	113
Winter 2018 in 2nd grade	Spring 2018 in 3rd grade	0.71	3,514	129
Spring 2018 in 2nd grade	Spring 2018 in 3rd grade	0.71	4,596	178
Fall 2018 in 3rd grade	Spring 2019 in 3rd grade	0.73	21,951	396
Winter 2019 in 3rd grade	Spring 2019 in 3rd grade	0.74	19,679	370

---

[14] These standardized regression coefficients were obtained from hierarchical linear models and account for the clustering of students within schools. No other covariates were included in the model. Standardized regression coefficients were obtained using Stata's (version 16.0) "beta" command. These estimates were very similar to estimates obtained from conducting simple pairwise correlations among the test scores at the student level.

Another way to examine the validity of ISIP in predicting student performance on the ISAT ELA is to examine how the ISIP tiers corresponded to the ISAT ELA performance levels, as shown in Figure 4. Note that this analysis is descriptive only, but it provides insight into how ISIP tier levels mapped onto ISAT ELA performance levels. ISIP tiers are similar to response-to-intervention tiers, where the majority of students fall into Tier 1, students who need a little additional support fall into Tier 2, and students who need a lot of additional support fall into Tier 3. The ISAT ELA categorizes student performance into four performance levels, ranging from 1 (lowest) to 4 (highest).

For second graders, the majority of students who were Tier 1 on the ISIP spring assessment scored at either Level 3 or 4 on the ISAT ELA when in the third grade. The majority of second graders who were Tier 2 on the ISIP spring assessment scored at either Level 1 or 2 on the ISAT ELA when in the third grade. Finally, the majority of second graders who were Tier 3 on the ISIP spring assessment scored at Level 1 on the ISAT ELA when in the third grade. Findings were nearly identical for how second and third graders' fall and winter ISIP scores mapped onto their ISAT ELA scores. These data indicate the general overlap between the ISIP tiers and ISAT ELA levels, but given the variability of students included in each tier and level, these mappings are not precise.

Figure 4: Mapping of ISIP tiers onto ISAT ELA performance levels
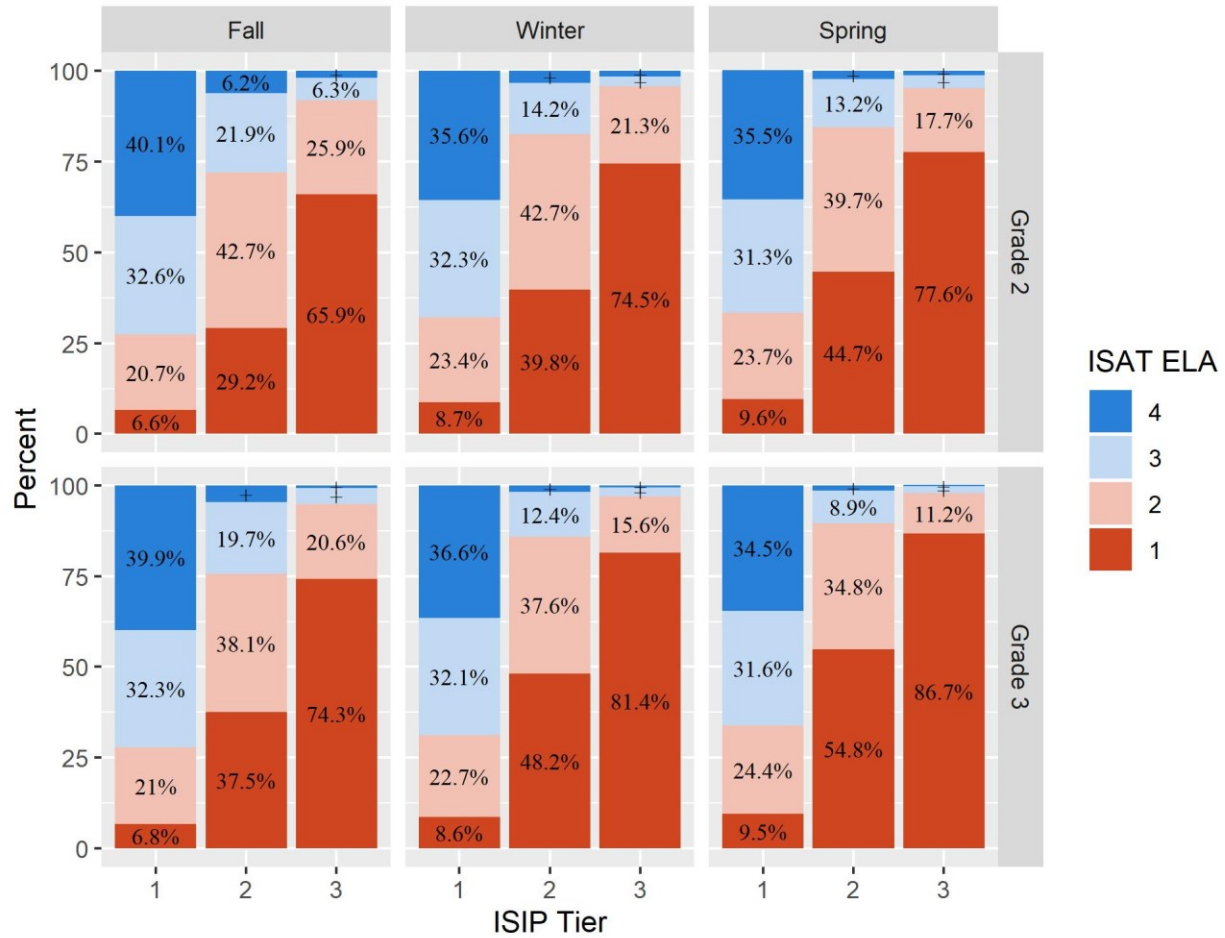
Summary

- Second and third grade student scores from the fall, winter, and spring administrations of the ISIP were all good predictors of student scores on the ISAT ELA, which is administered in the spring of students' third grade year.
- The approximate correlation between second and third grade students' ISIP scores and students' scores on the ISAT ELA was around .70.

Pilot school usage and achievement outcomes

As described in the introductory section of this report, a QED study was performed to explore whether students in schools that piloted the Istation ER program in 2017-18 outperformed students in similar, comparison schools on assessments in the springs of 2018 and 2019. Sampling equivalence of the two groups was demonstrated by analyses indicating similar baseline achievement and student demographics (except for a small difference in economically disadvantaged percentages). We first present usage data to reflect levels of Istation ER implementation by the pilot schools and accordingly, to inform the interpretation of student achievement outcomes.

Istation usage for pilot schools. Our analysis of usage data indicated that not all students in pilot schools used Istation progress monitoring or curriculum tools. The following figures display the extent of Istation use in the pilot school sample. It may be recalled that this sample included pilot schools that opted to use Istation before the ISIP became the state-mandated assessment for students in grades K–3,[15] as well as similar, comparison schools.

Figure 5 shows the medians and interquartile ranges for the percentages of students enrolled in each school who used Istation, either for progress monitoring or curriculum purposes. As shown, the median school percentage of students who had used Istation for any reason out of the school's total enrollment was 77%. In other words, in the majority of Istation pilot schools, about three-quarters of students had some exposure to Istation. Figure 5 also shows the percentages of students enrolled in each school who used the Istation curriculum, as well as the percentages of students enrolled in each school who used the Istation curriculum at the recommended levels. For schools where at least some students used the Istation curriculum[16], the median school percentage of students who used Istation's curriculum tools was 68%, and the median school percentage of students who used Istation tools at the recommended levels was 13%. Therefore, in pilot schools that used Istation curriculum, about half of students had participated in the curriculum in the majority of schools, while low percentages of students used the Istation curriculum at the recommended levels.

Figure 5

Medians and Interquartile ranges for the percentage of Istation users in the school in the 2017 – 18 school year, by type of user

---

[15] The sample was restricted to the pilot schools selected for the main analysis, as outlined in Appendix A.
[16] This was determined if the percentage of students in the school who had used the curriculum was greater than 0.
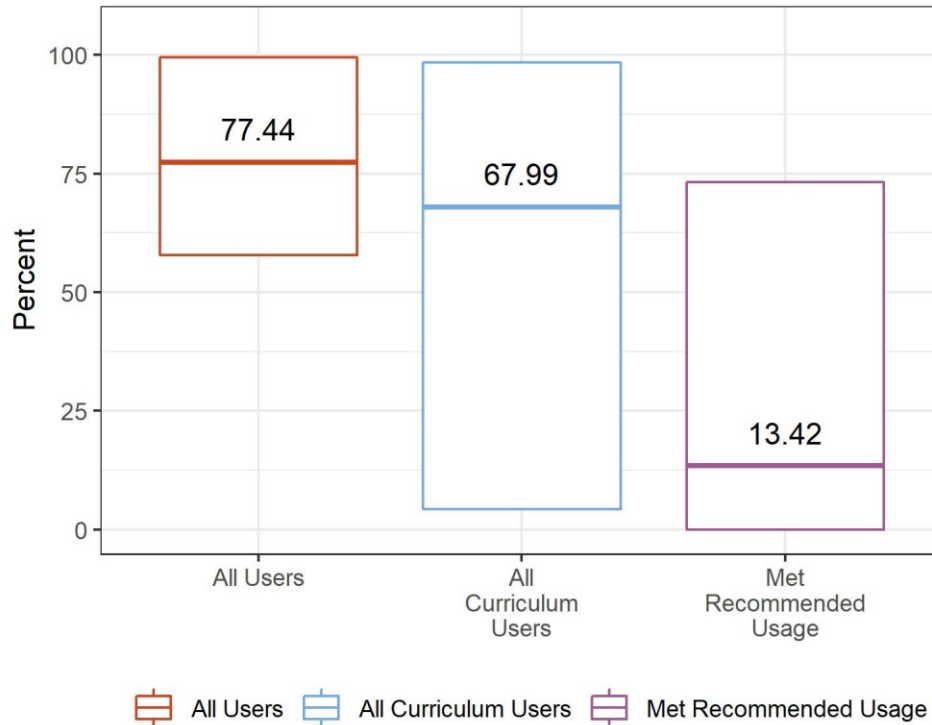
NOTE—Only schools with some Istation curriculum users were included in the estimates of the percentages of students in the school who used the curriculum, but nearly all schools that piloted Istation during the 2017–18 school year used the Istation curriculum at least to some extent.

When examining the average number of Istation sessions and weeks of usage, the median pilot school used Istation for 8–13 sessions (e.g., number of times students logged in) over the course of 8–10 weeks, depending on grade level. The numbers of sessions and weeks of usage were similar across grades, with slightly higher usage levels in the first grade than in other grades. Additionally, in about 25% of schools piloting Istation during the 2017–18 school year, students in grades 1–3 used Istation on average for approximately 40+ sessions over the course of 20+ weeks. Figure 6 shows the medians and interquartile ranges of the number of sessions and weeks of Istation usage by grade.

Figure 6: Medians and interquartile ranges for the number of sessions and weeks by grade in the 2017–18 school year
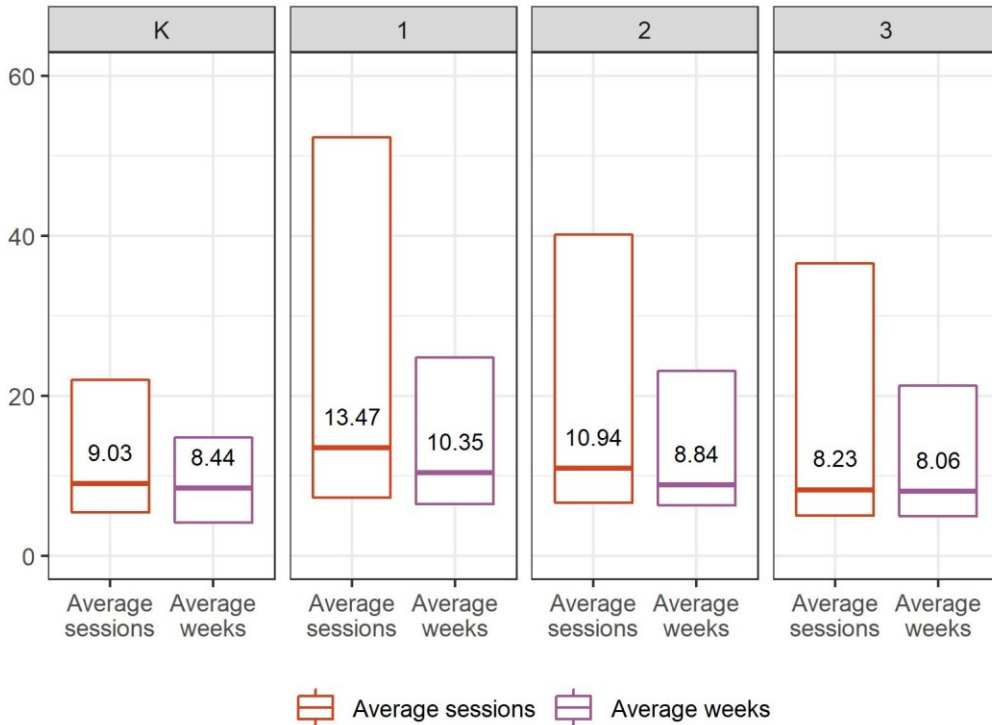
In terms of total time spent using Istation, students in the typical pilot school used Istation for at least an average of 2.5–5 hours total, with students in the first grade averaging the most time using Istation. When pilot schools used the curriculum as well, the time students in the typical Istation pilot school spent using the curriculum ranged from an average of 2–9 hours per student, with more time spent by students in grades 1–3 compared with students in kindergarten. Figure 7 shows the medians and interquartile ranges for total and curriculum minutes of Istation usage by grade.

Figure 7: Medians and interquartile ranges for the total and curriculum minutes of Istation usage by grade in the 2017–18 school year
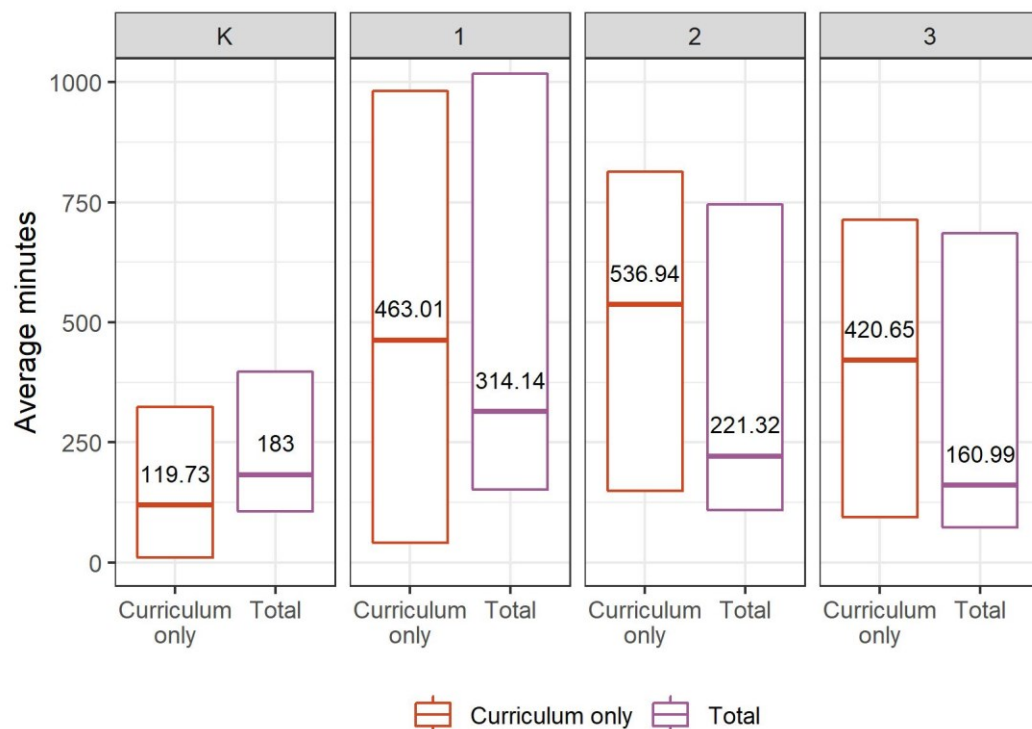
NOTE— Only schools with non-zero curricular usage were included in the estimates of time spent on Istation curriculum only, but nearly all schools that piloted Istation during the 2017–18 school year used the Istation curriculum at least to some extent.

Pilot school comparative achievement outcomes. The above usage data reveal tangible efforts by the pilot schools in general to introduce both diagnostic testing and the curriculum in advance of the statewide implementation of Istation in 2018-19. Although the limited amount of exposure to Istation by the typical student (only 2.5 to 5 total hours during the school year) represents the realistic first-year implementation for the participating pilot schools, it unfortunately precludes a strong test of program efficacy on end-of-year summative assessments, such as the IRI or ISAT. This constraint notwithstanding, it is still informative to explore possible benefits for students in the pilot schools. It could also be the case that Istation implementation benefitted certain types of reading skills more than others. We were able to explore this question through analyses of IRI and ISAT subtests.

In initial analyses, we examined the differences between the pilot and comparison school samples in reading on the IRI or ISAT ELA in the springs of 2018 and 2019, after controlling for prior student achievement and a host of other background variables. As shown in Table7, the average score for comparison students was 2.58 on the previous IRI, and the average score for Istation students was virtually

identical. On the ISIP administered in spring 2019, the average score for comparison students was 240.45, and the average difference for Istation students was +0.55 with a standard error of 0.52. Neither of these differences were statistically significant. Similarly, there were no statistically significant differences in student performance between Istation and comparison students on the ISAT ELA in either spring 2018 or 2019. Because the average differences in student performance between Istation and comparison students were very small, the standardized effect sizes (the average difference for Istation students in terms of standard deviations) were also very small and close to zero.

Table 7: Overall outcomes for pilot relative to comparison schools

Year	Outcome	Average score for comparison students	Average pilot student difference	Standardized effect size[17]
Spring 2018	Previous IRI Composite	2.58	0.00 (0.02)	0.00
Spring 2018	ISAT ELA	2,430.90	-0.10 (3.75)	0.02
Spring 2019	New IRI (ISIP)	240.45	0.55 (0.52)	0.00
Spring 2019	ISAT ELA	2,453.13	0.48 (3.30)	0.01

NOTES—1. The overall pilot vs. comparison school effect was not statistically significant in any analysis. 2. The numbers in parentheses are the standard errors. 3. The standardized effect size is in terms of standard deviations. 4. The model also controlled for student grade level, English learner status, special education status, low-income status, race/ethnicity, and fall 2017 IRI score, as well as LEA enrollment, percentage of teachers at the school with 6+ years of experience, charter school status, school mean prior achievement, percentages of students in the school at various grade levels, and percentages of students in the school who were English learners, low-income, or special education.

We also explored whether there appeared to be differences between students in pilot and comparison schools on particular subtests of state assessments. For example, the IRI contained three subtests (letter naming fluency, letter sound fluency, and reading curriculum-based measure), the ISAT contained four subtests (reading, writing, listening, and research), and the ISIP contained numerous subtests. While these analyses explore possible trends associated with the pilot initiative on student knowledge and skills in a particular area, we caution that statistically significant findings may also arise due to chance alone when conducting a large number of analyses.

---

[17] The standardized effect size was calculated by the pilot effect divided by the pooled unadjusted standard deviation of the outcome measure for pilot and comparison students per What Works Clearinghouse Standards Handbook Version 4.0.

As shown in Table 8, pilot students outperformed similar comparison peers by an average of 2.05 points (p<.05) on the letter sound fluency subtest of the spring 2018 IRI, when controlling for letter sound fluency as of fall 2017, as well as other student and school characteristics.[18] On the spring 2019 IRI (ISIP), pilot students outperformed similar comparison peers by an average of 3.60 points (p<.01) on the text fluency subtest and by an average of 2.05 points (p<.01) on the vocabulary subtest, when controlling for fall 2017 IRI composite score. For all other subtests of the IRI and ISAT ELA, there were no statistical differences in average student performance for pilot students and similar comparison peers. In summary, relative to comparison students, pilot students appeared to have made larger gains in letter sound fluency in the 2017–18 school year, as well as larger gains in text fluency and vocabulary in the 2018–19 school year. Possibly this outcome reflects Istation usage being more focused on these types of skills than on others. However, we cannot rule out all other factors that may have influenced these results.

Table 8: Effects of piloting relative to comparison schools on subtests

Outcome	Subtest	Average for comparison	Average piloting difference	ES
IRI Spring 2018	Letter sound	50.60	2.05* (1.00)	0.10
	Letter naming	39.52	0.41 (1.12)	0.02
	Reading curriculum	107.36	-1.94 (3.25)	-0.04
ISAT ELA Spring 2018	Writing	2421.60	2.33 (4.48)	0.02
	Research	2414.44	0.27 (5.12)	0.00
	Reading	2435.33	-2.84 (3.67)	-0.03
	Listening	2446.04	-3.19 (4.47)	-0.03
IRI (ISIP) Spring 2019	Text fluency	56.67	3.60** (1.21)	0.09
	Vocabulary	247.92	2.05** (0.79)	0.07
	Listening comprehension[19]	209.01	5.00 (3.64)	0.27
	Comprehension	244.18	0.84 (0.62)	0.03
	Alphabetic decoding	223.53	0.27 (0.75)	0.01
	Spelling	235.88	0.04 (0.52)	0.00
	Phonemic awareness	206.44	-0.61 (0.96)	-0.04
	Letter knowledge	203.83	-0.88 (1.33)	-0.05
ISAT ELA Spring 2019	Writing	2447.34	1.00 (4.29)	0.01
	Reading	2457.07	0.99 (3.22)	0.01
	Research	2432.85	-1.00 (4.07)	-0.01

[18] These analyses otherwise employ the same hierarchical linear model used to estimate the overall results in the previous section.

[19] The ISIP listening comprehension subtest was administered only to kindergarten students in the 2018–19 school year. As a result, this analysis had a small sample size. The small sample size, along with relatively large standard error of the pilot school difference, were the reasons that the average difference of 5.0 points between pilot and comparison students on this measure was not statistically significant.

Outcome	Subtest	Average for comparison	Average piloting difference	ES
	Listening	2469.11	-2.02 (4.08)	-0.02

NOTES—1. *p<.05, **p<.01. 2. The numbers in parentheses are the standard errors. 3. The ES (effect size) is in terms of standard deviations. 4. The model also controlled for student grade level, English learner status, special education status, low-income status, race/ethnicity, and prior achievement, as well as LEA enrollment, percentage of teachers at the school with 6+ years of experience, charter school status, school mean prior achievement, percentages of students in the school at various grade levels, and percentages of students in the school who were English learners, low-income, or special education.

Outcomes for pilot schools with different Istation usage levels. The next set of analyses allowed us to investigate the associations between the extent of Istation usage (progress monitoring and/or curriculum) and improved student performance in reading, relative to students in comparison schools. As previously described, we first categorized usage by quartiles as follows:

- Low usage: Usage less than or equal to the 25th percentile
- Mid usage: Usage greater than the 25th percentile and less than or equal to the 50th percentile
- Mid-high usage: Usage greater than the 50th percentile and less than or equal to the 75th percentile
- High usage: Usage greater than the 75th percentile

These quartiles were estimated using usage data from pilot schools only. In addition, the usage indicators that specifically captured usage of Istation curricular tools were estimated only for pilot schools with non-zero curriculum usage.

The results showed isolated significant differences between pilot and comparison students on both the IRI and ISAT ELA in spring 2018, when controlling for prior achievement and student and school characteristics.[20] All statistically significant effects favored the pilot schools in relatively high usage categories (see Table 9). Specifically, for the approximately 25% of pilot schools in which between 78–99% of the student population was estimated to have participated in Istation, either in progress monitoring or curriculum, pilot students outperformed their comparison peers on the spring 2018 IRI by an average of 0.06 points ($p<.05$). Similarly, for the approximately 21% of pilot schools in which between 69–98% of the student population used Istation curriculum, pilot students outperformed their comparison peers on the spring 2018 IRI by an average of 0.08 points ($p<.05$). Finally, students who used Istation curriculum for an average of 10.7 hours or more outperformed their comparison peers by an average of 0.11 points ($p<.05$). Given that the IRI composite score ranged from only 1 to 3, these effects are practically meaningful. Full regression results are available in Appendix B.

---

[20] The model was the same as the one used to estimate the overall results. The only difference was that the pilot indicator (pilot versus comparison) was replaced with the categorical usage indicators to test differences in student performance for pilot schools with various levels of usage, relative to the comparison group.

Table 9: Effects of schoolwide Istation usage on student spring 2018 IRI scores relative to comparison schools

Usage indicator	Usage amount	Average IRI score
Progress monitoring + curriculum		*Comparison average 2.56*
Average sessions		
Low	<=6.5	2.57
Mid	6.6-10.5	2.55
Mid-high	10.6-35.5	2.60
High	>35.5	2.60
Average weeks		
Low	<=6	2.59
Mid	6.1-8.5	2.55
Mid-high	8.6-19.5	2.58
High	>19.5	2.61
Average min		
Low	<=117	2.57
Mid	118-230	2.54
Mid-high	231-706	2.60
High	>706	2.61
Percent users		
Low	<=58%	2.58
Mid	59-77%	2.56
Mid-high	78-99%	2.63*
High	>99%	2.55
Curriculum only		
Average curriculum min		
No curriculum usage	0	2.50
Low	<=27	2.61
Mid	28-310	2.60
Mid-high	311-643	2.64
High	>643	2.68*
Percent curriculum users		
No curriculum users	0%	2.51
Low	<=4%	2.62
Mid	5-68%	2.60
Mid-high	69-98%	2.65*
High	>98%	2.65

NOTES—1. *$p<.05$. 2. The model also controlled for student grade level, English learner status, special education status, low-income status, race/ethnicity, and prior achievement, as well as LEA enrollment, percentage of teachers at the school with 6+ years of experience, charter school status, school mean

prior achievement, percentages of students in the school at various grade levels, and percentages of students in the school who were English learners, low-income, or special education.

While there were few statistically significant differences in the aforementioned analyses, Figure 8 generally depicts a positive directional trend between greater Istation usage and increased student performance on the spring 2018 IRI. However, this study cannot rule out other factors that may be confounded with greater Istation usage and improved student performance, such as higher principal and teacher quality.

Figure 8: Average student spring 2018 IRI scores at various usage levels relative to comparison schools
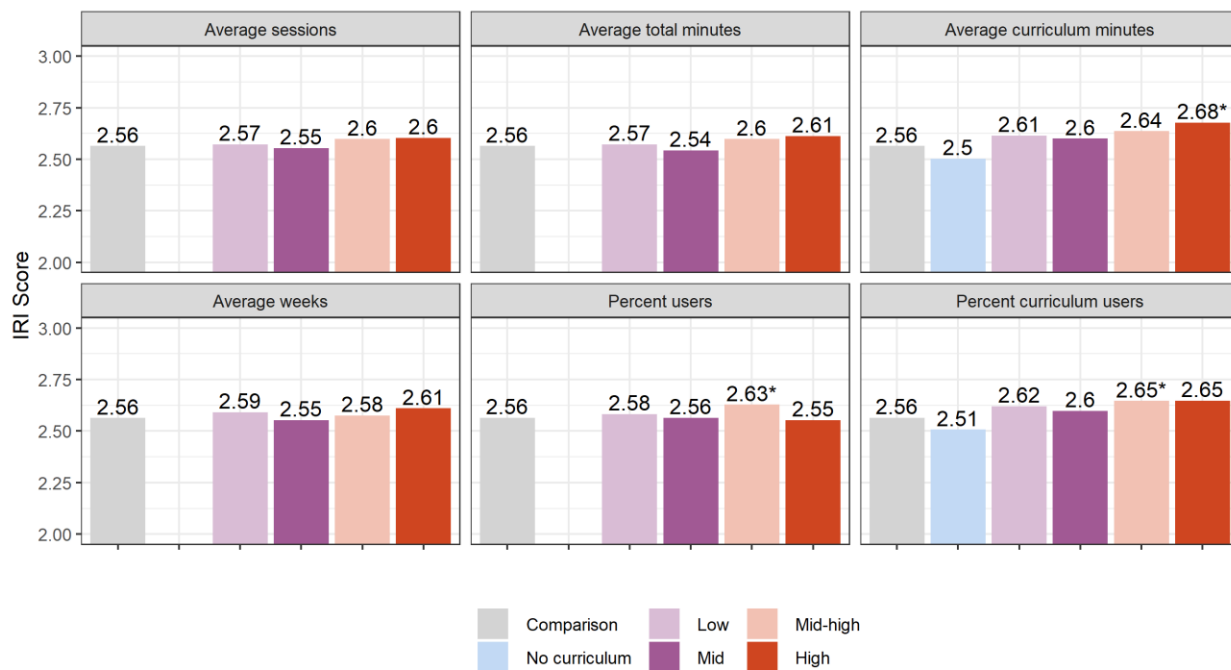
NOTES—1. *p<.05. 2. The statistical significance refers to the difference in average student scores relative to the comparison group.

Analyses of the relationship between increased Istation usage and scores on the ISAT ELA in spring 2018 yielded similar results. Although there were no statistically significant effects directional trends favored pilot schools with increased Istation usage over the comparison schools. Specifically, as shown in Figure 9, third grade students scored directionally (but not statistically significantly) higher on the ISAT ELA in schools with greater Istation usage specifically at the third grade level. Table 10 provides the associations between improved student performance and Istation usage levels, relative to the average student performance for comparison students. In terms of interpreting the magnitude of the effect, the percentile for an ISAT score of 2432-2433 (see score for subgroups with middle to high total minutes) is 52, while the percentile for an ISAT

score of 2428 for comparison students is 50.[21] This is a difference of about 2 percentile points for the average higher user relative to the average comparison student.

Table 10: Effects of third grade Istation usage on student spring 2018 ISAT ELA scores relative to comparison schools

Usage indicator	Usage amount	Average ISAT score
Progress monitoring + curriculum		*Comparison average 2428*
Average sessions		
Low	<=5	2421
Mid	6-8	2427
Mid-high	9-40	2430
High	>40	2430
Average min		
Low	<=73	2421
Mid	74-161	2428
Mid-high	162-706	2428
High	>706	2432
Average curriculum min		
No curriculum usage	0	2427
Low	<=78	2421
Mid	79-421	2429
Mid-high	422-727	2432
High	>727	2433

NOTES—1. *p<.05. 2. The model also controlled for student grade level, English learner status, special education status, low-income status, race/ethnicity, and prior achievement, as well as LEA enrollment, percentage of teachers at the school with 6+ years of experience, charter school status, school mean prior achievement, percentages of students in the school at various grade levels, and percentages of students in the school who were English learners, low-income, or special education.

---

[21] See Link to ISAT 2018 percentile ranks

Figure 9: Average student spring 2018 ISAT ELA scores at various usage levels for the third grade relative to comparison schools
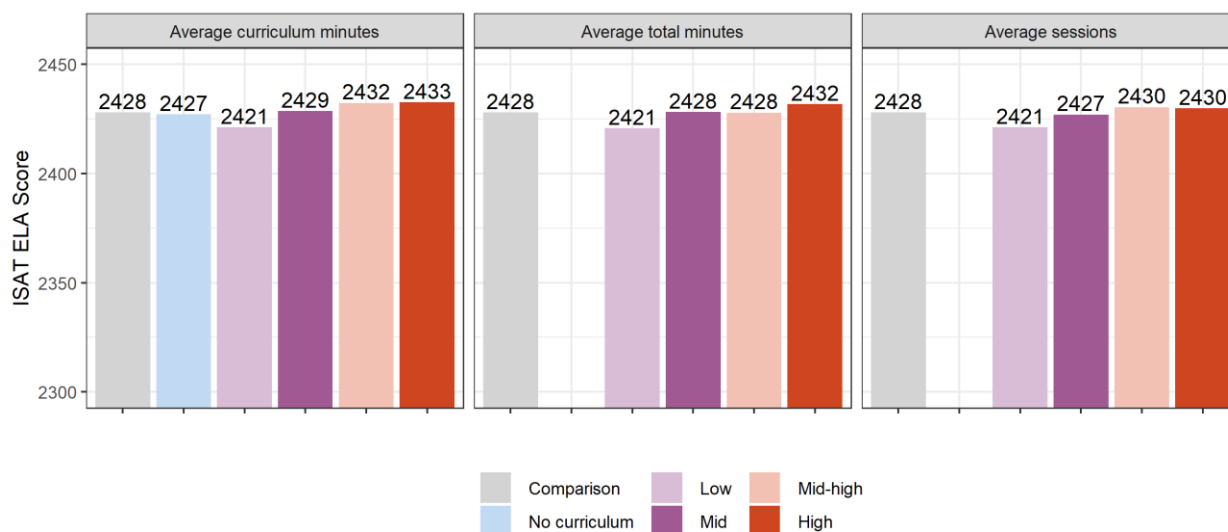
NOTES—1. There were no statistically significant differences in average student performance between Istation students with any usage level and comparison students. 2. These usage indicators were estimated based on data for third grade students only.

In summary, while there were only isolated statistically significant effects in the usage analyses, findings indicate that greater Istation usage was associated with improved student performance in some instances. More specifically, pilot students outperformed similar peers on the spring 2018 IRI when the majority of students in the school used Istation, either for progress monitoring and/or curriculum purposes. This finding suggests that a schoolwide implementation of Istation may yield greater benefits than a more piecemeal approach, but more research is needed to disentangle the effects of the various components of Istation's Early Reading program, particularly given that overall usage levels, even for schools in the highest usage categories, were relatively low relative to total instructional time presumably allocated for reading and ELA during the school year.

Findings should be interpreted with caution, however, because this study cannot rule out other, unobservable differences among the school and students in the various Istation usage categories. Additionally, due to the large number of statistical analyses conducted, some of these statistically significant findings may be the result of chance alone. Nevertheless, these findings point to possible benefits in increased student performance relative to the comparison group when a substantial proportion of students in the school participated in Istation, for progress monitoring and/or curriculum.

Summary

- Pilot schools made tangible efforts to implement Istation progress monitoring and to a lesser extent, the curriculum program. For schools that piloted Istation

during the 2017–18 school year, about three-quarters of students in the typical school used Istation, for either progress monitoring and/or curriculum. For schools where at least some students used the Istation curriculum, about half of the students in the typical school used the curriculum. Depending on grade level, the typical pilot school used Istation for 8–13 sessions over the course of 8–10 weeks. The numbers of sessions and weeks of usage were similar across grades K–3, with slightly higher usage levels in the first grade than in other grades.

- These implementation efforts notwithstanding, in this initial year of program experiences, overall usage was relatively limited. The typical student in pilot schools received only from 2.5-5 hours of total exposure to Istation products, and when schools also used the Istation curriculum, the typical student received only 2–9 hours of exposure.
- Taking the modest usage levels into account, the overall findings from student achievement analyses revealed some positive trends for the pilot schools. However, in the overall analyses of IRI and ISAT scores, the difference in average student performance between pilot and similar comparison schools was not statistically significant.
- Pilot students outperformed similar peers on the spring 2018 IRI when the vast majority of students in the school used Istation, for either progress monitoring and/or curriculum purposes.
- Additionally, in subtest analyses, students in pilot schools had higher average growth on the spring 2018 IRI letter sound fluency subtest and on the vocabulary and text fluency subtests of the spring 2019 IRI (ISIP) relative to comparison students. Importantly, students in Istation schools did not underperform students in comparison schools on any IRI or ISAT subtest.

Educator perceptions of Istation

In the final section of this report, we discuss findings from interviews and focus groups with educators whose schools had piloted Istation during the 2017–18 school year. We interviewed 21 educators from eight schools in Idaho in spring 2019. The schools were selected by the Idaho State Department of Education based on their geographic diversity and relatively high student growth on the ISIP from fall to spring of the 2017–18 school year. Table 11 shows a summary of the interviewees. Interview and focus group data were analyzed via standard qualitative techniques to identify emerging trends and themes across participant responses.

Table 11: Educator interview sample

	N	% of Sample
Principals	7	33.3%
Instructional coaches	2	9.5%
Teachers	12	57.1%
Total	21	100%

Educators were asked questions about the following topics:

- Support for implementation of Istation:
    - Initial supports
    - Ongoing supports
- How educators used various components of Istation:
    - Progress monitoring
    - Data and data reports
    - On-demand (e.g., online) assessments for students
    - Teacher-directed lessons
- Perceived impacts on students:
    - Student achievement
    - Student engagement
- Perceived impacts on teacher practice
- Program feedback:
    - Suggestions for improvement
    - Challenges to implementation

Findings from the interviews are organized according to these topics.

Support for implementation. We asked the educators to describe the types of support they received for the initial implementation of Istation. With the exception of one school, most schools sent 1–2 delegates, typically an instructional coach and/or teachers, to attend the Istation training and then tasked those individuals with training their colleagues upon return. Educators reported that the initial training mainly focused on the technical navigation and usability of the program and its website, with one teacher noting:

> *The Istation training was really helpful. We learned how to use the program, how to do on-demand, how to use the resources on the website, ways to manipulate the data…all of that was very helpful.*

This training information was also available via online modules, which were well-received for the purposes of familiarizing users with Istation. Some educators reported, however, that they would have preferred to attend the trainings themselves as opposed to learning from their colleagues. One person commented:

> *[This was] not the best model. It would have been better if we all went, heard the same thing at the same time. We've all played grapevine with that information.*

Another teacher reported that while the online modules provided sufficient supports for some teachers, other teachers were reluctant to use them and would have benefitted from more "hands-on" support where "they could ask questions."

According to respondents, the initial training was not sufficient for taking full advantage of Istation and its features. For instance, many commented on how overwhelming the program seemed at first. Educators were not fully aware of how to generate or interpret the data reports until they had sufficient experience with the program and had attended multiple training sessions throughout the school year and summer.

In general, educators found that trainings at their school where the principal encouraged a hands-on, collaborative approach were most conducive to learning how to use the program. Teachers predominantly found the in-house training provided by Istation consultants to be most useful because, by this time, teachers had direct experience with the program and had developed specific questions that Istation consultants could answer best. While colleagues were a good resource, many educators still valued the support provided by Istation. According to one teacher:

It was sink or swim until I attended the mid-year training from Istation.

Educators reiterated the importance of the Istation consultant, who offered continuous support and whose school visits were highly anticipated.

Respondents also elaborated on the type of professional development that would be helpful now, after having used Istation for an extended period of time. Overwhelmingly, respondents highly valued the data they had the knowledge to access and pointed to the need for more training on how to better understand and interpret the data reports that are generated by Istation. Teachers expressed an interest in learning more "tips and tricks" for efficient ways to generate data reports.

Several respondents also expressed an interest in knowing more about how the Istation data points on students were calculated. For instance, a respondent queried why two students who scored the same had different tier levels. A principal also noted:

It would have been great to get more information on how they calculate when students are ready to move from one category to the next. How does that work? Right now, it's a mystery.

Educators were genuinely interested in gaining a clearer understanding of how students were scored.

Respondents also wanted to understand how the Istation assessment correlated with other assessments administered to students. For example, one teacher commented

that her school had opted not to use the Istation assessment with upper elementary students because they were unsure how scores would correlate with scores on the Smarter Balanced assessment.

Principals stressed the need for ongoing, "refresher" professional development opportunities. Other calls for professional development included providing strategies for designing instruction based on assessment data and facilitating more networking opportunities through joint professional development sessions with other Istation schools.

Summary

- The initial training provided useful guidance on how to navigate and use Istation, but it was provided to a limited number of educators, and most educators gained the information secondhand from their colleagues.
- Educators found the in-house training provided by Istation consultants to be very effective and useful.
- Many educators expressed interest in additional professional development on Istation. Specifically, educators were interested in learning how to better use the data and reports. Some educators wanted professional development opportunities that go beyond how to use the program and extend into how the assessment works.

How educators used various components of Istation. Educators described the ways in which they used various features of Istation.

Progress monitoring. Educators indicated that they administered the Istation assessment to their students about once a month throughout the school year. In some cases, assessments were administered more frequently at the beginning of the year. After each assessment, educators would typically meet in small groups to review student progress and plan instruction and interventions for students. Meetings were typically held with grade-level teams, and some teams met monthly whereas other teams met more frequently. In these meetings, principals and teachers often would examine the data first at the grade level, then at the class level, followed by the tier and individual student level. These discussions informed teachers' decisions to assign or adjust interventions for students and prescribe on-demand assessments. On-demand assessments were online assignments that teachers could assign as independent work to individual students to target specific skill gaps. One teacher described the progress monitoring routine that was typical for most teachers:

> *Every month, we do the assessment and take that data and adjust groupings for small groups.*

Another teacher described the process as follows:

We've been using Istation in 1ˢᵗ grade to determine the effectiveness of our interventions. I'll look at changes in growth for this.

Most respondents indicated that teachers logged into Istation once per week, on average. Some schools provided general guidelines for usage, whereas others were more hands-off in their approach to implementation. In general, teachers developed their own routines based on whether or not there were specific administrative guidelines for using Istation. One teacher described the school's approach:

We are required to have an hour a week in instruction. Computer lab time [for Istation] was thirty minutes, two times per week.

Teacher opinions about Istation appeared to be related to the extent to which teachers used the program, with teachers using the program to a greater extent often expressing more positive opinions.

Data and data reports. Of the available data reports generated by Istation, respondents expressed that they found the class and student summary reports the most beneficial. Student summaries, which detail individual progress, were especially helpful for conferencing with parents but also provided insight into the type of instructional intervention that was needed for a particular student. Similarly, class summaries were useful for teacher self-evaluation and for monitoring class progress over time. The class summaries provided a broader overview of student performance than the individual student reports. Teachers also liked the skill growth summary. While the class and student summaries provided a wealth of data, the skill growth summary gave a more targeted view of individual student progress, which was highly valued by teachers.

Beyond the data reports mentioned above, teachers also reported using priority, Lexile trend, and tier movement reports. As summarized by one respondent:

Everyone in our building likes to look at [data] a different way.

Several teachers commented that they believed they were only accessing a portion of the information that was available to them and that they would benefit from additional training on how to extract and interpret all available reports. Teachers also mentioned that they would like to be able to do, or know how to do, the following with reports:

- Track individual student growth by viewing a current score along with a score from the beginning of the month.
- Access a graph of classroom-level scores depicting the initial scores and on-demand scores for the month.
- Allow their students to view and track their own progress.

- Remove student ranking from reports shared with parents. Some teachers found this data point was a distractor for parents.
- View monthly progress in the summary report.
- View all months on a single page or screen in the skill growth report.

On-demand assessments. Not all schools utilized the on-demand assessments regularly. This feature was used mostly for individual students in Tiers 2 and 3, or for students who performed below their benchmarks and needed additional intervention. As described by one teacher:

> *[Students] get monitored once a month. Anyone who is a Tier 2 or a 3, they go on-demand… When we meet about the data we have an intervention set-up that changes every eight weeks. Teachers meet every two weeks to look at the data in grade level groups.*

Another explained:

> *For our lowest achieving students, we are working on letter/sound and they are doing an on-demand every two weeks. So, we only use now if we think we really need it.*

Respondents implied that it was up to individual teachers to decide whether to use the on-demand assessment with low-performing students, as opposed to school-wide implementations.

Teacher-directed lessons. The teacher-directed lessons were used to varying degrees by different teachers, and in at least one case, the principal was entirely unaware that this feature existed. Teacher comments on this component of the program included:

> *It's a different approach to teaching so I was able to take that approach and use it within my core curriculum. My students enjoyed the manipulations. I have taken some of that and put it in my core instruction.*

> *I really like them. The teacher-directed lessons are the same as we've been trained to teach. [They] gave us the format for the test – very helpful.*

For others, this program feature was less beneficial. One teacher described the lesson piece as being "more hassle than it's worth" and instead opted to use core curriculum with which they were more familiar. The general sentiment towards the teacher-directed lessons was that they were useful for some teachers but unnecessary for others for whom the existing curriculum already addressed the core, grade-level content. One teacher commented:

The lessons look good, but we don't need them.

Another teacher elaborated that they were already using other intervention programs:

> *We are focused on using research-proven interventions especially for our intervention and response to intervention (RTI) process. We have the personnel to support these interventions: Phonics for Reading, Chipper Chat, and Zoophonics.*

It seemed as if teachers already had interventions at their disposal, as well as curricula that they knew and liked, and therefore did not spend as much time familiarizing themselves with the teacher-directed lessons.

Summary

- The Istation ISIP ER was administered to students in pilot schools about once a month throughout the school year. After each assessment, educators would typically meet as grade-level teams to review student progress and plan instruction and interventions for students.
- Teachers especially valued the class and student summary reports for a quick overview of student progress data. Teachers also liked the skill growth summary and used it to monitor individual student progress and determine whether on-demand assessment was warranted.
- On-demand assessments were used in targeted ways predominantly for low-performing students. Many teachers reported not using this feature.
- Teacher-directed lessons appeared to be under-utilized. The reason for this was that some teachers felt they already had a strong curriculum in place with which they were already familiar and trusted.

Perceived impacts on students. With regard to the impact of Istation on student achievement, most respondents were apprehensive to draw a straight line from Istation to learning outcomes, but many acknowledged how the program has made student progress more visible for both the teacher and student. Respondents indicated that Istation was more beneficial for low-performing than high-performing students because it provided foundational instruction. A teacher qualified this, saying:

> *My gifted students seem bored.*

Another teacher talked about advanced 3rd grade students "clicking through" the assessment because of their lack of interest, resulting in "wildly inaccurate" data.

Although the general consensus was that Istation was more beneficial for low-performing students, students with special needs—either special education students or

English learners (Els)—struggled most with the program. For instance, some teachers expressed concern that the program requires students to do too many things at one time, which inhibited these students' ability to properly understand and answer the question. One teacher explained:

> I do think that it is very challenging for special education and EL students... [Istation] tests their working memory and vocabulary, not phonemic awareness. They are asked which word starts with the same sound as something represented in four pictures. These students can't remember the words and don't have enough time to touch all the icons to hear the sounds. If I just ask them a question to test this, they will get it 100% of the time. If they could turn off the time portion of this, it would work for ELs. Having the option to hear the question again would also benefit these students.

For many respondents, however, it was just too early to determine the effects of Istation on academic achievement.

Many participants noted that by and large, student engagement has been positively impacted. Although the novelty of the program has worn off, students have remained engaged because they were invested in watching their progress over time. In particular, the "trend line" that tracks growth had a gamification-like quality, where students were motivated and encouraged by their individual growth charts. One teacher commented:

> It's made them more engaged in their performance. When they take the ISIP they love to see that their line is going up. When it drops, they feel SO bad. My students are really invested in that line and doing their best to make it go up. They are invested in their improvement and they can see if they're improving.

Another commented about the assessment:

> They are always eager to take it. They want to know their scores. They like to see themselves improving. The program is engaging and entertaining.

Teachers also noted that the activities and games were especially engaging for students.

Some respondents reported engagement issues, however. The primary criticisms here were that the assessments were too lengthy and placed an excessive amount of cognitive load on special education students, as described above. One principal indicated:

> *I don't think it has impacted [students]. In the beginning, it was a novelty, but they aren't as excited now. In K and 1st grade, 30 minutes is a long testing time (too long).*

Still, the majority response was that students were engaged while using Istation.

Perceived impacts on teacher practice. Positive impacts on teacher practice came across as a clear benefit of using Istation. Respondents emphatically agreed that Istation has positively impacted their practice. Istation provides teachers with a more holistic picture on student progress than they previously had. Istation data inform not just *how* students performed on assessment but *why* they did so. These data insights helped teachers identify gaps in student learning and customize instruction to address the gaps. Several comments from teachers included:

> *It's just a better, holistic assessment. It provides an overall picture of the student. As I become more familiar, I trust the results more because I understand them more.*

> *Before Istation it was really a guessing game. Now we know, we have targeted information letting us know.*

> *We wouldn't have identified these specific gaps [in learning] without Istation.*

> *It has benefitted teacher practice. Initially teachers were looking at Istation … as being outside the curriculum but now use to master curriculum.*

Teachers also indicated that using Istation has helped them better communicate with parents about their children's academic progress. Respondents noted that Istation conveniently provided them with current and objective data that they could discuss with parents during parent-teacher conferences. Many teachers reported using the student and skill growth summaries during conferences, in particular.

Summary

- Teachers perceived Istation to benefit low-performing students more so than high-performing students, with low-performing students seeing more noticeable gains in their learning progress. The exception here was that Istation was especially difficult for special education and EL students to access.
- While the initial novelty of Istation has worn off, student engagement levels still remained relatively high, and students were motivated to track their progress over time.

- Teachers asserted that Istation has impacted teacher practice in positive ways, equipping them with more information to make better instructional decisions and target gaps in their students' knowledge.

Program feedback. Educators were asked to describe what they liked best about Istation. Above all else, respondents liked the immediate feedback and robust data on student learning and areas of weaknesses. Teachers noted:

The wealth of information we get on each student is incredible – I'm a believer now. I see the value. I hope they will stay with [it]. It's been very valuable to us.

It really gives us a lot of target information – I love that. Before all we had was fluency. We get all of it in one test – I love that.

Teachers also recognized how Istation generates real-time data and allows for "instant feedback":

I love that I can quickly look at how kids are growing. I use the tier movement report. It's a report you can pull by grade level and by teacher, it's an at a glance sort of thing. I can see that immediate[ly] after they test. I love that.

Teachers also credited the program with pinpointing key areas for intervention. Respondents appreciated that Istation data provide both the overall student profile and also the detailed information on specific skills, available in the student's skill growth report and the changes in percentile over time:

I like the percentiles that show where students are, and I like being able to see the growth in the subtests.

In short, Istation helped teachers to identify student weaknesses and prescribe targeted interventions and supports for struggling students, which ultimately customized the learning experience for students while saving teachers time. Teachers found this component of Istation to be highly valuable.

Suggestions for improvement. When asked what could be improved about Istation, respondents cited a number of recommendations. Their responses generally fell into one of two categories: suggestions for software and content improvements.

Software fixes:

- Change the program response when students select a wrong answer. The owl popping up and the buzzer sound were discouraging and jolting for some students. Teachers also pointed out that students only hear a notification

sound for wrong answers but not for correct answers, limiting positive reinforcement.
- Fine tune the program for younger students to not require as advanced motor skills. In the comprehension and fluency sections, younger students sometimes hit the item *beside* the one they want, and younger students needed to be overly precise in tapping their selections on the screen to get the answer correct.
- Remind students that they have to hit the speaker button in order to hear the content. For the vocabulary section, it is assumed that students know that they have to use the speaker to hear the words.
- Offer demonstration videos for iPads and other devices to help with navigability and usability across different kinds of devices.

Content adjustments:

- Revise tricky and/or misleading questions. One of the questions on vocabulary asked students to click the picture that showed "land" but also included a picture of a helicopter, which lands. There was another question about the body of water where one of the options showed a picture of a bath tub filled with water.
- Reduce cognitive overload. For students in grades K–1, the test gives four multiple choice options. By the time students get through all four of the possible choices, they have forgotten the question.
- Improve quality of images. Some of the images for kindergarten students were outdated, or of low-quality. When pictures were compressed, they were hard to see.
- Include context with vocabulary. Vocabulary items tested whether students knew the meaning of words but did require students to also figure out the meaning of the word using context clues.
- Slow down and enunciate sounds more clearly. Sometimes the speaker's accent was hard to understand for ELs because of how quickly the speaker was talking.

Challenges in implementation. Respondents also pointed to challenges in broader implementation of Istation. Generally speaking, new initiatives may be met with resistance. However, positive experiences with using Istation will facilitate broader implementation. As explained by one teacher:

> *Anytime there's a change you have a bit of a struggle. Show them that it's worth the time and effort, there's a reward – it's a time saver. Teachers have seen programs come and go. Positive experiences and positive outcomes make it easier to get their buy-in.*

Administrative support for Istation was also cited as being important for successful implementation. Some respondents noted the lack of enthusiasm from administration, and in at least one case, a failure on the part of leadership to endorse Istation altogether. In cases where principals harbor resentment towards an initiative that has been perceived to be imposed upon them from the district or state level, teachers sense a lack of commitment and may not fully invest in the program. Instead, successful implementation requires explicit administrative support coupled with a teacher-led organic approach.

Despite these implementation challenges, when asked whether or not they would recommend Istation to others, the majority of educators said that they would recommend it, but with some caveats. Several teachers indicated that they needed additional time and experience using Istation before recommending it to other educators. Some teachers noted their uneasiness with using Istation as the sole form of assessment. This uneasiness was largely due to the fact that teachers were not yet sure to what extent Istation scores would correlate with scores from other assessments.

Summary

- Above all else, educators appreciated the immediate feedback and rich data on student progress that Istation afforded them. Teachers used the data to both inform their instruction and target interventions to individual students.
- Respondents listed a number of recommendations for program improvement, including both software fixes and tweaks in content.
- Administrative support for Istation was cited as being particularly impactful for successful implementation of Istation during its initial year.
- When asked if they would recommend Istation to other educators, the majority of respondents indicated that they would, but many are withholding their recommendations until Istation has been in place for a longer period of time.

Discussion

This study examined the relationship between the implementation of Istation, for either progress monitoring and/or use of the online curriculum, and improved student performance in reading in the state of Idaho in the 2017–18 and 2018–19 school years. The first research question concerned the relationship between Istation usage and student performance for all schools statewide during the 2018–19 school year. Results consistently showed a positive association between increased Istation usage and improved student performance, relative to schools with the lowest levels of Istation usage. Moreover, the sample size was large and included all public schools serving grades K–3 in Idaho. Therefore, this study supports the conclusion that moderate amounts of Istation usage related positively to improved student performance in reading, relative to schools with the lowest levels of Istation usage. Findings should be interpreted with caution, however, because this study cannot rule out systematic

differences between schools with different Istation usage levels. As such, it provides "promising" as opposed to causal evidence of the efficacy of Istation in improving student performance in reading per the Every Student Succeeds Act (ESSA).

In addressing the second research question, findings showed that the ISIP also was a good predictor of students' scores on the ISAT ELA, which is the state accountability assessment for students. The correlation between students' ISIP and ISAT ELA scores was relatively high approximating .70. Therefore, educators can look to second and third grade students' ISIP scores to forecast how their students will fare on the ISAT ELA, which is administered in the spring of students' third grade year.

The study also addressed a third, exploratory question comparing student outcomes for a group of schools that piloted the Istation assessment and curriculum components in 2017-18 and similar comparison schools. Relative to students in comparison schools, students in pilot schools did not perform significantly higher on either ISIP or ISAT achievement tests. However, they did outperform comparison students on some IRI subtests in the spring of 2018 and 2019. Additionally, students in pilot schools outperformed similar comparison peers on the spring 2018 IRI when the vast majority of students in the school participated in Istation. Usage data confirmed that the typical pilot school made tangible efforts to implement Istation components. However, in this initial year, usage was limited in both absolute and relative terms compared with time presumably devoted to regular reading and ELA instruction. For example, the typical pilot student received only 2–3.5 hours of total exposure to Istation during the school year.

A fourth research question concerned the perceptions of educators who piloted Istation in 2017-18. Results revealed reports of mostly positive experiences. Educators believed that the ISIP provided them with more nuanced information about their students' performance than did the state's previous IRI, which enabled them to adapt their instruction and target specific gaps in learning.

Conclusion

This study yielded a number of generally favorable correlational results of the benefit of Istation products for student performance in reading. However, the study design was limited for examining the efficacy of the Istation performance monitoring and curriculum due to what appears to have been limited or at best modest implementation at many schools. For example, in the typical school in the statewide sample, Istation was used (progress monitoring and/or curriculum) with only two-thirds of students in the school[22] and curriculum was used for only 9% of students. Therefore,

---

[22] The usage metric was determined at the school level because student-level usage data for progress monitoring were not available; therefore, a greater percentage of students in grades K-3 participated in the progress monitoring and mandatory testing than all students schoolwide.

while these results depict demonstrated statewide patterns for usage and associated performance outcomes, they may understate effects for schools that used the program at optimal levels. Another limiting factor was that many teachers were implementing Istation for the first time and therefore still learning how to employ it effectively.

More research is needed on the efficacy of Istation, whereby student performance in schools opting to implement the progress monitoring and curriculum with high fidelity is compared over multiple years to the performance of students in schools that do not participate in Istation or participate only in the progress monitoring. Having reliable student and classroom level usage metrics would further allow for rigorous analyses examining the relationship between usage of Istation and improved student performance.

Appendix A: Sample Selection

A comparison sample of schools that were similar to pilot schools in the 2017–18 school year were selected according to the following process. Student-level demographic and prior achievement data for students in grades K–3 were aggregated to the school level. Schools that piloted Istation were more likely to be located in rural areas, be in districts or LEAs with smaller enrollments, have lower percentages of teachers with 6+ years of experience, and serve students with lower average reading achievement, relative to all public elementary schools in the state. Pilot schools also served a greater average percentage of economically disadvantaged students than non-Istation schools, but the difference was not statistically significant. Table 12 shows the characteristics for schools that used and did not use Istation during the 2017–18 school year.

Table 12: School characteristics for Pilot versus non-pilot schools

	Pilot	**Non-pilot**	**Difference**
Mean district/LEA enrollment	5,083	10,837	***
Mean school enrollment	353	374	
Rural	64%	43%	***
Charter school	6%	10%	
Teachers with 6+ years of experience	62%	68%	*
Special education	11%	10%	
English learner	10%	7%	
Economically disadvantaged	57%	52%	
White	75%	78%	
Latino	18%	15%	
Other race	7%	6%	
School N	81	290	

NOTES—1. Estimates in this table are averages of school-level averages. 2. The difference denotes the statistical significance of the difference in the averages between Istation and non-Istation schools based on t-tests. 3. *p<.05, ***p<.001.

A logistic regression model was used to calculate the probability of being an Istation school based on school-level characteristics. The logistic model included the following variables: district or LEA enrollment, charter school status, economic disadvantage community eligibility, average student achievement on the fall 2017 IRI, the percentages of students in different grade levels, the percentage of special education, English learner, new English learner, continuing English learner, and economically disadvantaged students, and the percentages of White, Latino, and students of other races/ethnicities.

For each pilot school, a comparison school with a similar probability based on school characteristics was selected using one-to-one matching with propensity scores.

Eleven out of the 81 pilot schools did not have good comparison school matches and were therefore excluded from the sample.[23] These excluded pilot schools had lower average student reading achievement, greater percentages of special education, English learner, and Latino students and lower percentages of White students relative to the full sample of pilot schools. The excluded pilot schools were also more likely to be located in rural districts or LEAs relative to the full sample of Istation schools. Table 13 shows similar characteristics for selected pilot and comparison schools. The analyses comparing student achievement in pilot and comparison schools are based on this sample.

Table 2: School characteristics for selected pilot and comparison schools

	Pilot	**Comparison**
Mean district/LEA enrollment	5,563	5,945
Mean school enrollment	367	337
Rural	60%	61%
Charter school	7%	7%
Teachers with 6+ years of experience	62%	67%
Special education	11%	11%
English learner	8%	8%
Economically disadvantaged	55%	55%
White	79%	79%
Latino	16%	16%
Other race	6%	5%
School N	70	70

NOTES—1. Estimates in this table are averages of school-level averages. 2. There were no statistically significant differences at p<.05 or lower between selected Istation and comparison schools on any characteristic.

Additional pilot and comparison schools were not included in the sample above for the following reasons:

- Comparison schools with enrollments of less than 50 students because all pilot schools had student enrollments of at least 50 students.
- Schools that piloted Istation but only for half of the school year.[24] Schools that piloted Istation starting in later spring 2018 were eligible to be included in the comparison group of schools, and 16 were selected by the process described above.

---

[23] The following 11 pilot schools did not have a good comparison school match and were therefore excluded from the pilot versus comparison school analyses: Lakeside, Driggs, Fort Hall, Paris, Shoshone, Van Buren, Victor, Wendell, Riverside, Donald Stalker, and North Gem. These schools were included in the 2018–19 usage analyses, however.

[24] These schools included Bryan Elementary, Gooding Elementary, Heritage Community Charter, and Mullan Trail Elementary.

- Pilot schools with only a handful of Istation users or students with non-missing ISIP scores. Pilot schools were excluded from the sample when less than 20% of the student population AND fewer than 30 students had non-missing ISIP scores.

These schools were eligible to be included in the 2018–19 usage analyses for all schools, however, and were included when students had non-missing test scores.

Appendix B: Regression Results

Regression Results for Statewide Analyses

Association between Istation usage and student learning gains. This section contains regression results from statewide analyses that compared student performance in schools with lower versus higher amounts of Istation usage in the 2018–19 school year. The low, mid, mid-high, and high usage levels were determined based on quartiles of usage in the 2018–19 school year for all schools in Idaho. As with the previous analyses, the quartiles for the curriculum usage variables were calculated after first restricting to the sample of schools that had used the Istation curriculum.

Table 34: Associations between Istation usage levels and improved student performance

	Estimate		Standard error	p-value
Outcome: ISIP Spring 2019				
Usage indicator: Average sessions for school				
Istation mid usage	2.12	***	0.44	0.0000
Istation mid-high usage	3.08	***	0.45	0.0000
Istation high usage	1.96	***	0.45	0.0000
Intercept (low usage)	230.46	***	0.33	0.0000
Student N	84603			
School N	399			
Usage indicator: Average sessions – grade K only				
Istation mid usage	1.84	*	0.73	0.0121
Istation mid-high usage	3.77	***	0.75	0.0000
Istation high usage	3.09	***	0.75	0.0000
Intercept (low usage)	203.16	***	0.54	0.0000
Student N	20172			
School N	382			
Usage indicator: Average sessions – grade 1 only				
Istation mid usage	2.36	***	0.57	0.0000
Istation mid-high usage	2.97	***	0.59	0.0000
Istation high usage	1.91	**	0.59	0.0013
Intercept (low usage)	222.87	***	0.43	0.0000
Student N	21115			
School N	389			
Usage indicator: Average sessions – grade 2 only				
Istation mid usage	2.17	***	0.51	0.0000
Istation mid-high usage	2.33	***	0.54	0.0000
Istation high usage	1.73	**	0.54	0.0013
Intercept (low usage)	240.50	***	0.39	0.0000
Student N	21532			
School N	388			
Usage indicator: Average sessions – grade 3 only				
Istation mid usage	2.65	***	0.51	0.0000
Istation mid-high usage	3.16	***	0.54	0.0000
Istation high usage	2.03	***	0.53	0.0001
Intercept (low usage)	252.98	***	0.39	0.0000
Student N	21784			
School N	389			
Usage indicator: Average weeks for school				
Istation mid usage	2.49	***	0.43	0.0000

	Estimate		Standard error	p-value
Istation mid-high usage	2.95	***	0.45	0.0000
Istation high usage	2.26	***	0.45	0.0000
Intercept (low usage)	230.31	***	0.33	0.0000
Student N	84603			
School N	399			
Usage indicator: Average total minutes for school				
Istation mid usage	2.37	***	0.44	0.0000
Istation mid-high usage	2.52	***	0.46	0.0000
Istation high usage	1.84	***	0.46	0.0001
Intercept (low usage)	230.56	***	0.33	0.0000
Student N	84603			
School N	399			
Usage indicator: Average total minutes – grade K only				
Istation mid usage	1.85	*	0.72	0.0104
Istation mid-high usage	3.54	***	0.77	0.0000
Istation high usage	3.07	***	0.75	0.0000
Intercept (low usage)	203.23	***	0.53	0.0000
Student N	20172			
School N	382			
Usage indicator: Average total minutes – grade 1 only				
Istation mid usage	2.68	***	0.58	0.0000
Istation mid-high usage	2.71	***	0.63	0.0000
Istation high usage	2.02	**	0.61	0.0010
Intercept (low usage)	222.82	***	0.45	0.0000
Student N	21115			
School N	389			
Usage indicator: Average total minutes – grade 2 only				
Istation mid usage	2.57	***	0.53	0.0000
Istation mid-high usage	2.43	***	0.57	0.0000
Istation high usage	2.09	***	0.56	0.0002
Intercept (low usage)	240.30	***	0.41	0.0000
Student N	21532			
School N	388			
Usage indicator: Average total minutes – grade 3 only				
Istation mid usage	2.45	***	0.51	0.0000
Istation mid-high usage	3.44	***	0.56	0.0000
Istation high usage	2.02	***	0.54	0.0002
Intercept (low usage)	252.93	***	0.39	0.0000
Student N	21784			
School N	389			
Usage indicator: Average curriculum minutes for school				

	Estimate		Standard error	p-value
Istation low usage	0.04		0.45	0.9355
Istation mid usage	0.56		0.48	0.2419
Istation mid-high usage	0.54		0.49	0.2634
Istation high usage	0.44		0.47	0.3528
Intercept (no curriculum)	232.07	***	0.27	0.0000
Student N	84603			
School N	399			
Usage indicator: Average curriculum minutes – grade K only				
Istation low usage	-0.73		0.77	0.3453
Istation mid usage	-0.76		0.84	0.3660
Istation mid-high usage	0.18		0.84	0.8323
Istation high usage	0.82		0.81	0.3073
Intercept (no curriculum)	205.42	***	0.43	0.0000
Student N	20172			
School N	382			
Usage indicator: Average curriculum minutes – grade 1 only				
Istation low usage	-0.25		0.59	0.6690
Istation mid usage	0.18		0.64	0.7734
Istation mid-high usage	0.00		0.65	0.9981
Istation low usage	0.21		0.62	0.7330
Intercept (no curriculum)	224.72	***	0.35	0.0000
Student N	21115			
School N	389			
Usage indicator: Average curriculum minutes – grade 2 only				
Istation low usage	0.82		0.54	0.1265
Istation mid usage	1.13	*	0.58	0.0494
Istation mid-high usage	0.86		0.60	0.1475
Istation high usage	0.39		0.56	0.4858
Intercept (no curriculum)	241.57	***	0.32	0.0000
Student N	21532			
School N	388			
Usage indicator: Average curriculum minutes – grade 3 only				
Istation low usage	0.02		0.55	0.9640
Istation mid usage	0.56		0.61	0.3652
Istation mid-high usage	-0.12		0.60	0.8472
Istation high usage	-0.13		0.58	0.8214
Intercept (no curriculum)	254.94	***	0.32	0.0000
Student N	21784			
School N	389			
Usage indicator: Percent users in school				
Istation mid usage	0.28		0.48	0.5556

	Estimate		Standard error	p-value
Istation mid-high usage	1.02	*	0.49	0.0365
Istation high usage	0.50		0.50	0.3169
Intercept (low usage)	231.86	***	0.36	0.0000
Student N	84603			
School N	399			
Usage indicator: Percent curriculum users in school				
Istation low usage	1.83	*	0.73	0.0118
Istation mid usage	3.27	***	0.76	0.0000
Istation mid-high usage	3.25	***	0.75	0.0000
Istation high usage	2.81	***	0.76	0.0002
Intercept (no curriculum)	229.70	***	0.67	0.0000
Student N	84603			
School N	399			
Outcome: ISAT ELA Spring 2019				
Usage indicator: Average sessions for school				
Istation mid usage	7.54	*	3.01	0.0123
Istation mid-high usage	4.42		3.12	0.1568
Istation high usage	3.39		3.10	0.2739
Intercept (low usage)	2424.60	***	2.29	0.0000
Student N	21698			
School N	390			
Usage indicator: Average sessions – grade 3 only				
Istation mid usage	8.12	**	2.86	0.0046
Istation mid-high usage	0.50		3.03	0.8679
Istation high usage	3.07		3.02	0.3090
Intercept (low usage)	2425.65	***	2.15	0.0000
Student N	21698			
School N	390			
Usage indicator: Average weeks for school				
Istation mid usage	8.17	**	2.93	0.0054
Istation mid-high usage	2.45		3.10	0.4301
Istation high usage	3.78		3.12	0.2263
Intercept (low usage)	2424.82	***	2.27	0.0000
Student N	21698			
School N	390			
Usage indicator: Average total minutes for school				
Istation mid usage	3.76		3.00	0.2098
Istation mid-high usage	4.25		3.19	0.1819
Istation high usage	2.14		3.15	0.4963
Intercept (low usage)	2425.97	***	2.30	0.0000
Student N	21698			

	Estimate		Standard error	p-value
School N	390			
Usage indicator: Average total minutes – grade 3 only				
Istation mid usage	6.39	*	2.88	0.0265
Istation mid-high usage	2.79		3.15	0.3756
Istation high usage	2.63		3.05	0.3885
Intercept (low usage)	2425.58	***	2.20	0.0000
Student N	21698			
School N	390			
Usage indicator: Average curriculum minutes for school				
Istation low usage	0.70		2.94	0.8110
Istation mid usage	2.72		3.22	0.3982
Istation mid-high usage	-2.15		3.23	0.5046
Istation high usage	1.36		3.09	0.6606
Intercept (no curriculum)	2428.22	***	1.81	0.0000
Student N	21698			
School N	390			
Usage indicator: Average curriculum minutes – grade 3 only				
Istation low usage	-2.43		2.98	0.4153
Istation mid usage	1.58		3.33	0.6355
Istation mid-high usage	-1.72		3.25	0.5971
Istation high usage	-1.36		3.17	0.6667
Intercept (no curriculum)	2429.26	***	1.68	0.0000
Student N	21698			
School N	390			
Usage indicator: Percent users in school				
Istation mid usage	0.53		3.17	0.8681
Istation mid-high usage	2.42		3.29	0.4618
Istation high usage	-0.79		3.42	0.8170
Intercept (low usage)	2428.07	***	2.47	0.0000
Student N	21698			
School N	390			
Usage indicator: Percent curriculum users in school				
Istation low usage	-7.62		5.08	0.1336
Istation mid usage	-0.23		5.29	0.9650
Istation mid-high usage	-1.24		5.26	0.8141
Istation high usage	-5.84		5.33	0.2730
Intercept (no curriculum)	2432.09	***	4.74	0.0000
Student N	21698			
School N	390			

NOTES—1. *p<.05, **p<.01, ***p<.001. 2. The models also controlled for student prior achievement (fall 2018 ISIP score), grade level, gender, special education status, English learner (EL) status, new or

continuing EL, economic disadvantage status, race/ethnicity; LEA enrollment; and school percentage of teachers with 6+ years of experience, charter school status, mean prior achievement, percentage EL, percentage special education, percentage economically disadvantaged, and the percentage of students at various grade levels.

Istation usage and student learning gains by subtest. We also analyzed the relationship between usage of Istation (measured continuously) and improved student outcomes on each subtest of the ISIP and ISAT ELA in spring 2019. These analyses were exploratory and are summarized in the table below. These results show the relationship between usage of Istation and improved outcomes on each subtest, relative to schools with lower levels of Istation usage.

Greater Istation usage, defined in multiple ways, was associated with improved scores on the listening comprehension, letter knowledge, and vocabulary ISIP subtests. There was no association between Istation usage and improved scores on the alphabetic decoding, spelling, and comprehension ISIP subtests.

Table 45: Summary of associations between Istation usage and improved student performance by subtest in 2018–19

Subtest	Usage indicators	Relationship	Regression coefficient (standard error)
Outcome: ISAT ELA Spring 2019			
Reading	Average sessions	NS	
Listening	Avg. total min	NS	
Writing	Avg. curriculum min	NS	
Research	**3rd grade only	NS	
Outcome: IRI (ISIP) Spring 2019			
Listening comprehension	Average sessions	**	0.04 (0.02)
	Average weeks	***	0.17 (0.04)
	Avg. total min	**	0.13 (0.04) *in hours*
	Avg. curriculum min	*	0.11 (0.05) *in hours*
Letter knowledge	Average sessions	**	0.04 (0.01)
	Average weeks	*	0.10 (0.04)
	Avg. total min	**	0.11 (0.04) *in hours*
	Avg. curriculum min	**	0.12 (0.04) *in hours*
Vocabulary	Average sessions	*	0.02 (0.01)
	Average weeks	**	0.10 (0.03)
	Avg. total min	**	0.08 (0.03) *in hours*
	Avg. curriculum min	*	0.07 (0.03) *in hours*
Text fluency	Average sessions	NS	
	Average weeks	NS	
	Avg. total min	NS	
	Avg. curriculum min	NS	
Alphabetic decoding	Average sessions	NS	
	Average weeks	NS	
	Avg. total min	NS	

Subtest	Usage indicators	Relationship	Regression coefficient (standard error)
	Avg. curriculum min	NS	
Spelling	Average sessions	NS	
	Average weeks	NS	
	Avg. total min	NS	
	Avg. curriculum min	NS	
Comprehension	Average sessions	NS	
	Average weeks	NS	
	Avg. total min	NS	
	Avg. curriculum min	NS	

NOTES—1. NS=not statistically significant, *p<.05, **p<.01, ***p<.001. 2. As noted above, the usage indicators were continuous measures of Istation usage, measured at either the school level for the ISIP analyses or the third grade school level for the ISAT ELA analyses. 3. The models also controlled for student prior achievement (fall 2018 composite ISIP score for the ISAT sample, and each ISIP subtest score in fall 2018 for the ISIP sample), grade level, gender, special education status, English learner (EL) status, new or continuing EL, economic disadvantage status, race/ethnicity; LEA enrollment; and school percentage of teachers with 6+ years of experience, charter school status, mean prior achievement, percentage EL, percentage special education, percentage economically disadvantaged, and the percentage of students at various grade levels.

Regression Results for Pilot versus Comparison Analyses

Overall effects of Istation. The following regression results show the overall results from the hierarhical linear models comparing student performance for students in pilot and comparison schools. The models controlled for baseline achievement and student and school characteristics, as outlined earlier in the report. All covariates were grand-mean centered to facilitate the interpretation of the intercept. The models were estimated separately by outcome variable and year.

Table 56: Overall effects of piloting relative to comparison schools

	Estimate		Standard error	p-value
Outcome: IRI Spring 2018				
Pilot effect	0.00		0.02	0.8914
Intercept	2.58	***	0.01	0.0000
Student N	29,353			
School N	140			
τ^2	0.01			
σ^2	0.28			
ICC	0.04			
Outcome: ISAT ELA Spring 2018				
Pilot effect	-0.10		3.75	0.9789
Intercept	2427.45	***	2.64	0.0000
Student N	7,341			

	Estimate		Standard error	p-value
School N	136			
τ^2	335.81			
σ^2	4012.26			
ICC	0.08			
Outcome: ISIP Spring 2019				
Pilot effect	0.55		0.52	0.2897
Intercept	240.16	***	0.37	0.0000
Student N	20,695			
School N	139			
τ^2	6.59			
σ^2	210.88			
ICC	0.03			
Outcome: ISAT ELA Spring 2019				
Pilot effect	0.48		3.30	0.8841
Intercept	2,448.73	***	2.29	0.0000
Student N	13172			
School N	136			
τ^2	258.40			
σ^2	4649.84			
ICC	0.05			

NOTES—1. *p<.05, **p<.01, ***p<.001. 2. The models also controlled for student prior achievement (fall 2017 IRI score), grade level, gender, special education status, English learner (EL) status, new or continuing EL, economic disadvantage status, race/ethnicity; LEA enrollment; and school percentage of teachers with 6+ years of experience, charter school status, mean prior achievement, percentage EL, percentage special education, percentage economically disadvantaged, and the percentage of students at various grade levels. 3. τ^2 is the variance of the school-level intercepts, and σ^2 is the variation of the student-level residuals. ICC is the intraclass correlation coefficient, indicating what percentage of the variation in the outcome can be attributed to between-school differences in student performance, conditional on the covariates.

Effects of piloting by subtest. We also explored whether there appeared to be effects of piloting on particular subtests of state assessments. For example, the IRI contained three subtests (letter naming fluency, letter sound fluency, and reading curriculum-based measure), the ISAT contained four subtests (reading, writing, listening, and research), and the ISIP contained numerous subtests. While these analyses inform whether piloting appeared to have an effect on student knowledge and skills in a particular area, we caution that statistically significant findings may also arise due to chance alone when conducting a large number of analyses.

Table 67: Effects of piloting relative to comparison schools by subtest

	Estimate		Standard error	p-value
Outcome: IRI Spring 2018				

	Estimate		Standard error	p-value
Subtest: Letter naming fluency				
Pilot effect	0.41		1.12	0.7107
Intercept	40.61	***	0.79	0.0000
Student N	6874			
School N	135			
Subtest: Letter sound fluency				
Pilot effect	2.05	*	1.00	0.0415
Intercept	50.71	***	0.70	0.0000
Student N	14270			
School N	140			
Subtest: Reading-curriculum based measure				
Pilot effect	-1.94		3.25	0.5518
Intercept	106.52	***	2.27	0.0000
Student N	22455			
School N	140			
Outcome: ISAT ELA Spring 2018				
Subtest: Reading				
Pilot effect	-2.84		3.67	0.4392
Intercept	2432.66	***	2.61	0.0000
Student N	7328			
School N	136			
Subtest: Writing				
Pilot effect	2.33		4.48	0.6027
Intercept	2417.48	***	3.16	0.0000
Student N	7328			
School N	136			
Subtest: Listening				
Pilot effect	-3.19		4.47	0.4754
Intercept	2444.27	***	3.18	0.0000
Student N	7328			
School N	136			
Subtest: Research				
Pilot effect	0.27		5.12	0.9582
Intercept	2409.88	***	3.62	0.0000
Student N	7328			
School N	136			
Outcome: IRI (ISIP) Spring 2019				
Subtest: Listening comprehension				
Pilot effect	5.00		3.64	0.1695
Intercept	206.02	***	2.57	0.0000
Student N	109			

	Estimate		Standard error	p-value
School N	65			
Subtest: Letter knowledge				
Pilot effect	-0.88		1.33	0.5091
Intercept	203.52	***	0.95	0.0000
Student N	1026			
School N	129			
Subtest: Phonemic awareness				
Pilot effect	-0.61		0.96	0.5235
Intercept	206.50	***	0.68	0.0000
Student N	2100			
School N	137			
Subtest: Alphabetic decoding				
Pilot effect	0.27		0.75	0.7173
Intercept	223.80	***	0.53	0.0000
Student N	6736			
School N	139			
Subtest: Comprehension				
Pilot effect	0.84		0.62	0.1762
Intercept	243.78	***	0.44	0.0000
Student N	20572			
School N	139			
Subtest: Vocabulary				
Pilot effect	2.05	**	0.79	0.0094
Intercept	247.52	***	0.55	0.0000
Student N	20691			
School N	139			
Subtest: Spelling				
Pilot effect	0.04		0.52	0.9461
Intercept	235.73	***	0.36	0.0000
Student N	20573			
School N	139			
Subtest: Text fluency				
Pilot effect	3.60	**	1.21	0.0028
Intercept	55.33	***	0.85	0.0000
Student N	18624			
School N	139			
Outcome: ISAT ELA Spring 2019				
Subtest: Reading				
Pilot effect	0.99		3.22	0.7586
Intercept	2453.14	***	2.25	0.0000
Student N	13167			

	Estimate		Standard error	p-value
School N	136			
Subtest: Writing				
Pilot effect	1.00		4.29	0.8158
Intercept	2442.69	***	2.98	0.0000
Student N	13167			
School N	136			
Subtest: Listening				
Pilot effect	-2.02		4.08	0.6201
Intercept	2465.51	***	2.85	0.0000
Student N	13167			
School N	136			
Subtest: Research				
Pilot effect	-1.00		4.07	0.8067
Intercept	2426.60	***	2.84	0.0000
Student N	13167			
School N	136			

NOTES—1. *p<.05, **p<.01, ***p<.001. 2. The models also controlled for student prior achievement (fall 2017 IRI composite or subtest score), grade level, gender, special education status, English learner (EL) status, new or continuing EL, economic disadvantage status, race/ethnicity; LEA enrollment; and school percentage of teachers with 6+ years of experience, charter school status, mean prior achievement, percentage EL, percentage special education, percentage economically disadvantaged, and the percentage of students at various grade levels.

Effects of Piloting for schools with different usage levels. This section contains regression results for the models that compared student performance for comparison students and pilot students with different levels of usage. Models were estimated separately for each usage indicator, which was defined at either the school or grade level. In other words, only one indicator of usage was included in the model at a time. The low, mid, mid-high, and high usage levels were determined based on quartiles of usage for the schools that had used Istation during the 2017–18 school year. The quartiles for the curriculum usage variables were calculated after first restricting to the sample of schools that had used the Istation curriculum, although nearly all Istation pilot schools used the curriculum at least to some extent during the 2017–18 school year.

Table 78: Effects of piloting for schools with different usage levels relative to comparison schools

	Estimate		Standard error	p-value
Outcome: IRI Spring 2018				

	Estimate		Standard error	p-value
Usage indicator: Average sessions for school				
Istation low usage	0.01		0.03	0.7844
Istation mid usage	-0.01		0.03	0.7587
Istation mid-high usage	0.03		0.03	0.2671
Istation high usage	0.04		0.03	0.1921
Comparison	2.56	***	0.02	0.0000
Student N	29353			
School N	140			
Usage indicator: Average sessions – grade K only				
Istation low usage	0.03		0.06	0.5650
Istation mid usage	-0.02		0.06	0.7315
Istation mid-high usage	0.03		0.06	0.6312
Istation high usage	0.10		0.06	0.0764
Comparison	2.69	***	0.04	0.0000
Student N	6889			
School N	135			
Usage indicator: Average sessions – grade 1 only				
Istation low usage	0.02		0.05	0.6726
Istation mid usage	-0.03		0.05	0.5181
Istation mid-high usage	0.02		0.05	0.7676
Istation high usage	0.05		0.05	0.3738
Comparison	2.47	***	0.03	0.0000
Student N	7391			
School N	140			
Usage indicator: Average sessions – grade 2 only				
Istation low usage	-0.01		0.03	0.7043
Istation mid usage	0.00		0.03	0.9002
Istation mid-high usage	0.06		0.03	0.0526
Istation high usage	0.02		0.03	0.4394
Comparison	2.49	***	0.02	0.0000
Student N	7675			
School N	140			
Usage indicator: Average sessions – grade 3 only				
Istation low usage	-0.02		0.02	0.5095
Istation mid usage	-0.02		0.02	0.5163
Istation mid-high usage	0.01		0.02	0.5350
Istation high usage	-0.01		0.02	0.6829
Comparison	2.63	***	0.01	0.0000
Student N	7398			
School N	136			
Usage indicator: Average weeks for school				

	Estimate		Standard error	p-value
Istation low usage	0.03		0.03	0.3572
Istation mid usage	-0.01		0.03	0.7296
Istation mid-high usage	0.01		0.03	0.6822
Istation high usage	0.05		0.03	0.1271
Comparison	2.56	***	0.02	0.0000
Student N	29353			
School N	140			
Usage indicator: Average total minutes for school				
Istation low usage	0.01		0.03	0.7806
Istation mid usage	-0.02		0.03	0.4949
Istation mid-high usage	0.03		0.03	0.2657
Istation high usage	0.05		0.03	0.1092
Comparison	2.56	***	0.02	0.0000
Student N	29353			
School N	140			
Usage indicator: Average total minutes – grade K only				
Istation low usage	0.04		0.06	0.5078
Istation mid usage	-0.02		0.06	0.6885
Istation mid-high usage	0.08		0.06	0.1611
Istation high usage	0.04		0.06	0.4471
Comparison	2.69	***	0.04	0.0000
Student N	6889			
School N	135			
Usage indicator: Average total minutes – grade 1 only				
Istation low usage	-0.02		0.05	0.6800
Istation mid usage	-0.02		0.05	0.6798
Istation mid-high usage	0.04		0.05	0.3989
Istation high usage	0.05		0.05	0.3846
Comparison	2.47	***	0.03	0.0000
Student N	7391			
School N	140			
Usage indicator: Average total minutes – grade 2 only				
Istation low usage	0.00		0.03	0.9784
Istation mid usage	-0.01		0.03	0.7453
Istation mid-high usage	0.04		0.03	0.2374
Istation high usage	0.04		0.03	0.1583
Comparison	2.49	***	0.02	0.0000
Student N	7675			
School N	140			
Usage indicator: Average total minutes – grade 3 only				
Istation low usage	-0.02		0.02	0.4129

	Estimate		Standard error	p-value
Istation mid usage	-0.01		0.02	0.6960
Istation mid-high usage	0.00		0.02	0.9871
Istation high usage	0.00		0.02	0.9298
Comparison	2.63	***	0.01	0.0000
Student N	7398.00			
School N	136.00			
Usage indicator: Average curriculum minutes for school				
Istation no curriculum	-0.06		0.04	0.0971
Istation low usage	0.05		0.05	0.2956
Istation mid usage	0.04		0.04	0.3363
Istation mid-high usage	0.07		0.04	0.0590
Istation high usage	0.11	*	0.05	0.0178
Comparison	2.56	***	0.02	0.0000
Student N	29353			
School N	140			
Usage indicator: Average curriculum minutes – grade K only				
Istation no curriculum	-0.09		0.06	0.1268
Istation low usage	0.04		0.08	0.6220
Istation mid usage	0.14		0.08	0.0796
Istation mid-high usage	0.11		0.07	0.1605
Istation high usage	0.17	*	0.08	0.0254
Comparison	2.69	***	0.04	0.0000
Student N	6889			
School N	135			
Usage indicator: Average curriculum minutes – grade 1 only				
Istation no curriculum	-0.07		0.06	0.2908
Istation low usage	0.07		0.09	0.4344
Istation mid usage	0.00		0.07	0.9839
Istation mid-high usage	0.09		0.07	0.1670
Istation high usage	0.11		0.08	0.1727
Comparison	2.47	***	0.03	0.0000
Student N	7391.00			
School N	140.00			
Usage indicator: Average curriculum minutes – grade 2 only				
Istation no curriculum	-0.02		0.04	0.5561
Istation low usage	-0.02		0.04	0.6089
Istation mid usage	0.05		0.04	0.1738
Istation mid-high usage	0.06		0.04	0.1144
Istation high usage	0.07		0.05	0.1524
Comparison	2.49	***	0.02	0.0000
Student N	7675.00			

	Estimate		Standard error	p-value
School N	140.00			
Usage indicator: Average curriculum minutes – grade 3 only				
Istation no curriculum	-0.03		0.03	0.2485
Istation low usage	0.00		0.03	0.9392
Istation mid usage	0.00		0.04	0.9539
Istation mid-high usage	0.04		0.03	0.2083
Istation high usage	0.03		0.03	0.3768
Comparison	2.63	***	0.01	0.0000
Student N	7398.00			
School N	136.00			
Usage indicator: Percent users in school				
Istation low usage	0.02		0.03	0.5639
Istation mid usage	0.00		0.03	0.9889
Istation mid-high usage	0.06	*	0.03	0.0356
Istation high usage	-0.01		0.03	0.7263
Comparison	2.56	***	0.02	0.0000
Student N	29353			
School N	140			
Usage indicator: Percent curriculum users in school				
Istation no curriculum	-0.06		0.04	0.1215
Istation low usage	0.06		0.05	0.2415
Istation mid usage	0.03		0.04	0.3512
Istation mid-high usage	0.08	*	0.04	0.0491
Istation high usage	0.08		0.05	0.0724
Comparison	2.56	***	0.02	0.0000
Student N	29353			
School N	140			
Outcome: ISAT ELA Spring 2018				
Usage indicator: Average sessions for school				
Istation low usage	-9.17		5.45	0.0923
Istation mid usage	4.41		5.79	0.4459
Istation mid-high usage	2.32		5.43	0.6697
Istation high usage	0.30		5.46	0.9566
Comparison	2427.89	***	3.01	0.0000
Student N	7341			
School N	136			
Usage indicator: Average sessions – grade 3 only				
Istation low usage	-6.77		5.55	0.2225
Istation mid usage	-0.94		5.73	0.8697
Istation mid-high usage	2.50		5.58	0.6547
Istation high usage	1.91		5.53	0.7304

	Estimate		Standard error	p-value
Comparison	2427.89	***	3.03	0.0000
Student N	7341			
School N	136			
Usage indicator: Average weeks for school				
Istation low usage	-7.44		5.58	0.1822
Istation mid usage	2.60		5.57	0.6405
Istation mid-high usage	-1.37		5.77	0.8118
Istation high usage	2.01		5.54	0.7162
Comparison	2428.08	***	3.03	0.0000
Student N	7341			
School N	136			
Usage indicator: Average total minutes for school				
Istation low usage	-11.22	*	5.36	0.0365
Istation mid usage	2.72		5.70	0.6331
Istation mid-high usage	5.34		5.42	0.3246
Istation high usage	1.36		5.37	0.7995
Comparison	2427.77	***	2.96	0.0000
Student N	7341			
School N	136			
Usage indicator: Average total minutes – grade 3 only				
Istation low usage	-7.20		5.55	0.1945
Istation mid usage	0.20		5.60	0.9711
Istation mid-high usage	-0.08		5.54	0.9884
Istation high usage	3.85		5.53	0.4857
Comparison	2427.88	***	3.03	0.0000
Student N	7341			
School N	136			
Usage indicator: Average curriculum minutes for school				
Istation no curriculum	1.66		6.74	0.8055
Istation low usage	-6.07		8.90	0.4948
Istation mid usage	-7.36		6.86	0.2834
Istation mid-high usage	3.58		6.83	0.5998
Istation high usage	0.38		8.67	0.9650
Comparison	2427.90	***	3.01	0.0000
Student N	7341			
School N	136			
Usage indicator: Average curriculum minutes – grade 3 only				
Istation no curriculum	-0.82		5.93	0.8895
Istation low usage	-6.76		6.58	0.3040
Istation mid usage	0.67		9.03	0.9404
Istation mid-high usage	4.41		7.11	0.5355

	Estimate		Standard error	p-value
Istation high usage	4.70		7.90	0.5519
Comparison	2427.91	***	3.01	0.0000
Student N	7341.00			
School N	136.00			
Usage indicator: Percent users in school				
Istation low usage	-7.64		5.61	0.1729
Istation mid usage	-1.14		5.54	0.8366
Istation mid-high usage	8.43		5.54	0.1281
Istation high usage	-3.51		5.69	0.5365
Comparison	2427.91	***	3.00	0.0000
Student N	7341			
School N	136			
Usage indicator: Percent curriculum users in school				
Istation no curriculum	6.06		6.66	0.3627
Istation low usage	-14.76		8.70	0.0898
Istation mid usage	-0.57		6.63	0.9314
Istation mid-high usage	-1.82		7.44	0.8070
Istation high usage	-4.82		8.34	0.5633
Comparison	2427.79	***	3.02	0.0000
Student N	7341			
School N	136			

NOTES—1. *p<.05, **p<.01, ***p<.001. 2. The models also controlled for student prior achievement (fall 2017 IRI score), grade level, gender, special education status, English learner (EL) status, new or continuing EL, economic disadvantage status, race/ethnicity; LEA enrollment; and school percentage of teachers with 6+ years of experience, charter school status, mean prior achievement, percentage EL, percentage special education, percentage economically disadvantaged, and the percentage of students at various grade levels.

Appendix C: Sensitivity Analysis

In selecting the Istation and comparison school sample (see Appendix A), 16 out of the 70 schools that were included in the comparison school sample started piloting Istation towards the end of the 2017–18 year. While these schools were included in the comparison sample for the main analyses, we conducted a sensitivity analysis to determine whether results changed if these schools were excluded from the sample.

Using the same models as in the main analysis, results were mostly the same when excluding the "late adopters" of Istation except for on the ISIP that was administered in 2019. When "late adopters" were excluded, pilot schools outperformed comparison schools on the spring 2019 ISIP by an average of 1 point (p <. 05). This finding implies that schools that administered ISIP for two years in a row had improved student performance relative to schools with only one year of ISIP administration. This finding is not surprising, given that educators and students may both become more familiar with assessments over time, resulting in higher student performance.

Table 8: Effects of piloting removing Istation "late adopters" in 2017–18

	Estimate		Standard error	p-value
Outcome: IRI Spring 2018				
Pilot effect	0.01		0.02	0.6134
Intercept	2.56	***	0.02	0.0000
Student N	25,799			
School N	124			
Outcome: ISAT ELA Spring 2018				
Pilot effect	-0.71		4.04	0.8608
Intercept	2427.54	***	3.27	0.0000
Student N	6,369			
School N	120			
Outcome: ISIP Spring 2019				
Pilot effect	1.01	*	0.50	0.0454
Intercept	239.74	***	0.37	0.0000
Student N	18,197			
School N	138			
Outcome: ISAT ELA Spring 2019				
Pilot effect	2.63		3.46	0.4464
Intercept	2447.97	***	2.54	0.0000
Student N	11,496			
School N	135			

NOTES—1. *p<.05, **p<.01, ***p<.001. 2. The models also controlled for student prior achievement (fall 2017 IRI score), grade level, gender, special education status, English learner (EL) status, new or continuing EL, economic disadvantage status, race/ethnicity; LEA enrollment; and school percentage of

teachers with 6+ years of experience, charter school status, mean prior achievement, percentage EL, percentage special education, percentage economically disadvantaged, and the percentage of students at various grade levels.

It is also important to note that for the sample of schools with Istation "late adopters" excluded, baseline equivalence on the pretest measure was achieved (and less than 0.25 standard deviations) in all cases. One difference in this subsample of schools and the sample of schools included in the main analysis, however, is that there were greater differences between the pilot and comparison schools in this sample in terms of percentages of economically disadvantaged and EL students. Fifty-five percent of pilot students in this sample were economically disadvantaged compared with 49% of comparison students, and 10% of pilot students were ELs compared with 8% of comparison students. These demographic differences were relatively small, but we cannot rule out other factors that may have contributed to the improved student performance on the ISIP in spring 2019 for pilot schools in their second year of implementation.