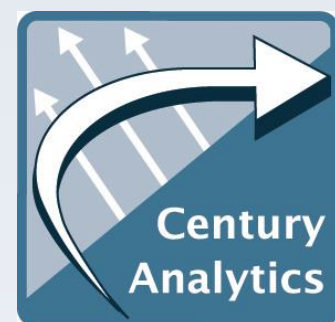


Impacts of Study Island on Student Reading Achievement

June 2019

Bruce Randel, Ph.D.
Century Analytics, Inc.
Centennial CO
www.centuryanalytics.com



Executive Summary

Century Analytics conducted a rigorous evaluation of Edmentum's Study Island to estimate its impact on student achievement in reading in Grade 2 through Grade 8. Study Island is a flexible product that provides students with practice and instructional support on standards-based topics in mathematics, reading, science, and social studies. This study's quasi-experimental design (QED), analyses, and measures meet the What Works Clearinghouse (WWC) 4.0 standards needed to achieve a rating of Meets WWC Group Design Standards with Reservations (WWC, 2017). This study also meets the Every Student Succeeds Act (ESSA) guidance for Moderate Evidence (U.S. Department of Education, 2016).

Students were included in the Study Island group if they had completed at least half of the Study Island practice topics available in their grade level between their fall and winter administration of the Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) assessment. The MAP is a standardized educational assessment that meets WWC standards for validity and reliability. Students who had completed no Study Island topics during the same period were matched to the Study Island students in each grade level on their fall MAP test scores and the duration between their fall and winter MAP test administrations. At each grade level, Study Island students were within the WWC threshold for baseline equivalence on both fall MAP test scores and duration.

Analyses revealed statistically significant positive impacts of Study Island use on winter MAP reading achievement in Grades 2 through 8. Effect sizes ranged from 0.30 for Grade 2 to 0.07 for Grade 4 and improvement indexes ranging from 11.79 to 2.79. Improvement indexes show the expected change in percentile rank for an average comparison student if he or she had been in the intervention group.

Results suggest that students who complete at least half of the Study Island topics available in their grade level will make statistically significant gains in achievement relative to students who do not complete any Study Island topics. These results suggest that Study Island is providing students with practice in standards-based academic content that improves students' scores on standardized tests.

This study is not without limitations. The definition for the Study Island group focused solely on topic completion. The study does not shed any light on the potential impact of any other of Study Island's student resources (e.g., lessons, games, flash cards, printable worksheets) or the impact of Study Island when integrated into classroom instruction.

Many students in the Study Island group completed more than the minimum number of topics needed to be included in the Study Island group. In some cases, students completed more than double the number of topics available at their grade level. Whether this amount of practice can be attributed to Study Island or some characteristics of the students is a question unanswered by this study. In addition, the lack of any student demographic variables limits the generalizability of the study's findings.

Future research on Study Island should incorporate a broader definition of student usage in order to estimate the impact of the many student resources available beyond practice topics. This research also should include student demographic characteristics to help understand which groups of students may benefit most from Study Island and to support generalizing study findings. Although Study Island is designed primarily to provide students with practice on standards-based academic topics, it also has resources for teachers to help integrate Study Island into classroom instruction. Future research also should include student-level criteria and classroom-level criteria for inclusion in the Study Island group to better understand the impacts of Study Island on student achievement.

Introduction

Study Island is a flexible formative assessment tool that can be used for practice and instructional support on standards-based topics, test preparation, classroom assessment, and intervention support in four academic subjects: math, reading, science, and social studies (Edmentum, 2018). Study Island provides students with in-depth practice, immediate feedback on progress, and remediation when needed.

Study Island is structured around academic topics: a grouping of conceptual material within a subject area and grade aligned to one or more state standards. For each subject and grade, the topics cover the state standards in their entirety. Each Study Island topic includes practice items, lessons, and supporting instructional materials (e.g., games, flash cards, printable worksheets). If a student works through all topics in a grade and subject area, he/she will have practiced the full range of learning goals covered by the standards.

Students gain practice on topics by completing practice sessions. Practice sessions are online formative assessments composed of at least 10 multiple-choice practice items or technology enhanced practice items that are scored online. Students can earn virtual trophies by completing a practice session with a score of 70% correct or more. Students receive immediate feedback on their answer to each item in the practice session. Explanations of the correct answer appear if the student gets the item wrong, providing real-time remediation based on individual student performance. Students can revisit any items they answer incorrectly at the end of each practice session.

Study Island's real-time feedback and links to supportive instructional materials provide students with remediation based on student performance. For example, if the student scores 40% or lower on a practice session, Study Island routes the student to practice sessions for lower level topics that are building blocks for higher-level topics. Once the student demonstrates success at the lower level, they are routed back up to the higher level.

Study Island allows students to learn at their own pace, choosing which topics to work on in addition to topics assigned to them by their teachers. The program provides students with access to a variety of supporting instructional materials targeted to their learning needs. In this way, Study Island provides each student with an individual learning trajectory.

In addition to practice sessions and instructional materials targeted at individual students' needs, Study Island also provides a variety of features to support teachers and administrators. Teachers can customize Study Island practice to individual students by assigning students specific topics, selecting specific practice items, and customizing the score needed to earn virtual trophies. Teachers also can access lessons and supporting instructional materials aligned to each topic at the beginning and/or end of the practice activity.

Study Island's array of customizable reports allow teachers and administrators to identify student weaknesses and monitor student progress toward the mastery of standards-based content. If students fall behind or advance ahead, teachers can assign additional practice targeted to individual students while continuing to monitor progress. Teachers also can use information from Study Island to inform instruction or remediation within the classroom. With the formative information from Study Island integrated into classroom instruction, teachers can create individualized instructional assignments based on demonstrated student need to help students meet specific content benchmarks and standards.

One of the supporting features available to Study Island customers is Study Island’s NWEA™ MAP® Link. This feature allows customers to upload student MAP scores and integrate them with Study Island to help target specific Study Island content and automatically target learner’s individual needs. The program is available for Grade 2 through high school.

The following research question guided the design and analyses used in this study:

What is the impact of Study Island usage on student reading achievement in Grade 2 through Grade 8 relative to students who did not use Study Island?

Method

The purpose of this study was to provide a rigorous estimate of the impact of Study Island use on student achievement in reading. Rigorous studies of educational interventions and estimates of impacts are needed by state and local education agencies to select and implement interventions that improve academic outcomes for students (U.S. Department of Education, 2016).

The study was designed to meet the What Works Clearinghouse (WWC) 4.0 standards for quasi-experimental designs (QED) necessary to achieve a rating of *Meets WWC Group Design Standards with Reservations* (WWC, 2017). In meeting WWC standards, the study also was designed to meet the requirements of the Every Student Succeeds Act (ESSA) guidance for *Moderate Evidence* (U.S. Department of Education, 2016).

Data

Century Analytics obtained student data from Edmentum to conduct this study. These data included unique student identifiers, student grade level, identifiers for academic subjects, Study Island usage data, number of Study Island topics completed, and scores from the fall and winter administrations of the NWEA MAP assessment. No student demographic variables other than location were available for analysis. The study used data on fall MAP scores as the baseline measure and winter MAP scores as the outcome measure.

Study Island usage data. Data were compiled to summarize the number of distinct Study Island topics each student completed between the baseline (fall) MAP assessment and the outcome (winter) MAP assessment. Topics completed below and above students’ grade level were included to capture personalized remedial or enrichment practice using Study Island. Actual assessments of topics (e.g., diagnostic tests) were not included in the number of topics completed. Practice sessions that were not completed also were not included. Customer-generated topics or customer-generated practice questions, although very rare, were filtered out. Any topics where the student was practicing in Study Island’s “Game Mode,” in which the student is motivated to complete practice topics with brief, interstitial video games, were excluded to ensure that the total number of topics completed was based only on topic practice that was directed towards improving academic skills.

A small number of students using Spanish reading programs, approximately one tenth of one percent of the total participants, were removed from the sample because the topics focused on reading Spanish-language texts not aligned to the outcome used, MAP scores on English reading tests.

NWEA MAP data. The data provided also included NWEA MAP scores uploaded to Study Island by Study Island customers. Data cleaning rules were applied to ensure the quality of the test score data. Only students with a valid test score of 100 or greater were included. Student records were dropped if multiple MAP scores were uploaded for a given content area and testing period in order to identify one score per test and student, per season. The MAP score data were merged to the Study Island usage data by student ID.

Measures

Student achievement, both at baseline (fall) and outcome (winter) was measured using the NWEA MAP. The MAP is a computer adaptive interim assessment that measures achievement in mathematics, reading, and other subjects for students in Grades 2 through 12 (NWEA, 2011). The MAP is designed to measure growth over time with all items anchored to a vertically-aligned equal-interval scale (RIT scale) across grade levels to provide a longitudinal measure of student growth. As such, the MAP provides scores for student achievement even if the student is performing below or above typical grade level expectations. The MAP is aligned to state standards and covers a range of skills across grade levels. The MAP meets the WWC standards for outcomes in terms of validity and reliability. Because the MAP measures content aligned to national and state standards it is not over-aligned to the Study Island intervention.

Design

This study used a quasi-experimental design (QED) in order to meet WWC (4.0) standards with reservations. According to the WWC, a QED uses a non-random process to form the intervention and comparison conditions (WWC, 2017). The WWC allows groups to be formed using a variety of methods as long as the groups are mutually exclusive. That is, units (e.g., students or schools) can only be analyzed as a member of a single group. Further, in a QED, the WWC accepts assignment to the intervention based on observed characteristics. Assignment to study conditions for this study was conducted at the student level.

Students included in this study were selected from Grade 2 through Grade 8 students enrolled in a Study Island MAP Link program during the 2018-19 school year. The sample excludes students enrolled in the Study Island Benchmark program. To be eligible for inclusion in the study, student records had to meet the following eligibility criteria:

- Baseline MAP scores occurring between August 5, 2018 and September 16, 2018;
- Outcome MAP scores occurring between November 25, 2018 and January 20, 2019; and
- Duration between baseline and outcome test score dates of 84 to 140 days.

This study focused on students in Grades 2 through 8. Students in Kindergarten and Grade 1 were not able to be included in the study because the Study Island MAP Link program does not include mappings between NWEA scores and Study Island topics for kindergarten and first grade. Grades 9 and above were not included in the study because sufficient sample sizes were not available.

Students meeting the eligibility criteria were then selected for inclusion in either the Study Island (intervention) group or the comparison group. The Study Island group included students who completed a total number of distinct topics equal to at least half of the available topics for that grade (see Table 1). The rationale for this is that students who have completed half of the available topics approximately

halfway through the school year are on track to have completed all available topics by the end of the year and therefore have been exposed to a broad range of topics that cover grade level expectations. Topics from below the student’s grade level (remediation) or above the student’s grade level (enrichment) were included as contributing to the total number completed because these represent practice targeted to students’ individual skill levels. (See Appendix A for the number of lessons completed by students in the Study Island group.)

Table 1. Study Island (Treatment) Group Criterion

Grade	Minimum Reading Topics Completed
2	22
3	24
4	26
5	24
6	22
7	22
8	22

Students eligible for inclusion in the comparison group included those students who met the above eligibility criteria and had completed zero Study Island topics. Propensity score matching was used to match eligible comparison students to Study Island group students separately by grade. One-to-four matching was conducted using the nearest neighbor method and a logistic regression model with the fall MAP score and duration (time between fall and winter MAP test administrations) as the matching variables. Duration is an exogenous variable (i.e., not related to or affected by group selection) because it was based on the MAP test score data uploaded to Study Island by customers prior to the conduct of this study.

Students in both the Study Island group and the comparison group were administered the MAP and were exposed to any features or services available through the MAP. Thus, the difference between the two groups in terms of the educational instruction and implementation (i.e., the intervention contrast) was use of Study Island. In this way, this study design estimates the impact of student use of Study Island, defined as completion of the prescribed number of topics.

Baseline Equivalence

In order to meet WWC standards with reservation for a QED, baseline equivalence must be established for the analytic samples of the Study Island (intervention) and comparison groups. In addition, baseline equivalence needs to be established separately for each grade level included in the analyses. Finally, baseline equivalence must be established using a measure that meets WWC standards.

Baseline equivalence was established using fall MAP scores and duration separately for the analytics sample in each grade level. As described above in the Measures section, the MAP assessment meets WWC standards for baseline and outcome measures. All grades had baseline differences between students in the Study Island and comparison groups that were under the WWC threshold for baseline equivalence (i.e., ≤ 0.25 standard deviation) using a WWC accepted method for calculating baseline differences (WWC, 2017). Appendix B provides the statistics for baseline equivalence. Also included in Appendix B is information on the distribution of students in both groups across states.

Analyses and Results

Data were analyzed to estimate differences between students in the Study Island (intervention) group and the comparison group on the winter MAP reading assessment. Impact analyses were conducted using the following linear regression model fit to the data separately for each grade level.

$$Y_i = \beta_0 + \beta_1(\text{GROUP})_i + \beta_2(\text{FALLMAP})_i + \beta_3(\text{DURATION})_i + e_i$$

Where: Y_i is student i 's winter reading MAP score. β_0 is the regression adjusted comparison group mean. β_1 is the adjusted mean difference between the intervention and comparison groups, and GROUP represents the group status of student i coded as 0 = comparison and 1 = intervention. β_2 is the regression slope for the fall (baseline) MAP reading score and β_3 is the regression slope for duration. FALLMAP is student i 's fall MAP score in reading, and DURATION is the number of days between the fall MAP administration and winter MAP administration for student i . e_i is the residual for student i .

Impact analyses yielded statistically significant positive impacts for all grade levels in reading except Grade 8 (Table 2). Output from the regression analyses are provided in Appendix C. Adjusted mean differences between the intervention and comparison groups ranged from 1.11 points on the MAP RIT scale for Grade 4 to 4.37 points for Grade 2. These differences translate into effect sizes ranging from 0.30 for Grade 2 to 0.07 for Grade 4.

In addition to translating the impacts of Study Island use into effect sizes, the improvement index is another useful method to aid in the interpretation of the practical importance of impacts. The improvement index represents the difference in percentile rank at the mean (i.e., the 50th percentile) between the intervention group and the comparison group (WWC, 2017). The improvement index shows the expected change in percentile rank for an average comparison student if he or she had received the intervention. Percentile improvements for Study Island use in reading ranged from 2.79 for Grade 4 to 11.79 for Grade 2. Most improvement indexes were greater than 5. This is equivalent to a comparison student improving from the 50th percentile to the 55th percentile. An improvement index of 11.79 is equivalent to a comparison student improving from the 50th percentile to nearly the 62nd percentile.

Finally, normative data from the MAP can help interpret results from this study. These normative data show the average achievement levels by grade at the beginning, middle, and end of the school year based on data collected from 2011 to 2014 (NWEA, 2015). The MAP norms also provide educators with information on the normative growth on the MAP assessment.

The average achievement levels from the norms can be compared to the average achievement for the study groups on the winter administration of the MAP (Table 3). These data show that the students in this study had similar levels of achievement to the national norms. In the lower and upper grade levels, students participating in the study had slightly higher average scores than the norms, and in the middle grade levels study students had average achievement similar to the norms. Rather than indicating impacts, the comparison of achievement levels between the students in this study and the norms shows that the students in this study are likely to represent typical students in terms of achievement and not either especially low or especially high achieving students. This suggests this study's results are broadly generalizable.

Table 2. Impacts on Reading Achievement

	N	Mean MAP Score	SD	Adjusted Mean Difference (SE)	Pooled Standard Deviation	Effect Size	Improve Index
Grade 2							
Comparison	680	184.47	14.73	4.37***	14.46	0.30	11.79
Study Island	136	188.84	13.05	(0.75)			
Grade 3							
Comparison	1450	198.06	13.36	1.18*	13.46	0.09	3.59
Study Island	290	199.25	13.97	(0.47)			
Grade 4							
Comparison	1500	201.25	15.94	1.11*	15.74	0.07	2.79
Study Island	300	202.36	14.68	(0.48)			
Grade 5							
Comparison	1355	208.15	14.94	2.31***	14.74	0.16	6.36
Study Island	271	210.46	13.70	(0.49)			
Grade 6							
Comparison	1385	211.67	15.99	2.40***	15.84	0.15	5.96
Study Island	277	214.07	15.05	(0.51)			
Grade 7							
Comparison	610	218.35	14.84	2.12**	15.19	0.14	5.17
Study Island	122	220.47	16.86	(0.77)			
Grade 8							
Comparison	610	220.74	16.17	3.02***	15.87	0.19	7.53
Study Island	122	223.77	14.27	(0.79)			

Note. MAP scores are on the RIT scale. Numbers in table rounded to 2 decimals. M = mean. SD = standard deviation. SE = Standard error. Improve Index = Improvement index. Effect size measured as Hedges' *g*.

* = *p*-value < .05. ** = *p*-value < .01. *** = *p*-value < .001.

Table 3. Study Island Impacts and MAP Norms

Grade	Middle Year MAP Mean	Comparison Group Average Winter Score	Study Island Group Average Winter Score	Study Island Impacts
2	184.2	184.47	188.84	4.37***
3	195.6	198.06	199.25	1.18*
4	203.6	201.25	202.36	1.11*
5	209.8	208.15	210.46	2.31***
6	214.2	211.67	214.07	2.40***
7	216.9	218.35	220.47	2.12**
8	219.1	220.74	223.77	3.02***

Notes. All scores are on the RIT scale. Normative data from NWEA (2015).

* = *p*-value < .05. ** = *p*-value < .01. *** = *p*-value < .001.

In addition to the average achievement provided by the norms, data on normative growth can be used to aid in the interpretation of the impacts found in this study. NWEA calculated the average amount of growth students show on the MAP between the beginning of the year and middle of the year administrations (the same time period used in this study). This normative growth is shown in the first column in Table 4. These normative growth values can be compared to the impacts of Study Island to provide some additional meaning to the difference in scores between students who used Study Island and those in the comparison group. Dividing the impact by the normative growth shows the impacts in terms of a percentage of normative growth (last column of Table 4). Using this method, the impacts of Study Island represent between 16% and 159% of the typical growth seen by students during the first half of the school year. In grade 2 for example, students in the Study Island group scored 4.37 points higher than students in the comparison group. This 4.37 point difference represents 46% of the typical growth on the MAP for Grade 2 students. The impact relative to growth for Grade 8 is very high because Map growth in Grade 8 is only 1.9 points.

Table 4. MAP Normative Growth and Study Island Impacts by Grade

Grade	Beginning to Middle Year Normative Growth	Study Island Impacts	Impact as Percent of Normative Growth
2	9.5	4.37***	46%
3	7.3	1.18*	16%
4	5.4	1.11*	21%
5	4.2	2.31***	55%
6	3.2	2.40***	75%
7	2.5	2.12**	85%
8	1.9	3.02***	159%

Notes. All scores are on the RIT scale. Normative data from NWEA (2015).

* = p -value < .05. ** = p -value < .01. *** = p -value < .001.

Interpretation of the Study Island impacts as a percentage of normative growth should not be confused with a difference in growth between Study Island and comparison group students. For example, Study Island students in 2nd grade did **not** grow 46% more than comparison group students. Study Island and comparison group students were not compared in terms of growth but only compared on their average winter MAP assessment scores (regression adjusted).

Summary

Using a rigorous design and analytic strategy, this study estimated the impact of Study Island use on students' reading achievement as measured by the Measures of Academic Progress (MAP) at the level of rigor needed to meet WWC standards with reservations (WWC, 2017). Baseline equivalence was established between the Study Island (intervention) group and the comparison group. The Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) was used as the baseline measure and the outcome measure of student reading achievement. The MAP is a standardized educational assessment that meets WWC standards for validity and reliability. The baseline and outcome measures are aligned to national and state academic content standards and so are not over-aligned to the Study Island intervention. The study had no confounds.

The study also meets criteria set forth by the Every Students Succeeds Act (U.S. Department of Education, 2016). The Department of Education considers a quasi-experimental study to be "well-designed and well-implemented" if it receives a *Meets WWC Design Standards with Reservations* rating

or is of equal quality (U.S. Department of Education, 2016). The study also meets the ESSA criteria for statistically significant positive effects. These two aspects of the study mean it qualifies as providing Moderate evidence (Level 2) of Study Island’s effectiveness.

Study Island had a statistically significant impact on student reading achievement at every grade level. These impacts occurred on the winter administrations of the MAP assessment. Students who met the definition for inclusion in the Study Island (intervention) group—completion of at least half the grade level topics in reading between their fall and winter MAP assessments—showed greater gains in reading achievement than students who completed zero Study Island topics.

The results of this study suggest that students who use Study Island and complete topics in reading will make gains in achievement relative to students who do not complete any topics. The statistically significant gains made by students in the Study Island group over those students in the comparison group also suggest that practice on academic topics using Study Island helps students improve their reading knowledge and skills.

The data on topic completion in Appendix A show that many students completed many more than the minimum number of topics needed to be included in the Study Island group. In some cases, students completed more than double the number of topics available at their grade level. These data suggest that many students in the Study Island group spent a considerable amount of time practicing academic content, perhaps more than can be reasonably expected of typical students. But Study Island is designed to provide students with the resources to practice as much as they want. Few would doubt that time spent practicing academic content is positively correlated with test scores. The question unanswered by this study is whether Study Island is the cause of this time spent in practice, if the practice time is driven by student characteristics, or some combination of the two.

Limitations

This study has several limitations worth noting. It used a focused definition for the intervention group: students who had completed at least half of the available Study Island topics for a specific grade (Table 1). But Study Island involves much more than practice of academic topics, and use of Study Island cannot simply be measured as completion of topics. Study Island includes additional resources, such as lessons, games, flash cards, and printable worksheets; summarizing Study Island use as simply the number of 10-item practice sessions completed likely represents an incomplete picture of usage.

This study did not estimate the impact of use of any of Study Island’s additional resources. Students in the Study Island group, or in the comparison group, may or may not have used these additional resources to varying degrees. The use, or lack of use, of these resources could be influencing student achievement and, therefore, the findings in this study. Given the study design, it is possible that the key difference between the Study Island group students and the comparison group students was not completion of topics but use of additional resources, and the additional resources are, in fact, responsible for the higher levels of achievement among the Study Island students.

This study used a design sufficient to meet WWC standards with reservations. The Study Island students and comparison students were equivalent on baseline reading achievement (fall MAP scores). Students in the two groups also were equivalent in terms of the duration between their fall and winter MAP administrations. Both the students’ fall reading scores and their test durations were used as statistical

adjustments for estimating impacts on reading. No other student characteristics, however, were available or included in the study.

This study used a rigorous quasi-experimental design (QED) that is acceptable to meet WWC standards with reservations. Along with the statistically significant positive impacts, this study meets ESSA Level 2 standards. That said, the study was unable to control for student characteristics other than baseline (fall) achievement and duration. It is possible that students in the two groups differed on some key characteristic(s), and this difference was responsible for the difference in winter MAP scores rather than the difference in completion of Study Island topics. The lack of student demographic characteristics also limits the generalizability of the study results. It is unclear from this study what types of student were included in the Study Island group or if students of differing backgrounds experienced differing impacts from Study Island usage.

Students were the unit of assignment to the intervention and comparison groups. Study Island usage typically differs by students, so using students as the unit of assignment is appropriate. Study Island, however, is not used by students in isolation. Study Island also makes resources available to teachers and can be used to supplement regular classroom instruction. Teachers can use Study Island to assign individual students practice in areas of need, select specific practice items, and revise the score needed to advance to the next topic. Like the additional resources available to individual students, the features of Study Island for classroom and school use likely have an effect on student achievement. This study, however, was unable to estimate the impacts of any of these additional features of Study Island.

Further Research

This study provides a rigorous estimate of the impact of student completion of Study Island topics on student achievement in reading. Additional research is needed to understand how other features of Study Island impact student achievement. Future research also should address the limitations of this study. In addition to including student demographic characteristics as part of future analyses, further research also should examine other aspects of student usage and how these might impact student achievement. These could include student use of lessons, flash cards, and printable worksheets.

Although Study Island is designed primarily as a resource for students to gain practice and instructional support on standards-based topics as a means of test preparation, Study Island is still used within the classroom context. Integration of students' use of Study Island into classroom instruction likely magnifies the benefits of its use. A future study at the classroom level, that includes both student-level criteria and classroom-level criteria for inclusion in the Study Island group, could shed additional light on the impacts of Study Island on student achievement.

An unbiased estimate of Study Island's impact can only be provided by a randomized controlled trial (RCT). In this type of study, students (or classrooms) are randomly assigned to either use Study Island or conduct business as usual, creating two groups that are equivalent in expectation on all characteristics, known and unknown. This equivalence means any difference in achievement between the study groups can be attributed to Study Island usage. A well-conducted RCT eliminates the possibility that differences between intervention and comparison groups on outcome measures are caused by differences in baseline characteristics rather than the intervention under study, a limitation of the present study.

Finally, impact studies are best conducted in parallel with studies of implementation fidelity. Findings from the two studies complement each other and aid in the interpretation of results. Studies of

implementation fidelity inform the impact research by aiding in the definition of intervention group and communicating to the research audience what level of usage resulted in the impacts. Studies of impact inform implementation research by estimating impacts at different levels of implementation and helping to focus on how much usage is needed to produce statistically significant and meaningful increases in student achievement.

References

- Edmentum. (2018). *Research report: Examination of Study Island practice and benchmarks for Allentown School District*. Bloomington, MN: Author.
- Northwest Evaluation Association. (2011). *Technical manual for Measures of Academic Progress® (MAP®) and Measures of Academic Progress for Primary Grades (MPG)*. Portland, OR: Author.
- Northwest Evaluation Association. (2015). *2015 NWEA MAP Growth normative data*. Retrieved from <https://www.nwea.org/resource-library/research/2015-normative-data-3>.
- U.S. Department of Education, Office of Elementary and Secondary Education. (2016, September). *Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments*. Retrieved from <https://www2.ed.gov/policy/elsec/leg/essa/guidanceuseseseinvestment.pdf>.
- What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2017, October). *What Works Clearinghouse: Standards Handbook (Version 4.0)*. Retrieved from <http://whatworks.ed.gov>.

Appendix A
Study Island Reading Topics Completed

Table A.1. Number of Study Island Students Completing Reading Topics by Grade Level

Topics completed	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
22 Topics	11				27	19	16
23 Topics	7				34	11	17
24 Topics	10	27		18	22	12	16
25 Topics	7	26		11	22	9	14
26 Topics	13	31	38	24	26	5	9
27 Topics	9	23	31	17	18	6	7
28 Topics	10	19	34	15	21	9	6
29 Topics	11	12	18	14	18	5	5
30 Topics	10	14	19	13	13	4	7
31 Topics	8	16	24	9	11	6	3
32 Topics	1	11	17	11	9	0	3
33 Topics	1	15	12	8	11	3	6
34 Topics	3	5	13	9	6	4	1
35 Topics	3	9	15	10	9	2	3
36 Topics	3	5	14	11	4	0	2
37 Topics	4	5	9	6	3	5	0
38 Topics	1	7	9	6	7	1	1
39 Topics	2	6	8	5	4	3	0
40 Topics	5	4	3	7	2	2	1
1-50 Topics	14	33	15	41	9	12	2
51-60 Topics	2	11	12	26	1	3	2
61 or more Topics	1	11	9	10	0	1	1
Total	136	290	300	271	277	122	122

Note. Minimum number of topics completed varies by grade level. See Table 1.

Appendix B
Baseline Equivalence

Table B.1. Fall MAP Reading Baseline Equivalence by Grade

	N	M	SD	Adjusted Mean Difference	Pooled SD	Effect Size
Grade 2						
Comparison	680	176.16	15.34	0.99	15.18	0.07
Study Island	136	177.15	14.32			
Grade 3						
Comparison	1450	191.83	13.66	0.03	13.65	0.00
Study Island	290	191.87	13.58			
Grade 4						
Comparison	1500	196.43	16.80	0.33	16.69	0.02
Study Island	300	196.77	16.16			
Grade 5						
Comparison	1355	204.54	15.52	0.24	15.50	0.02
Study Island	271	204.78	15.41			
Grade 6						
Comparison	1385	209.67	16.08	0.39	15.96	0.02
Study Island	277	210.06	15.34			
Grade 7						
Comparison	610	216.31	15.71	-0.58	15.95	-0.04
Study Island	122	215.73	17.07			
Grade 8						
Comparison	610	218.66	16.58	-0.75	16.46	-0.05
Study Island	122	217.91	15.88			

Note. Numbers in table rounded to 2 decimals. M = mean. SD = standard deviation. Effect size measured as Hedges' *g*.

Table B.2. Duration Baseline Equivalence by Grade

	N	M	SD	Adjusted Mean Difference	Pooled SD	Effect Size
Grade 2						
Comparison	680	118.26	13.43	-0.43	13.15	-0.03
Study Island	136	117.83	11.62			
Grade 3						
Comparison	1450	111.13	14.84	-0.65	14.30	-0.05
Study Island	290	110.48	11.18			
Grade 4						
Comparison	1500	112.17	13.74	-0.18	13.47	-0.01
Study Island	300	111.99	12.02			
Grade 5						
Comparison	1355	115.00	14.42	-0.17	14.37	-0.01
Study Island	271	114.83	14.10			
Grade 6						
Comparison	1385	114.53	11.94	-0.44	11.77	-0.04
Study Island	277	114.08	10.88			
Grade 7						
Comparison	610	108.31	15.30	0.34	15.14	0.02
Study Island	122	108.66	14.29			
Grade 8						
Comparison	610	105.20	12.66	-1.08	13.14	-0.08
Study Island	122	104.11	15.34			

Note. Numbers in table rounded to 2 decimals. M = mean. SD = standard deviation. Effect size measured as Hedges' *g*.

Table B.3. Student Island and Comparison Group Students by Grade and State

State	Grade 2		Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	SI	C	SI	C	SI	C	SI	C	SI	C	SI	C	SI	C
AR									11		20	4	3	1
AZ			12	40	3	68	18	43	5	59	1	49	1	36
CA		1	23	1	14	2	35	1	87	4	7	2	28	
GA	125	37	144	53	149	34	143	19	277	48	112	1	80	4
IL	24	4	51	3	75	5	40		96	8	24	2	61	
IN	128	37	264	63	241	57	69	52	201	92	124	31	102	20
KS	23		71		31		35	14	11		38		16	
KY	54	3	230	16	202	63	219	5	168	6	120	6	145	6
MI		1	1		5		2	3	26		13	1	3	
MN							1							
MO	2	10	1				2		10	4	1			1
MS	4	2	10	16	8	9	4	18	3	4	3	1	19	
NE			4		28		48		8		5	14	3	
OH	233	38	478	57	558	33	502	86	399	35	128	5	116	46
OK												2		
PA	16		33	23	33	1	50	4	24	2	2			
SC	1	2	30	12	7	17		2	1		4		14	
SD	5		2		4		7		8		3		1	
TN							3		2					2
TX	65	1	96	6	142	11	148	4	26	10	5	4	18	6
WI							29	20	22	5				
Total	680	136	1,450	290	1,500	300	1,355	271	1,385	277	610	122	610	122

Note. SI = Study Island group. C = comparison group.

Appendix C
Regression Analysis Output

Table C.1. Grade 2 Estimates of Regression Coefficients

Parameter	Coefficient	Std. Error	<i>t</i> value	<i>p</i> -value	95% Conf. Int.	
Study Island	4.37	0.75	5.84	0.00	2.90	5.83
Fall MAP score	0.80	0.02	42.6	0.00	0.76	0.83
Duration	0.00	0.02	0.1	0.92	-0.04	0.04
Intercept	43.94	4.53	9.7	0.00	35.05	52.83

Std. Error = standard error

95% Conf. Int. = 95% confidence interval

Table C.1. Grade 3 Estimates of Regression Coefficients

Parameter	Coefficient	Std. Error	<i>t</i> value	<i>p</i> -value	95% Conf. Int.	
Study Island	1.18	0.47	2.5	0.01	0.26	2.11
Fall MAP score	0.83	0.01	63.77	0.00	0.80	0.85
Duration	0.06	0.01	5.06	0.00	0.04	0.09
Intercept	32.77	2.88	11.39	0.00	27.13	38.41

Std. Error = standard error

95% Conf. Int. = 95% confidence interval

Table C.1. Grade 4 Estimates of Regression Coefficients

Parameter	Coefficient	Std. Error	<i>t</i> value	<i>p</i> -value	95% Conf. Int.	
Study Island	1.11	0.48	2.32	0.02	0.17	2.05
Fall MAP score	0.83	0.01	77.39	0.00	0.81	0.85
Duration	0.00	0.01	0.14	0.89	-0.02	0.03
Intercept	38.56	2.62	14.72	0.00	33.42	43.70

Std. Error = standard error

95% Conf. Int. = 95% confidence interval

Table C.1. Grade 5 Estimates of Regression Coefficients

Parameter	Coefficient	Std. Error	<i>t</i> value	<i>p</i> -value	95% Conf. Int.	
Study Island	2.31	0.49	4.71	0.00	1.35	3.27
Fall MAP score	0.82	0.01	69.15	0.00	0.80	0.84
Duration	0.04	0.01	3.24	0.00	0.02	0.07
Intercept	35.75	2.74	13.07	0.00	30.38	41.11

Std. Error = standard error

95% Conf. Int. = 95% confidence interval

Table C.1. Grade 6 Estimates of Regression Coefficients

Parameter	Coefficient	Std. Error	<i>t</i> value	<i>p</i> -value	95% Conf. Int.	
Study Island	2.40	0.51	4.69	0.00	1.40	3.40
Fall MAP score	0.87	0.01	72.47	0.00	0.84	0.89
Duration	0.01	0.02	0.72	0.47	-0.02	0.04
Intercept	28.91	3.13	9.23	0.00	22.77	35.06

Std. Error = standard error

95% Conf. Int. = 95% confidence interval

Table C.1. Grade 7 Estimates of Regression Coefficients

Parameter	Coefficient	Std. Error	<i>t</i> value	<i>p</i> -value	95% Conf. Int.	
Study Island	2.12	0.77	2.77	0.01	0.62	3.62
Fall MAP score	0.82	0.02	45.25	0.00	0.78	0.85
Duration	0.02	0.02	0.94	0.35	-0.02	0.06
Intercept	39.36	4.17	9.43	0.00	31.16	47.55

Std. Error = standard error

95% Conf. Int. = 95% confidence interval

Table C.1. Grade 8 Estimates of Regression Coefficients

Parameter	Coefficient	Std. Error	<i>t</i> value	<i>p</i> -value	95% Conf. Int.	
Study Island	3.02	0.79	3.83	0.00	1.47	4.58
Fall MAP score	0.83	0.02	45.54	0.00	0.80	0.87
Duration	0.01	0.02	0.43	0.67	-0.04	0.05
Intercept	37.63	4.24	8.88	0.00	29.32	45.95

Std. Error = standard error

95% Conf. Int. = 95% confidence interval