

**Curriculum Associates** RESEARCH

# Impact of *i-Ready Personalized Instruction* on Massachusetts Comprehensive Assessment System English Language Arts Scores in Grades 4 and 5

---

Reading

Molly K. Duncan, Ph.D. and Madison A. Holzman, Ph.D.

Research Report, November 2023

# Impact of *i-Ready Personalized Instruction* on Massachusetts Comprehensive Assessment System English Language Arts Scores in Grades 4 and 5

---

## Reading

Molly K. Duncan, Ph.D. and Madison A. Holzman, Ph.D.

Research Report, November 2023

## Summary

This study analyzed the impact of using *i-Ready Personalized Instruction* on Grades 4 and 5 students' scores on the Massachusetts Comprehensive Assessment System English language arts assessments (MCAS ELA). Two different treatment groups of students were examined. The first treatment group included students who had used any amount of *i-Ready Personalized Instruction* and their comparison group counterparts. This portion of the analysis indicated that any use of *i-Ready Personalized Instruction* was associated with higher MCAS ELA scores compared to similar students who did not use *i-Ready Personalized Instruction*. The second treatment group included only students who met Curriculum Associates' guidance for *i-Ready Personalized Instruction* usage and their comparison group counterparts. This portion of the analysis indicated that the use of *i-Ready Personalized Instruction* for Reading, according to Curriculum Associates' guidance, is associated with improved MCAS ELA scores for Grade 4 students. The results for Grade 5 students were not statistically significant, possibly because the strict standards for inclusion resulted in a small sample of students. This study was designed to meet the standards for ESSA Level 2 evidence.

## Introduction

Schools have always needed efficient and effective methods of teaching students, but recent years have brought new challenges such as unfinished learning related to school closures and disruptions related to the COVID-19 pandemic. These challenges make it all the more important to find efficient solutions that can meet students at their level. One such solution is Curriculum Associates' *i-Ready Personalized Instruction*, a digital instructional supplement designed to provide instruction delivered at each individual student's level and target their unique needs. *i-Ready Personalized Instruction* is available for Mathematics and for Reading instruction at the levels of Grades K–8. This study focused on the use of *i-Ready Personalized Instruction* for reading in Grades 4 and 5. Most students begin the program by taking the *i-Ready Diagnostic*, which provides detailed information about what grade-level content the student is prepared for in several different domains. Students can progress through a queue of lessons selected by the program to target the areas where the student needs the most support, or they can work on lessons assigned by an educator. Most lessons focus on one or two discrete concepts and include instruction, interactive practice, and a short quiz that assesses whether the student is ready to move on to the next concept.

This study examined how the use of *i-Ready Personalized Instruction* for Reading impacted scores on the ELA portion of a statewide comprehensive examination, the MCAS. Notably, by applying the same methodology to two different treatment groups, this study provided information about the impact of *i-Ready Personalized Instruction* when used in any amount and when used according to Curriculum Associates' guidance.

This study uses an outcome measure that is of great interest to the education community: a statewide comprehensive exam. This study's focus on students in Grades 4 and 5 in 2021–2022 is also relevant to education decision-makers, as these students were in Grade 3 in the years that contained the most disruption from the COVID-19 pandemic (i.e., 2019–2020 and 2020–2021), and Grade 3 is widely considered to be a pivotal year in reading instruction (e.g., Hernandez, 2011; Shanahan et al., 2010; Stanley et al., 2018). Although previous research (Cook & Ross, 2022; Curriculum Associates, 2022; Holzman & Duncan, 2022; Randel et al., 2020) has examined the impact of *i-Ready Personalized Instruction* usage on the MCAS, the data used for this study were further removed from the drastic disruptions of the pandemic and included a more typical administration of both the pretest measure and the MCAS.

## Study Purpose and Research Questions

This study was designed to examine the efficacy of *i-Ready Personalized Instruction* for improving ELA scores on a statewide year-end assessment. The following research questions were addressed:

1. How were students in Grades 4 and 5 using *i-Ready Personalized Instruction* for Reading?
2. What was the impact of using *i-Ready Personalized Instruction* for Reading on MCAS ELA achievement in Grades 4 and 5?
3. What was the impact of using *i-Ready Personalized Instruction* for Reading according to Curriculum Associates' guidance on MCAS ELA achievement in Grades 4 and 5?

## Methodology

All analyses were conducted in R version 4.1.3 (R Core Team, 2022). Packages included several packages from tidyverse version 1.3.2 (Wickham et al., 2019) for general data cleaning and management; MatchIt version 4.5.0 (Ho et al., 2011) for matching; and lme4 version 1.1-34 (Bates et al., 2015) and lmerTest version 3.1-3 (Kuznetsova et al., 2017) for multilevel models.

### Data

The data for this study were collected from Curriculum Associates' in-house databases and directly from participating districts in Massachusetts. Curriculum Associates had access to information about individual students' use of *i-Ready Personalized Instruction* and performance on the *i-Ready Diagnostic*. Districts provided additional information about students' demographics and MCAS results.

Students' use of *i-Ready Personalized Instruction* during the 2021–2022 school year determined whether an individual student was in the *i-Ready Personalized Instruction* (i.e., treatment) group or the comparison group. Curriculum Associates tracks and reports a variety of metrics about the usage of *i-Ready Personalized Instruction*, including the date and time of each login, the amount of time spent in the program, and information about performance on lesson quizzes.

Scores from students' *i-Ready Diagnostic* taken in fall 2021 served as a pretest measure for this study. Scores on the *i-Ready Diagnostic* are vertically aligned and range from 100–800. The *i-Ready Diagnostic* is strongly correlated with the MCAS in the same subject. Correlations of spring *i-Ready Diagnostic* for Reading scores with MCAS ELA scores range from .80 to .82 (Curriculum Associates, 2023a).

The demographic variables used were student race/ethnicity, special education status, low-income status, and English Learner status. These variable names are used throughout this paper because it was the terminology used by the Massachusetts Student Information Management System (SIMS) during the 2021–2022 school year, and they do not necessarily reflect Curriculum Associates' preferred terminology. In some cases, the more detailed information captured by SIMS had to be collapsed into fewer categories for use in this analysis because the sample sizes of the more detailed categories were too small.

Students' MCAS ELA scores from spring 2022 was the outcome measure for this study. The MCAS is the comprehensive end-of-year exam taken by most students in Massachusetts. Scores range from 440 to 560 for all grades, and these scores are divided into four categories of placements. Score ranges and placements for the MCAS are detailed in [Table 1](#). The ELA assessment measures both reading and writing skills (Massachusetts Department of Elementary and Secondary Education, 2023).

Table 1. MCAS Scores and Placement Levels

Score	Placement Level
440-469	Not Meeting Expectations
470-499	Partially Meeting Expectations
500-529	Meeting Expectations
530-560	Exceeding Expectations

### Sample

Before matching took place, students who had missing data for any of the variables that would be used for the matching or analysis were dropped. Descriptive information about the students remaining in the full sample (i.e., the pool of students who were eligible to be included in the matching process) is presented in Table 2.

Table 2. Sample Sizes and Fall *i-Ready Diagnostic* Scores for Full Sample

Grade Level	Student Group	Number of Students	Mean Fall <i>i-Ready Diagnostic</i> Score	Standard Deviation of Fall <i>i-Ready Diagnostic</i> Score
Grade 4	No <i>i-Ready Personalized Instruction</i> Usage	709	536.31	49.99
	<i>i-Ready Personalized Instruction</i> Usage	2,032	524.51	57.18
	<i>i-Ready Personalized Instruction</i> Usage According to Guidance	308	526.90	63.16
Grade 5	No <i>i-Ready Personalized Instruction</i> Usage	806	557.00	48.77
	<i>i-Ready Personalized Instruction</i> Usage	1,871	548.67	56.50
	<i>i-Ready Personalized Instruction</i> Usage According to Guidance	180	557.13	63.63

Note: The *i-Ready Personalized Instruction* Usage According to Guidance group is a subset of the *i-Ready Personalized Instruction* Usage group. These students are included in the calculations for both relevant rows.

This study was concerned with describing usage patterns (Research Question 1) and estimating the impacts of two distinct usage patterns (Research Questions 2 and 3): the impact of any use of *i-Ready Personalized Instruction* and the impact of *i-Ready Personalized Instruction* when used according to Curriculum Associates’ guidance.

The primary purpose of the first research question was to provide context for the questions about the impact of *i-Ready Personalized Instruction*, and therefore it utilizes the same two samples of students that are included in the research questions about impact.

To estimate the overall impact of *i-Ready Personalized Instruction*, the pool of students for the treatment group for the second research question included any student who completed at least one lesson. The final sample of students for this portion of the analysis consists of pairs of students (i.e., one from the treatment group and one from the comparison group) who were selected by the matching model. These analyses are referred to as the “observed usage” analyses.

Because it was expected that at least some of the students in the treatment group from the observed usage analyses would use *i-Ready Personalized Instruction* very little, there was also a need to estimate the impact of the program when used according to guidance. Therefore, we conducted the same analyses a second time on the subset of students who used *i-Ready Personalized Instruction* according to Curriculum Associates’ guidance and a matched comparison group. Curriculum Associates recommends that students spend 30–49 minutes per week on average in *i-Ready Personalized Instruction* in each subject and maintain a 70% lesson pass rate. Curriculum Associates also recommends that students continue usage throughout the year, and 18 weeks is commonly used as a minimum cutoff (e.g., Curriculum Associates, 2022; Holzman & Duncan, 2022). Therefore, a student was considered to have used *i-Ready Personalized Instruction* according to guidance if they passed at least 70% of their lessons, logged in to the program during at least 18 calendar weeks, and used the program for an average of 30–49 minutes during those weeks. The final sample of students for this portion of the analysis consists of pairs of students (i.e., one from the treatment group and one from the comparison group) who were selected by a second matching model. These analyses are referred to as the “usage according to guidance” analyses.

### Matching

Matching models can reduce bias in quasi-experimental studies by reducing the pre-existing differences between the treatment and control groups (Fortson et al., 2015; Fortson et al., 2012; Rubin, 1974; Shadish et al., 2002). For consistency and interpretability, a single matching method was selected for both Grades 4 and 5 for the observed usage analyses, and a separate single matching method was selected for both Grades 4 and 5 for the usage according to guidance analyses. For each analysis, several models were tested that varied along specific parameters. The model that resulted in retention of the largest sample without exceeding the maximum baseline difference in the pretest score, while maintaining balance on most demographic variables, was selected. The maximum difference that was allowed on the fall *i-Ready Diagnostic*, which served as the pretest for this analysis, was .25 SD of the comparison group’s fall *i-Ready Diagnostic* scores. This is similar to the recommendations of What Works Clearinghouse and is the preferred method for calculating baseline differences according to Evidence for ESSA (Evidence for ESSA, 2023; What Works Clearinghouse, 2022).

All tested models were propensity score matching models that utilized a one-to-one, caliper-limited, nearest-neighbor match without replacement. The models were allowed to vary in the order in which treatment units were matched to comparison units (i.e., whether treatment students with

the smallest or largest propensity scores received a match first), the caliper size (.10-.30 SD, tested in increments of .05), and the inclusion of a variable that indicated the number of calendar days between September 6 and the completion of the student’s fall *i-Ready Diagnostic*. September 6 was the first day of school for many of the schools in this study, so this variable served as a proxy for the number of instructional days a student had received before the pretest measure.

For observed usage analyses, the final matching model used a .15 SD caliper, and the treatment students with the largest propensity scores received the first matches. For the usage according to guidance analyses, the final matching model used a .20 SD caliper, and the treatment student with the largest propensity score received the first match. For both analyses, the propensity score models predicted the propensity of the student to be in the treatment group and included as predictors the student’s race/ethnicity, disability status, low-income status, English Learner status, and fall *i-Ready Diagnostic* score, which was centered at the mean for the grade-level group in the original dataset that included unmatched students. Additionally, the matching models for the observed usage analyses included the instructional days variable.

Information about sample size and baseline equivalence on the fall *i-Ready Diagnostic* for each grade and subject and for each research question are presented in Table 3. Further information about sample demographics is presented in the [Appendix–Table A1](#).

**Table 3. Descriptive Statistics for Matched Samples**

	Total Students	Total Schools	<i>i-Ready Personalized Instruction</i> Group Mean Fall Diagnostic Score (SD)	Comparison Group Mean Fall Diagnostic Score (SD)	Glass’s Delta for Fall Diagnostic Scores
<b>Observed <i>i-Ready Personalized Instruction</i> Usage</b>					
<b>Grade 4</b>	1,414	38	528.43 (55.68)	536.08 (49.87)	-.15
<b>Grade 5</b>	1,538	34	554.81 (54.24)	556.23 (48.77)	-.03
<b><i>i-Ready Personalized Instruction</i> Usage According to Guidance</b>					
<b>Grade 4</b>	588	36	530.00 (62.66)	534.45 (52.45)	-.08
<b>Grade 5</b>	356	31	557.27 (63.14)	563.71 (54.91)	-.11

### Descriptive Information about Usage

To answer the first research question, the median of several metrics of usage was calculated for the observed usage treatment group and for the usage according to guidance treatment group. Additionally, we calculated the number of students who met Curriculum Associates’ guidance for *i-Ready Personalized Instruction* (i.e., passed at least 70% of lessons, averaged 30–49 minutes of usage per week, and used the program throughout the school year, which was operationalized as 18 distinct weeks of usage).

## Impact Model

To account for the clustered nature of the data, we used hierarchical linear models with students clustered within schools (Raudenbush & Bryk, 2002). We first calculated the intraclass correlation coefficient (ICC) to confirm the magnitude of clustering by fitting unconditional random-intercepts-only models for each grade (i.e., Grades 4 and 5) and analysis type (i.e., observed *i-Ready Personalized Instruction* usage and *i-Ready Personalized Instruction* usage according to guidance), resulting in four separate models. Each model predicted student-level MCAS ELA scores. The ICC represents the variability in scores that is explained by school membership. The ICC for each model is presented in Table 4. The magnitude of these ICCs reinforced the need for the use of hierarchical linear modeling.

Table 4. Descriptive Statistics for Matched Samples

	Between Schools Variance	Within Schools Variance	ICC
<b>Observed <i>i-Ready Personalized Instruction</i> Usage</b>			
<b>Grade 4</b>	46.64	371.24	.11
<b>Grade 5</b>	44.77	344.74	.11
<b><i>i-Ready Personalized Instruction</i> Usage According to Guidance</b>			
<b>Grade 4</b>	75.72	380.96	.17
<b>Grade 5</b>	51.57	379.04	.12

After estimating the unconditional models, all covariates were entered as a block for the final impact models. As with the estimation of the unconditional model, four separate models were fit for each grade by analysis type combination. Each model took the following form:

Level 1 (Student):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{usage of } i\text{-Ready Personalized Instruction}_{ij}) + \beta_{2j}(\text{centered fall } i\text{-Ready Diagnostic score}) + \sum \beta_{3j}(\text{race/ethnicity}_{ij}) + \beta_{4j}(\text{low-income status}_{ij}) + \beta_{5j}(\text{special education status}_{ij}) + \beta_{6j}(\text{English Learner status}_{ij}) + e_{ij}$$

Level 2 (School):

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$



where  $Y_{ij}$  represents the expected MCAS ELA score for student  $i$  in school  $j$ ;  $\beta_{0j}$  represents the school-level intercept, that is, the average MCAS ELA score for a student in school  $j$  for whom all predictors are zero or the centered value;  $\beta_{1j}$  represents the difference in MCAS ELA score associated with usage of *i-Ready Personalized Instruction*<sup>1</sup>;  $\beta_{2j}$  represents the difference in MCAS ELA score associated with a one-point increase in fall *i-Ready Diagnostic* score;  $\beta_{3j}$  is a vector of values that represent the difference in MCAS ELA scores associated with the race/ethnicity reported for the student; and  $\beta_{4j}$  through  $\beta_{6j}$  represent the differences in MCAS ELA scores associated with the binary variables indicated in the equation. Information about the values of these variables is available in Table 5.

Table 5. Values of Student-Level Variables

Variable	Values
<b><i>i-Ready Personalized Instruction</i> User</b>	Student did not use <i>i-Ready Personalized Instruction</i> (reference value), OR student used <i>i-Ready Personalized Instruction</i> .
<b>Fall <i>i-Ready Diagnostic</i> Score</b>	This value was centered at the mean for the full pre-matched sample for the relevant analysis. The mean values are reported in the <a href="#">Appendix—Table A2</a> .
<b>Race/Ethnicity</b>	Black (reference value) OR Hispanic OR White OR Other race/ethnicity
<b>Special Education Status</b>	Student is not a special education student (reference value), OR student is a special education student.
<b>Low-Income Status</b>	Student is not low-income (reference value), OR student meets one or more of the low-income criteria, which are: - eligible for free or reduced-price lunch - receives Transitional Aid to Needy Families benefits - eligible for food stamps
<b>English Learner Status</b>	Student is not an English Learner (reference value) OR Student is an English Learner

Note: The variable names and values used here and throughout this paper are used because they are reflective of the variable names and values aligned with the Massachusetts SIMS during the 2021-2022 school year. In some cases, more detailed categories had to be combined for these analyses due to sample size.

<sup>1</sup>Although the models for both analysis types were specified identically, the interpretation of this estimate changes depending on which sample was used for the analysis in question. When the model is fit to the data that include students who used any amount of *i-Ready Personalized Instruction* and their comparison group matches, this estimate represents the average impact of usage in any amount that was observed. However, when the model is fit to the data that include only students who used *i-Ready Personalized Instruction* according to Curriculum Associates’ guidance and their comparison group matches, this estimate represents the average impact of usage according to guidance.

For each model, several graphical checks were conducted for common violations of the assumptions of hierarchical linear modeling. None of these checks produced any cause for concern about assumption violations.

### Additional Analyses for Contextualizing the Impact

To better contextualize the impact of the use of *i-Ready Personalized Instruction*, several additional metrics were calculated. These metrics included a standardized effect size metric, an improvement index, and a hypothetical change in proficiency rates for the comparison group if they had used *i-Ready Personalized Instruction*. The use of several different metrics for interpreting the effect of the intervention can help educators and other stakeholders better understand the impact of an intervention and make more informed decisions about resource allocation (Lipsey et al., 2012).

The first metric that was calculated is Glass's Delta, a standardized effect size metric calculated by dividing the covariate-adjusted mean difference in the outcome by the standard deviation of the outcome in the comparison group. This was selected as the standardized effect size metric because it is commonly used and is the preferred metric of Evidence for ESSA and an acceptable metric by What Works Clearinghouse (Evidence for ESSA, 2023; What Works Clearinghouse, 2022).

Another metric that was calculated is known as the improvement index, which can be considered a translation of the effect size into percentile points (What Works Clearinghouse, 2022). What Works Clearinghouse provides detailed instructions for this calculation. The improvement indices reported in this paper were calculated using the unrounded effect size. Conceptually, the improvement index can be understood as the covariate-adjusted mean difference added to the median outcome score for the comparison group. The percentile of this new score is calculated, and the difference between the two percentiles is reported as the improvement index.

Lastly, we also calculated the number of additional comparison group students who would have placed proficient or higher on the MCAS if their score had increased by the amount of points attributed to *i-Ready Personalized Instruction* usage. This metric of additional *students* rather than *points per student* may be more meaningful for policymakers and district decision-makers.

## Results

In order to learn more about usage differences, we examined patterns of usage among the *i-Ready Personalized Instruction* users in both analysis types. Summary statistics about the usage patterns of both groups in each grade are presented in [Table 6](#). The summary statistics from the analysis of usage according to guidance are presented in this table to provide a fuller picture of the differences between groups, although it is important to remember that this group was selected based on their usage patterns according to the same metrics that are summarized below. For example, because students were only included in the usage according to guidance group if they averaged between 30 and 49 minutes per week of *i-Ready Personalized Instruction*, the median of average minutes per week in this group is, of course, between 30 and 49 minutes and is within the guidance range.

Table 6. Usage Metrics for *i-Ready Personalized Instruction* in Treatment Group for Both Samples

	Median Lesson Pass Rate	Median Weeks of Usage	Median Average Weekly Minutes	Percentage of Students Who Met Guidance
<b>Observed <i>i-Ready Personalized Instruction</i> Usage</b>				
<b>Grade 4</b>	82%	15	28.48	15.28%
<b>Grade 5</b>	80%	16	27.66	10.92%
<b><i>i-Ready Personalized Instruction</i> Usage According to Guidance</b>				
<b>Grade 4</b>	85%	26	37.78	100%
<b>Grade 5</b>	84%	25	38.59	100%

As Table 6 demonstrates, the average usage rates in the observed *i-Ready Personalized Instruction* group are lower than is recommended in both the frequency of usage (i.e., the number of weeks *i-Ready Personalized Instruction* was used) and the duration of usage within those weeks (i.e., the average minutes per week). This indicates that the majority of students in this group did not meet the minimum guidance.

### What Is the Impact of Observed Usage of *i-Ready Personalized Instruction*?

We first examined the impact of *i-Ready Personalized Instruction* when used as observed in the full sample. Separate models for Grades 4 and 5 were fit to a matched sample that included students who used *i-Ready Personalized Instruction* and their comparison group counterparts. In both models, the use of *i-Ready Personalized Instruction* had a positive and statistically significant association with higher MCAS ELA scores (see [Table 7](#)). On average, Grade 4 students who used *i-Ready Personalized Instruction* for Reading scored 2.69 points higher ( $p < .05$ ) on the MCAS ELA than similar comparison group students. This corresponds to a standardized effect size of .14 and an improvement index of 5.42. If each student in the comparison group had scored 2.69 points higher as a result of using *i-Ready Personalized Instruction*, an additional 4% of the comparison group would have been proficient. On average, Grade 5 students who used *i-Ready Personalized Instruction* for Reading scored 2.20 points higher ( $p < .05$ ) on the MCAS ELA than similar comparison group students. This corresponds to a standardized effect size of .12 and an improvement index of 4.65. If each comparison group student had scored 2.20 points higher as a result of using *i-Ready Personalized Instruction*, an additional 7% of the comparison group would have been proficient. For the full table of parameter estimates from these models, see the [Appendix—Table A3](#).

Table 7. Impact of Observed Usage of *i-Ready Personalized Instruction*

Grade	MCAS ELA Score Difference	$p$	Confidence Interval	Comparison Group SD	Glass's Delta	Improvement Index	Additional Proficient Comparison Group Students
Grade 4	2.69	.01	.74–4.63	19.72	.14	5.42	4%
Grade 5	2.20	.01	.48–3.92	18.83	.12	4.65	7%

### What Is the Impact of Usage of *i-Ready Personalized Instruction* According to Guidance?

Next, we examined the impact of using *i-Ready Personalized Instruction* according to guidance as this group represents the impact of *i-Ready Personalized Instruction* when used as intended. Separate models for Grades 4 and 5 were fit with a smaller subset of students, specifically students who had used *i-Ready Personalized Instruction* according to Curriculum Associates' guidance and their matched comparison group counterparts. In both models, the use of *i-Ready Personalized Instruction* was associated with higher MCAS ELA scores, but this association was statistically significant only in the Grade 4 model (see [Table 8](#)). On average, Grade 4 students who used *i-Ready Personalized Instruction* for Reading according to guidance scored 4.13 points higher ( $p < .05$ ) than similar students who did not use *i-Ready Personalized Instruction* for reading. This corresponds to a standardized effect size of .20 and an improvement index of 7.99. If each comparison group student had scored 4.13 points higher as a result of using *i-Ready Personalized Instruction* according to guidance, an additional 8% of the comparison group would have been proficient. On average, Grade 5 students who used *i-Ready Personalized Instruction* for Reading according to guidance scored 2.16 points higher ( $p > .05$ ) than similar students who did not use *i-Ready Personalized Instruction* for Reading. This corresponds to an effect size of .11 and an improvement index of 4.26. If each comparison group student had scored 2.16 points higher, an additional 6% of the comparison group would have been proficient. While these results are not statistically significant, recall that only a small number of Grade 5 students used *i-Ready Personalized Instruction* according to guidance. Given such small sample sizes, models may not allow differences in outcomes to reach traditional levels of statistical significance. For the full table of parameter estimates from these models, see the [Appendix—Table A4](#).

Table 8. Impact of Usage of *i-Ready Personalized Instruction* According to Guidance

Grade	MCAS ELA Score Difference	$p$	Confidence Interval	Comparison Group SD	Glass's Delta	Improvement Index	Additional Proficient Comparison Group Students
Grade 4	4.13	.02	1.03–7.23	20.50	.20	7.99	8%
Grade 5	2.16	.16	-.68–5.07	20.16	.11	4.26	6%

## Discussion

The purpose of this study was to examine rates of usage of *i-Ready Personalized Instruction* for Reading and to estimate the impact of *i-Ready Personalized Instruction* for Reading on a statewide comprehensive exam. This study demonstrated evidence of the positive impact of *i-Ready Personalized Instruction* for Reading when used in any amount and when used according to Curriculum Associates' guidance on a statewide comprehensive ELA exam.

The examination of usage rates revealed that for the observed usage treatment group, the majority of the sample did not meet Curriculum Associates' guidance for usage. In fact, only around 15% of Grade 4 students and 11% of Grade 5 students used *i-Ready Personalized Instruction* according to Curriculum Associates' guidance (see [Table 6](#)). On average, students seem to have used the program far less than recommended. Although the individual metrics for median weekly time and the number of weeks of use of *i-Ready Personalized Instruction* are near the minimum guidance, when considered together, the different metrics suggest that the average student may have received far less time with the program than is advised. For example, the median weeks of usage and the median average weekly minutes of usage for a Grade 4 student in this sample would result in around 430 minutes of *i-Ready Personalized Instruction* usage during the school year (i.e., 15 weeks of use at about 28 minutes of use on average per week), whereas the minimum amount of recommended usage would have resulted in 540 minutes. Because these are the median rates, we know that half the sample used *i-Ready Personalized Instruction* for less time per week or fewer weeks. These low rates of usage also seemed to hold true in the larger sample of *i-Ready Personalized Instruction* users (that is, the pre-match sample), where 15% of *i-Ready Personalized Instruction* users in Grade 4 and 10% of *i-Ready Personalized Instruction* users in Grade 5 used the program according Curriculum Associates' guidance (see [Table 2](#)). This pattern of low usage in Grade 4 contributed to the weaker effect sizes in the observed usage analysis compared to the usage according to guidance analysis, and in Grade 5 it resulted in a small sample size for the usage according to guidance analyses. The medium effect sizes (Kraft, 2020) in the observed usage analyses are all the more remarkable because they are evident in a sample with relatively low average rates of usage. On average, even low rates of usage of *i-Ready Personalized Instruction* can make a meaningful difference for students, and usage according to Curriculum Associates' guidance may be especially beneficial.

Other studies of *i-Ready Personalized Instruction* usage in typical conditions show higher average usage rates than were found in this study (Curriculum Associates, 2021; Curriculum Associates, 2023b; Holzman & Duncan, 2022), which indicates that Curriculum Associates' usage guidance is generally attainable. Although examining possible reasons that usage was lower in this sample is outside of the scope of this study, it is important to remember these lower average usage rates when interpreting the results of the observed usage analysis and to remember the limited sample sizes when interpreting the results of the usage according to guidance analysis.

Despite the relatively low usage rates in the observed usage sample, the use of *i-Ready Personalized Instruction* was associated with higher MCAS scores. In Grade 4, the impact of *i-Ready Personalized Instruction* was positive and statistically significant in both analyses, though usage according to guidance produced a larger effect size. In Grade 5, the analysis of observed usage of *i-Ready Personalized Instruction* indicated that *i-Ready Personalized Instruction* had a positive and statistically significant impact. The analysis of *i-Ready Personalized Instruction* according to guidance also produced a positive result in Grade 5, although the effect was not quite strong enough to reach statistical significance. However, the relatively small sample size means that this analysis had limited statistical power to detect an effect.

This study utilized a rigorous quasi-experimental design, but future research could further strengthen and deepen the quality of evidence and the precision of the program impact estimate. Specifically, future research could collect additional information about instruction of reading and ELA. We were not able to obtain information about what instructional materials comparison group students used during the time that the *i-Ready Personalized Instruction* group was using *i-Ready Personalized Instruction*. Comparison group students may have dedicated this time to using a different digital instruction product, working one on one with an educator, participating in whole class instruction, or studying an entirely different subject. Although this study indicates that *i-Ready Personalized Instruction* has a positive impact on student test scores, more information about how comparison students were spending their instructional time may help educators make more informed decisions about how to fit *i-Ready Personalized Instruction* into the time allotted for the typical school day. Furthermore, because *i-Ready Personalized Instruction* is designed as an instructional supplement, collecting information about what it supplements may provide more insight into its optimal use.

This study's strong quasi-experimental design comparing groups who were similar on a baseline measure of achievement meets the requirements for ESSA Level 2 evidence, and it demonstrates that students who used *i-Ready Personalized Instruction* for Reading performed better on a statewide end-of-year exam compared to similar students who did not use *i-Ready Personalized Instruction*. Remarkably, this finding was evident even in a sample with average usage rates that were lower than Curriculum Associates' guidance and lower than has been observed in previous studies. For an individual student, *i-Ready Personalized Instruction* can be instrumental in providing instruction in important concepts that are tailored to the student's current understanding and focused on the domains with which the student needs the most support. For a classroom, school, or district, the value of *i-Ready Personalized Instruction* is evident in the number of additional students who could be placing proficient on a statewide exam. In this difficult educational landscape, having solutions that help more students reach grade-level proficiency is invaluable.

## Full Report References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Cook, M., & Ross, S. M. (2022). *The impact of i-Ready Personalized Instruction with fidelity on 2021 MCAS ELA achievement. Final report.* <https://files.eric.ed.gov/fulltext/ED622942.pdf>
- Curriculum Associates. (2021). *Impact of i-Ready Personalized Instruction time and lesson pass rates on student learning gains.* <https://www.curriculumassociates.com/-/media/mainsite/files/i-Ready/iready-instruction-dosage-study-2021.pdf>
- Curriculum Associates. (2022). *The impact of i-Ready Personalized Instruction during the 2020-2021 school year. Evidence to support historically marginalized student groups.* <https://www.curriculumassociates.com/-/media/mainsite/files/i-Ready/iready-efficacy-during-covid-2022.pdf>
- Curriculum Associates. (2023a). *The relationship between i-Ready Diagnostic and the 2022 Massachusetts Comprehensive Assessment System (MCAS).* <https://cdn.bfldr.com/LS6J0F7/as/bjj59t7mj5z5gbvt6wcsg7q/iready-full-correlation-brief-massachusetts>
- Curriculum Associates. (2023b). *Variation in attainment of i-Ready Stretch Growth® in reading and mathematics: What can we learn from students' usage of i-Ready Personalized Instruction?* [https://cdn.bfldr.com/LS6J0F7/at/ps8zr4p63mb6xgjxhmg39r2k/Stretch\\_Growth\\_Technical\\_Doc\\_-\\_Research\\_and\\_Usage.docxn](https://cdn.bfldr.com/LS6J0F7/at/ps8zr4p63mb6xgjxhmg39r2k/Stretch_Growth_Technical_Doc_-_Research_and_Usage.docxn)
- Evidence for ESSA. (2023). *Evidence for ESSA: Standards and procedures: Version 2.0.* [https://evidenceforessa.org/wp-content/themes/evidenceforessa/uploads/FINAL\\_Standards\\_and\\_Proc%2005\\_18\\_23.pdf](https://evidenceforessa.org/wp-content/themes/evidenceforessa/uploads/FINAL_Standards_and_Proc%2005_18_23.pdf)
- Fortson, K., Verbitsky-Savitz, N., Kopa, E., & Gleason, P. (2012). *Using an experimental evaluation of charter schools to test whether nonexperimental comparison group methods can replicate experimental impact estimates* (NCEE Technical Methods Report 2012-4019). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education. <https://ies.ed.gov/ncee/pubs/20124019/pdf/20124019.pdf>
- Fortson, K., Gleason, P., Kopa, E., & Verbitsky-Savitz, N. (2015). Horseshoes, hand grenades, and treatment effects? Reassessing whether nonexperimental estimators are biased. *Economics of Education Review*, 44, 100-113. <https://doi.org/10.1016/j.econedurev.2014.11.001>
- Hernandez, D. J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation.* The Annie E. Casey Foundation. <https://files.eric.ed.gov/fulltext/ED518818.pdf>
- Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28. <https://doi.org/10.18637/jss.v042.i08>

- Holzman, M. A., & Duncan, M. K. (2022). *Impact of i-Ready Personalized Instruction on the Massachusetts Comprehensive Assessment System achievement for Grade 5 in English language arts and mathematics*.  
<https://cdn.bfldr.com/LS6J0F7/at/xqqfvrjkf6f96mkq3gbqh3p7/iready-impact-on-mcas-report-2022.pdf>
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4). <https://doi.org/10.3102/0013189X20912798>
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.  
<https://doi.org/10.18637/jss.v082.i13>
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. Institute of Education Sciences, National Center for Special Education Research.  
<https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf>
- Massachusetts Department of Elementary and Secondary Education. (2023). *Massachusetts Comprehensive Assessment System: English language arts test design and development*. DESE.  
<https://www.doe.mass.edu/mcas/tdd/ela.html>
- R Core Team. (2022). *R (version 4.1.3)*. The R Project for Statistical Computing. <https://www.R-project.org>
- Randel, B., Swain, M., Norman Dvorak, R., Spratto, E., & Prendez, J. Y. (2020). *Impact evaluation of Reading “i-Ready” for striving learners using 2018–19 data: Final report*.  
<https://eric.ed.gov/?id=ED610441>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*. Sage Publications.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701. <https://doi.org/10.1037/h0037350>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage Learning.
- Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade (NCEE 2010–4038)*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education. <https://files.eric.ed.gov/fulltext/ED512029.pdf>
- Stanley, C. T., Petscher, Y., & Catts, H. (2018). A longitudinal investigation of direct and indirect links between reading skills in kindergarten and reading comprehension in tenth grade. *Reading and Writing*, 31, 133–153. <https://doi.org/10.1007/s11145-017-9777-6>



What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook: Version 5.0*. [https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final\\_WWC-HandbookVer5.0-0-508.pdf](https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5.0-0-508.pdf)

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pederson, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *The Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>

## Appendix

Table A1. Information about Student Samples after Matching

Grade	Group	Total Students	Percentage Black	Percentage Hispanic	Percentage White	Percentage Other Race/Ethnicity	Percentage Students with Disabilities	Percentage Low-Income Students	Percentage English Learner
<b>Observed <i>i-Ready Personalized Instruction</i> Usage</b>									
Grade 4	No <i>i-Ready Personalized Instruction</i> Usage	707	12.59%	16.83%	60.25%	10.33%	21.36%	48.09%	8.49%
	<i>i-Ready Personalized Instruction</i>	707	11.46%	19.66%	56.44%	12.45%	21.50%	49.08%	10.04%
Grade 5	No <i>i-Ready Personalized Instruction</i> Usage	769	12.22%	19.64%	59.69%	8.45%	19.77%	45.25%	5.98%
	<i>i-Ready Personalized Instruction</i>	769	11.57%	23.93%	54.23%	10.27%	20.81%	53.32%	6.89%
<b><i>i-Ready Personalized Instruction</i> Usage According to Guidance</b>									
Grade 4	No <i>i-Ready Personalized Instruction</i> Usage	294	9.86%	33.67%	40.48%	15.99%	16.67%	51.02%	9.86%
	<i>i-Ready Personalized Instruction</i> According to Guidance	294	8.50%	33.67%	41.84%	15.99%	19.39%	55.78%	10.54%
Grade 5	No <i>i-Ready Personalized Instruction</i> Usage	178	14.04%	32.58%	39.89%	13.48%	15.73%	44.94%	10.11%
	<i>i-Ready Personalized Instruction</i> According to Guidance	178	12.36%	34.83%	39.89%	12.92%	15.17%	48.88%	12.36%

Note: Demographic category names are based on the categories that were used by the Massachusetts Department of Education in the 2022–2023 school year and do not necessarily reflect Curriculum Associates’ preferred terminology.

Table A2. Mean Scores Used for Centering in Hierarchical Linear Models

	Total Students	Mean Fall <i>i-Ready Diagnostic</i> Score
<b>Observed <i>i-Ready Personalized Instruction</i> Usage</b>		
<b>Grade 4</b>	2,741	527.56
<b>Grade 5</b>	2,677	551.18
<b><i>i-Ready Personalized Instruction</i> Usage According to Guidance</b>		
<b>Grade 4</b>	1,017	533.46
<b>Grade 5</b>	986	557.02

Table A3. Parameter Estimates for Observed Usage of *i-Ready Personalized Instruction Models*

Grade	Estimate	Category	Unstandardized Estimate	Standard Error	<i>t</i>	<i>p</i>	
Grade 4	(Intercept)		492.47	1.50	327.79	<.01	
	<i>i-Ready Personalized Instruction Usage</i>		2.69	.99	2.71	.01	
	Fall <i>i-Ready Diagnostic Score</i>		.26	.01	34.43	<.01	
	Race/Ethnicity	Hispanic		-.03	1.29	-.02	.98
		White		.55	1.24	.45	.66
		Other		2.54	1.41	1.81	.07
	Low-Income Status		-2.54	.75	-3.38	<.01	
Disability Status		-5.49	.88	-6.24	<.01		
English Learner Status		-4.27	1.21	-3.54	<.01		
Grade 5	(Intercept)		493.53	1.34	368.98	<.01	
	<i>i-Ready Personalized Instruction Usage</i>		2.20	.88	2.49	.01	
	Fall <i>i-Ready Diagnostic Score</i>		.27	.01	38.68	<.01	
	Race/Ethnicity	Hispanic		-0.75	1.13	-0.67	0.50
		White		.52	1.08	.48	.63
		Other		1.73	1.33	1.30	.19
	Low-Income Status		-3.33	.66	-5.07	<.01	
Disability Status		-5.12	.79	-6.53	<.01		
English Learner Status		-1.36	1.31	-1.04	.30		

Table A4. Parameter Estimates for *i-Ready Personalized Instruction* Usage According to Guidance Models

Grade	Estimate	Category	Unstandardized Estimate	Standard Error	<i>t</i>	<i>p</i>	
Grade 4	(Intercept)		491.44	2.29	214.89	< .01	
	<i>i-Ready Personalized Instruction</i> Usage		4.13	1.60	2.58	.02	
	Fall <i>i-Ready Diagnostic</i> Score		.26	.01	22.66	< .01	
	Race/Ethnicity	Hispanic		2.15	1.96	1.09	.27
		White		4.29	2.06	2.08	.04
		Other		3.72	2.16	1.72	.09
	Low-Income Status		-2.83	1.19	-2.39	.02	
Disability Status		-5.25	1.52	-3.46	< .01		
English Learner Status		-2.32	1.81	-1.28	.20		
Grade 5	(Intercept)		495.66	2.32	213.77	< .01	
	<i>i-Ready Personalized Instruction</i> Usage		2.16	1.48	1.46	.16	
	Fall <i>i-Ready Diagnostic</i> Score		.26	.01	18.14	< .01	
	Race/Ethnicity	Hispanic		.10	2.08	.05	.96
		White		.66	2.18	.30	.76
		Other		3.64	2.49	1.46	.15
	Low-Income Status		-2.62	1.46	-1.80	.07	
Disability Status		-3.84	1.88	-2.04	.04		
English Learner Status		-1.65	2.41	-0.68	.49		