Impact of *i-Ready*® *Personalized Instruction* on the Massachusetts Comprehensive Assessment System Achievement for Grade 5 in English Language Arts and Mathematics

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

Curriculum Associates Research Report No. 07| March 2022

Executive Summary

With the onset of the COVID-19 pandemic also came an increase in the prevalence of digital learning solutions. *i-Ready Personalized Instruction* is one digital learning tool students and teachers can use to support learning. It is a supplemental, standards-aligned program designed to provide students with content geared toward what students know and what they need to learn next, and it is available for Reading and for Mathematics in Grades K–8. Students can complete lessons anywhere at any time they have a device that connects to the internet, which makes *i-Ready Personalized Instruction* a flexible tool that can be incorporated into a wide variety of learning models.

The current study evaluated the impact of *i-Ready Personalized Instruction* usage on achievement in English Language Arts (ELA) and Mathematics as measured by the Massachusetts Comprehensive Assessment System (MCAS) for Grade 5 students in six Massachusetts districts during the 2020–2021 school year. Students who used *i-Ready Personalized Instruction* and students who completed the *i-Ready Diagnostic* but did not use *i-Ready Personalized Instruction* were matched to create similar instruction and comparison groups based on a series of student-level covariates. Specifically, Coarsened Exact Matching (CEM) was used to match students based on fall *i-Ready Diagnostic* scores and several other important student characteristics. Impacts were measured using a two-level hierarchical linear model (HLM) that both accounted for the clustered nature of the data and allowed for the inclusion of variables at the student and school levels. This study was rigorously designed to meet the requirements of Every Student Succeeds Act (ESSA) Level 2 evidence, and it demonstrates that usage of *i-Ready Personalized Instruction* is related to significantly greater scores on the MCAS.

This report is structured into two parts: 1) primary research questions and analyses that answered questions related to the effect of *i-Ready Personalized Instruction* from an intent-to-treat perspective and 2) exploratory research questions and analyses that answered questions related to the effect of *i-Ready Personalized Instruction* from a treatment-on-the-treated perspective.

The primary conclusions of this report include the following:

- Students who used *i-Ready Personalized Instruction* for Reading scored, on average, 2.50 points higher on the ELA portion of the MCAS than similar students who did not use *i-Ready Personalized Instruction* for Reading.
- Students who used *i-Ready Personalized Instruction* for Mathematics scored, on average, 4.58 points higher on the Mathematics portion of the MCAS than similar students who did not use *i-Ready Personalized Instruction* for Mathematics.
- These score differences were apparent even though most of the *i-Ready Personalized Instruction* group did not meet Curriculum Associates' recommended usage guidelines, which demonstrates the program can be beneficial in a practical setting with varying implementations and less-than-optimal usage.

The exploratory conclusions of this report include the following:

- Students who used *i-Ready Personalized Instruction* for Reading and met recommended instruction usage guidelines scored, on average, 7.96 points higher on the ELA portion of the MCAS than similar students who did not use *i-Ready Personalized Instruction* for Reading.
- Students who used *i-Ready Personalized Instruction* for Mathematics and met recommended instruction usage guidelines scored, on average, 9.50 points higher on the Mathematics portion of the MCAS than similar students who did not use *i-Ready Personalized Instruction* for Mathematics.

Introduction

When schools shut down in spring 2020 due to the onset of the COVID-19 pandemic in the United States, students and families were forced to adjust to at-home learning, and educators had to get creative to keep students engaged in this new learning environment. Pre-COVID-19, prevalence of digital learning was already on the rise within classrooms (Globe Newswire, 2019). Since the onset of the COVID-19 pandemic, digital learning has expanded into classrooms in unprecedented ways and with the full expectation that the classroom landscape has changed and digital learning solutions are here to stay (Li & Lalani, 2020).

It is imperative to note that access to technology and broadband internet is a major hindrance for the use and benefits of digital learning solutions, and the lack of technological infrastructure compounds impacts for already marginalized groups of people (Richards et al., 2021; Rome & Lay, 2022; The Pew Charitable Trusts, 2020). Although it is outside of the purview of this report to delve into solutions for sparse technological infrastructure, it would be an oversight not to acknowledge this fact in this report. Nevertheless, the use of digital learning solutions is on the rise and, some communities might say, is commonplace in today's educational framework. With the opportunity to be completed in the classroom, at home, or at a library, digital learning solutions offer a flexible and responsive option in diverse learning environments where technology is supported, whether that be during a pandemic or in a non-traditional classroom.

One such digital learning solution is Curriculum Associates' *i-Ready*. *i-Ready* is a comprehensive package of assessments, instructional programs, and educator resources designed to: 1) identify students' reading and mathematics abilities in relation to grade-level knowledge and 2) support all learners in mastering grade-level skills and knowledge. One component of the suite is the *i-Ready Diagnostic*, which is an adaptive, online benchmark assessment designed to help learners, their families, and educators understand students' skills and knowledge in reading or mathematics. The *i-Ready Diagnostic* informs both normative- and criterion-referenced placements for students. Specifically, results from the *i-Ready Diagnostic* place students in relation to grade-level standards and national norms. The *i-Ready Diagnostic* may be completed three times per school year, allowing students' progress to be tracked throughout the year. Because the *i-Ready Diagnostic* is adaptive, it efficiently and accurately pinpoints students' abilities in reading or mathematics, integrating seamlessly with *i-Ready Personalized Instruction*, which offers a collection of supplemental, standards-aligned lessons. *i-Ready Personalized Instruction* lessons are fully online, consist of standards-aligned content followed by short quizzes, and are designed as mastery-based lessons. The intention of the lessons is for students to master key knowledge and skills in each lesson and for the sequence of lessons to scaffold students to grade-level knowledge and skills. Based on their *i-Ready Diagnostic* placement, students are routed to *i-Ready Personalized Instruction* lessons that align with their skills and knowledge in Reading and Mathematics domains. By providing students with supplemental instruction geared toward what they need to learn next, students have the opportunity to practice and master critical skills to move them forward academically.

Given the unfinished teaching and learning associated with COVID-19 (Curriculum Associates, 2021b; Dawson, 2021; Kuhfeld et al., 2020) on top of already persistent learning gaps, it is more important than ever that educators employ evidence-supported learning programs in their classrooms. Educators must understand the efficacy of programs in our current non-traditional learning environment, and they need evidence that their trusted programs continue to support student achievement outcomes in our COVID-19 learning world. The purpose of the current study was to generate evidence that further establishes the effectiveness of *i-Ready Personalized Instruction* on student achievement, particularly in a year of teaching and learning disruptions.

Study Purpose and Research Questions

To provide educators with the most robust evidence about the continued effectiveness of *i-Ready Personalized Instruction*, this study was designed to answer questions about the effectiveness of *i-Ready Personalized Instruction* in typical practice with varying usage and implementations as well as when used with fidelity according to Curriculum Associates' established usage recommendations.

First, the two following primary research questions were addressed. In a year of learning disruptions related to COVID-19:

1. What was the impact of *i-Ready Personalized Instruction* for Reading on MCAS ELA achievement for Grade 5 students?
2. What was the impact of *i-Ready Personalized Instruction* for Mathematics on MCAS Mathematics achievement for Grade 5 students?

These questions were designed to evaluate the effectiveness of *i-Ready Personalized Instruction* in typical practice with varying usage and implementations. However, Curriculum Associates has established instruction usage recommendations that are related to increased gains on achievement outcomes (Curriculum Associates, 2021a; Curriculum Associates, in press). Specifically, for a given subject, Curriculum Associates recommends students use *i-Ready Personalized Instruction* for at least 30 minutes per week consistently throughout the school year. It is also recommended that students maintain an average pass rate of 70% for the quizzes at the end of lessons.

Because meeting these recommendations is central to conversations with instruction users and we know meeting usage recommendations tends to provide an added learning benefit for students, it was important to not only understand the extent to which *i-Ready Personalized Instruction* was used by students in this study but also to explore the impact of using *i-Ready Personalized Instruction* according to established usage recommendations on MCAS achievement.

To contextualize and extend the findings from research questions one and two, additional exploratory research questions were addressed.

3. To what extent were instruction students using *i-Ready Personalized Instruction* for Reading?
4. To what extent were instruction students using *i-Ready Personalized Instruction* for Mathematics?
5. What was the impact of *i-Ready Personalized Instruction* for Reading on MCAS ELA achievement for Grade 5 students who used *i-Ready Personalized Instruction* according to established usage recommendations?
6. What was the impact of *i-Ready Personalized Instruction* for Mathematics on MCAS Mathematics achievement for Grade 5 students who used *i-Ready Personalized Instruction* according to established usage recommendations?

Primary Analyses

This section of the report includes methodology and results related to the primary research questions. The primary analyses were designed to answer the following questions for Grade 5 students:

1. What was the impact of *i-Ready Personalized Instruction* for Reading on MCAS ELA achievement?
2. What was the impact of *i-Ready Personalized Instruction* for Mathematics on MCAS Mathematics achievement?

Methodology

Unless otherwise noted, all analyses were completed using SAS® 9.4M7 software with Enterprise Guide version 7.1 (SAS Institute, 2017).

Data

Data were collected for Grade 5 students in six Massachusetts school districts in two stages. First, *i-Ready Diagnostic* scale scores and student demographic data, such as race, ethnicity, and special education status, were collected. This first round of data was collected in fall 2020. Second, ELA and Mathematics 2021 MCAS (Massachusetts Department of Elementary and Secondary Education (DESE), 2022) scale scores were collected in September 2021 after MCAS scores were publicly released. This two-round data-collection process ensured the study sample could be determined prior to collection of the 2021 MCAS data. In addition to district-provided data, publicly available 2020–2021 school data from the Massachusetts DESE were used.

Fall *i-Ready Diagnostic* scale scores served as the pre-achievement measure for this study. The *i-Ready Diagnostic* is developed and owned by Curriculum Associates and is an adaptive, online assessment designed to place students in relation to grade-level standards and national norms for either Reading or Mathematics. *i-Ready Diagnostic* scale scores range from 100–800. All students completed the fall *i-Ready Diagnostic* between August 1, 2020 and November 15, 2020, Curriculum Associates' standard fall testing window. MCAS assessment scale scores served as the achievement outcome measure for this study. The MCAS assessment is the statewide summative, standardized assessment completed by all Massachusetts students in the spring of each school year. MCAS scores range from 440–560. All achievement data were cleaned to ensure they were within valid score ranges. Note that MCAS assessments were not completed in spring 2020 due to COVID-19-related school closures, which is why fall *i-Ready Diagnostic* scale scores served as the pre-achievement measure for this study. A moderate relationship was observed between fall *i-Ready Diagnostic* and MCAS scale scores, with correlations of .70 and .72 for Reading and Mathematics, respectively, suggesting fall *i-Ready Diagnostic* scores served as a reasonable pre-achievement proxy for this study.

SAS® is a registered trademark of the SAS Institute, Inc.

Sample

Students were assigned to one of two groups in alignment with an intent-to-treat quasiexperimental design. Within an intent-to-treat framework, any students exposed to an educational intervention are analyzed as if they received the intervention, regardless of the extent to which they actually received the intervention. Estimated intervention effects from intent-to-treat studies represent the effect of being assigned an intervention in practice, with ranging implementations and dosage (Shadish et al., 2002). As such, students were assigned to the *i-Ready Personalized Instruction* group—hereafter referred to as instruction group—if they completed at least one *i-Ready Personalized Instruction* lesson for a given subject between August 1, 2020 and November 15, 2020. Students were assigned to the Diagnostic-only comparison group if they did not complete a lesson in the same time window. This window was chosen because it aligns with Curriculum Associates' established fall *i-Ready Diagnostic* testing window. Approximately two-thirds of students were assigned to the instruction group for each subject.

Note that while schools were initially selected for the study based on whether *i-Ready Personalized Instruction* was used or not, not all students within schools using *i-Ready Personalized Instruction* used the program. This is to be expected as *i-Ready Personalized Instruction* is a supplemental program, and we know that implementations may vary between classrooms (e.g., some schools require usage of *i-Ready Personalized Instruction* and others allow teachers to decide whether to use the program) and across student groups (e.g., more students who are not ready for grade-level material tend to use *i-Ready Personalized Instruction* compared to students who are ready for grade-level material). While there were some schools within our sample where all students within the school used *i-Ready Personalized Instruction* or all students did not use the program, some schools were mixed implementations. For this reason, combined with the expectation that not all students within a school will use the supplemental program, inclusion in the instruction or Diagnostic-only groups was assigned by student and whether or not they had the opportunity to use *i-Ready Personalized Instruction*.

Table 1 displays an overview of the sample descriptive information for each group. As is evidenced based on the raw percentage differences between groups, the instruction group was composed of fewer White students, more Black students, more Latino students, and more students with disabilities compared to the Diagnostic-only group. Moreover, the instruction group was lower performing on the fall *i-Ready Diagnostic* than the Diagnostic-only group. These student characteristics are theoretically or empirically related to achievement. To strengthen claims that any differences in MCAS achievement between instruction and Diagnostic-only groups is related to *i-Ready Personalized Instruction* usage, coarsened exact matching was used to minimize sample differences on these key covariates also related to achievement.

Coarsened Exact Matching

Once pre-achievement and demographic data were received and cleaned, but before MCAS scores were available, Coarsened Exact Matching (CEM) was used to achieve a sample that was more balanced on the pre-achievement measure and other important covariates. Like propensity score matching and other matching methods, CEM is designed to achieve a treatment group and control group that are similar on important, observed covariates in order to reduce the confounding influence of those covariates on an outcome and/or treatment effect (Iacus et al., 2012). CEM was chosen as the matching method for this study because it has superior ability to reduce imbalance and model dependence when compared to propensity score matching and many other common matching methods (Iacus et al., 2012; Wells et al., 2013), could ensure a close match on the pre-achievement measure while still allowing us to match on additional covariates, and could retain a large proportion of students in the initial sample. With CEM, continuous variables are temporarily "coarsened" into categorical variables and then all possible combinations of covariate values are used to create bins or strata. For example, in a study with a three-level categorical variable and a continuous variable that is coarsened into five categories, there are 15 strata. Each student, regardless of treatment status, is then placed into the stratum that matches that student's pattern of covariates. After students are assigned to the appropriate stratum, strata that contain treatment students but not comparison, or vice versa, are discarded. Students in the remaining strata are then assigned weights according to the ratio of treatment to comparison students in the stratum. For the purposes of impact analysis, the coarsened variables are returned to their original continuous scale, and each student's weight—hereafter referred to as the CEM weight—is included in the estimation.

All matching was completed using version 4.3.0 of the MatchIt package (Ho et al., 2011) in version 4.1.1 of R (R Core Team, 2021). Two separate, matched datasets were created, one for ELA and one for Mathematics. Matching ELA and Mathematics separately allowed us to include students who had a fall *i-Ready Diagnostic* score in only one subject as well as ensuring the closest possible match on *i-Ready Diagnostic* score in that subject specifically. After removing students who were missing data for any of the covariates used for matching, students were matched on the following characteristics: fall *i-Ready Diagnostic* score, which was coarsened into 20 percentile-based bins; a binary variable that indicated whether the entire *i-Ready Diagnostic* had been taken in school, according to a question the student answered each time they logged in to the *i-Ready Diagnostic*; whether the student was Latino; the student's race, coarsened into three categories of Black, White, and Other Race due to sample size constraints; and whether or not the student was classified as a student with a disability. Table 1 displays information about the differences between the instruction and Diagnostic-only groups on important covariates before and after matching. In both cases, only differences before attrition are displayed.

After matching, instruction and Diagnostic-only groups were balanced on all matching covariates. All standardized differences were well below <|.25|, a common benchmark used to evaluate group balance in educational quasiexperimental designs (Evidence for ESSA Standards and Procedures, n.d.; What Works Clearinghouse, 2020).

Table 1. Sample Descriptive Data before and after Matching

ELA										
		Student N	**School N**	**Mean Fall Diagnostic Score**	**% In-School Diagnostics**	**% Latino**	**% Black**	**% White**	**% Other Race**	**% Students with Disabilities**
Before Matching	***i-Ready Personalized Instruction***	1,862	43	531.64	13.21%	56.34%	24.38%	68.37%	7.25%	23.63%
	Diagnostic Only	859	30	565.05	5.12%	11.41%	11.18%	78.11%	10.71%	19.21%

After Matching (CEM Weighted)	***i-Ready Personalized Instruction***	1,397	43	565.49	4.46%	10.86%	10.86%	79.61%	9.53%	17.61%
	Diagnostic Only	829	30	565.55	4.46%	10.86%	10.86%	79.61%	9.53%	17.61%
	Instruction – Diagnostic Only (|Standardized Difference|)			-.06 (< .01)	.00% (.00)	.00% (.00)	.00% (.00)	.00% (.00)	.00% (.00)	.00% (.00)
Mathematics										
		Student N	**School N**	**Mean Fall Diagnostic Score**	**% In-School Diagnostics**	**% Latino**	**% Black**	**% White**	**% Other Race**	**% Students with Disabilities**
Before Matching	***i-Ready Personalized Instruction***	1,883	44	451.75	13.06%	58.84%	25.07%	67.87%	7.06%	23.95%
	Diagnostic Only	946	22	468.52	4.76%	9.09%	10.25%	79.18%	10.57%	18.50%
After Matching (CEM Weighted)	***i-Ready Personalized Instruction***	1,459	42	468.48	4.45%	8.57%	10.20%	79.39%	10.41%	17.14%
	Diagnostic Only	922	22	468.80	4.45%	8.57%	10.20%	79.39%	10.41%	17.14%
	Instruction – Diagnostic Only (|Standardized Difference|)			-.32 (< .01)	.00% (.00)	.00% (.00)	.00% (.00)	.00% (.00)	.00% (.00)	.00% (.00)

Note. The Other Race category combines the Massachusetts DESE's categories of Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, and all combinations of two or more races. Although combining these categories is not ideal or preferred, they had to be combined due to small sample sizes.

Note. The Students with Disabilities category was derived from an indicator referred to as "special education" provided by districts. "Students with Disabilities" is used here given it is preferred language to "special education," and the Massachusetts DESE (2021b) defines their special education variable in relation to the educational environment for a student with a disability.

Post-Attrition Sample Evaluation

Upon collection of 2021 MCAS scale scores, the matched sample was evaluated for attrition (i.e., students who had a valid fall *i-Ready Diagnostic* score and relevant demographic data but did not have a valid MCAS scale score for the respective subject). See Table 2 for attrition rates of the instruction and Diagnostic-only groups. For both ELA and Mathematics, differential attrition across instruction and Diagnostic-only groups was less than 4%, which is considered an acceptable rate of attrition in education research (Evidence for ESSA Standards and Procedures, n.d.). See Table 3 for descriptive information of the final sample after students without MCAS scale scores were removed. This sample represents the CEM-weighted sample and is representative of the sample as analyzed, not necessarily of individual students.

Table 2. Sample Attrition by Subject

	Instruction	Diagnostic Only	Differential Attrition
ELA	4.51%	3.74%	.77%
Mathematics	6.51%	3.25%	3.26%

Table 3. Final Sample Descriptive Information for *i-Ready Personalized Instruction* and Diagnostic-Only Groups

ELA									
	Student N	**School N**	**Mean Fall Diagnostic Score**	**% In-School Diagnostics**	**% Latino**	**% Black**	**% White**	**% Other Race**	**% Students with Disabilities**
i-Ready Personalized Instruction	1,334	43	565.70	4.42%	11.01%	11.12%	79.04%	9.84%	17.83%
Diagnostic Only	798	25	566.35	4.26%	10.27%	11.03%	79.45%	9.52%	17.17%
Mathematics									
	Student N	**School N**	**Mean Fall Diagnostic Score**	**% In-School Diagnostics**	**% Latino**	**% Black**	**% White**	**% Other Race**	**% Students with Disabilities**
i-Ready Personalized Instruction	1,364	41	468.81	4.53%	8.66%	10.13%	79.51%	10.35%	16.69%
Diagnostic Only	892	20	469.20	4.26%	8.07%	10.08%	79.48%	10.42%	17.15%

Note. The Other Race category combines the Massachusetts DESE's categories of Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, and all combinations of two or more races. Although combining these categories is not ideal or preferred, they had to be combined due to small sample sizes.
Note. The Students with Disabilities category was derived from an indicator referred to as "special education" provided by districts. "Students with Disabilities" is used here given it is preferred language to "special education," and the Massachusetts DESE (2021b) defines their special education variable in relation to the educational environment for a student with a disability.
Note. These numbers incorporate the CEM weights.

After removing students without MCAS scale scores, baseline equivalence was re-evaluated for each subject. As mentioned, the fall *i-Ready Diagnostic* (i.e., the pre-achievement measure) was the prioritized baseline equivalence variable, and thus was the only variable considered when re-estimating baseline equivalence after attrition. To align with the structure of the eventual impact model, a two-level, students-within-schools hierarchical linear model (HLM) was used to estimate the adjusted mean difference between groups on the fall *i-Ready Diagnostic*:

Level 1 (Student):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(Inst_{ij}) + e_{ij}$$

Level 2 (School):

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

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

where Y_{ij} is the estimated fall *i-Ready Diagnostic* score for student *i* in school *j*, β_{0j} is the adjusted average fall *i-Ready Diagnostic* score for Diagnostic-only students in school *j*, β_{1j} is the adjusted mean difference in fall *i-Ready Diagnostic* score between instruction and Diagnostic-only students, $Inst_{ij}$ is a dummy-coded grouping variable where instruction students have a value of 1 and Diagnostic-only students have a value of 0, e_{ij} is the within-school student variability not accounted for by the model, y_{00} is the predicted average fall *i-Ready Diagnostic* score across schools for Diagnostic-only students, and u_{0j} is the deviation between a school's average *i-Ready Diagnostic* score and the overall average *i-Ready Diagnostic* score across schools for Diagnostic-only students.

See Table 4 for the post-attrition adjusted mean difference between the instruction and Diagnostic-only groups as well as baseline equivalence effect sizes standardized to the standard deviation of the Diagnostic-only group (Evidence for ESSA Standards and Procedures, n.d.). At a standardized effect size well below $|.25|$, both subject samples were considered baseline equivalent on the fall *i-Ready Diagnostic* after accounting for attrition (Evidence for ESSA Standards and Procedures, n.d.; What Works Clearinghouse, 2020).

Table 4. Post-Attrition Baseline Equivalence on the Fall *i-Ready Diagnostic*

	Adjusted Mean Difference		Diagnostic-Only SD	Effect Size
	Estimate	Standard Error		
ELA	3.70	2.87	46.78	.08
Mathematics	.04	1.88	24.94	< .01

Note. These numbers incorporate the CEM weights. Mean difference is adjusted for school membership.

Impact Model

All impact model analyses were conducted in SAS 9.4 software using the Mixed Procedure (SAS Institute, 2017). All estimates incorporated the CEM weights. Note the CEM weights are not included as a covariate in the model representation below, as weights were incorporated via replacing fixed- and random-effects design matrices with CEM-weighted diagonal weight matrices (SAS Institute, 2017).

Intraclass Correlation

Prior to estimating the impact of *i-Ready Personalized Instruction* on MCAS achievement, the intraclass correlation coefficients (ICC) were estimated by subject to evaluate the extent of school-level clustering. When data are clustered within a unit, scores and their associated errors may no longer be independent of one another, resulting in a violation of the assumption of independence. Violating this assumption leads to deflated standard errors and increased chance of interpreting statistical significance as a true effect when the estimated effect is actually due to sampling error (i.e., not a true instructional effect of *i-Ready Personalized Instruction*; O'Connell & McCoach, 2008).

To estimate the ICC, subject-specific, empty random-intercepts-only models predicting MCAS achievement were fit to the data. The ICC represents the proportion of total variability in MCAS scores due to school clustering:

$$\frac{\tau_{00}}{\tau_{00} + \sigma^2}$$

where τ_{00} is the variance between schools, and σ^2 is the variance within school. For ELA, 14% of the variability in MCAS scores was explained by school membership. For Mathematics, 12% of the variability in MCAS scores was explained by school membership. See the Appendix for variance estimates. For both subjects, the effect of school was considered strong enough to account for school-based clustering in the data. Although there are several methods for correcting the violation of independence, in this study a two-level HLM was used. Using an HLM served two needs: 1) the need to account for clustering of students within schools and address the violation of independence and 2) the need to account for theoretically and/or empirically relevant school-level covariates (Raudenbush & Bryk, 2002).

Impact Model Specification

The impact model was determined by fitting a series of competing models to the data by subject. Competing models were compared via the Bayesian information criterion (BIC; Schwarz, 1978) due to its performance when selecting the most accurate model under the condition of ICCs in the .1 range (Whittaker & Furlow, 2009). The model with the lowest relative BIC was selected as the champion impact model. For both ELA and Mathematics, the following impact model was selected:

Level 1 (Student):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(Instruction_{ij}) + \beta_{2j}(Fall\ Score_{ij} - \overline{Fall\ Score}) + \beta_{3j}(Hispanic_{ij}) + \beta_{4j}(Disability_{ij}) + \beta_{5j}(White_{ij}) + \beta_{6j}(Fall\ Testing\ Location_{ij}) + e_{ij}$$

Level 2 (School):

$$\beta_{0j} = y_{00} + y_{01}(Percent\ White_j) + y_{02}(Student\text{–}Teacher\ Ratio_j) + u_{0j}$$

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

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

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

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

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

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

where Y_{ij} is the estimated 2021 MCAS scale score for student *i* in school *j*; β_{0j} is the average MCAS scale score for students in school *j* for the reference groups; β_{1j} is the adjusted *i-Ready Personalized Instruction* treatment effect; β_{2j} is the difference in MCAS scale score associated with achievement on the fall *i-Ready Diagnostic* for students in school *j* (grand-mean centered)*;* β_{3j} is the adjusted difference in MCAS scale score associated with being a Latino student in school *j;* β_{4j} is the adjusted difference in MCAS scale score associated with being a student with a disability in school *j;* β_{5j} is the adjusted difference in MCAS scale score associated with being of student of color racial status in school *j;* and β_{6j} is the adjusted difference in MCAS scale score associated with fall *i-Ready Diagnostic* testing location. For the definitions and reference group of each variable, see Table 5. Note that given demographic covariates were not grand-mean centered, they retain interpretations in relation to the reference group, and β_{0j} retains an interpretation in relation to the average fall *i-Ready Diagnostic* score for reference groups of each student-level demographic covariate.

Table 5. Impact Model Variable Definitions and Reference Groups

Student Covariates		
Variable	**Definition**	**Reference**
Instruction	Instruction or Diagnostic-only group membership	Diagnostic-only student
Fall Score	Fall *i-Ready Diagnostic* scale score	Continuous variable
Latino	Latino-identifying student	Non-Latino student
Disability	Disability status	Student without a disability
Student of Color	Race indicator, representing White students and students of color	White student
Fall Testing Location	Testing location of student's fall *i-Ready Diagnostic*	In school
School Covariates		
Variable	**Definition**	**Reference**
Percent White	Percentage of White students within a given school	Continuous variable
Student–Teacher Ratio	Ratio of students to teachers within a given school	18 students per teacher

Note. The Students of Color Race category combines the Massachusetts DESE's categories of Asian, American Indian or Alaska Native, Black, Native Hawaiian or Other Pacific Islander, and all combinations of two or more races. Although combining these categories is not ideal or preferred, they had to be combined due to small sample sizes.

Note. The Students with Disabilities category was derived from an indicator referred to as "special education" provided by districts. "Students with Disabilities" is used here given it is preferred language to "special education," and the Massachusetts DESE (2021b) defines their special education variable in relation to the educational environment for a student with a disability.

Assumptions for each subject-specific model were evaluated using the SAS macro mixed_DX (Bell et al., 2010). No major assumption violations were found. Effects for the selected impact model were estimated using restricted maximum likelihood estimation (REML). Although the number of school clusters for this study was not unduly small (see Table 3), REML was conservatively selected as the estimator given its performance when there are few school-level clusters and clusters are small (McNeish & Stapleton, 2016).

An effect size similar in calculation to Cohen's *d* was calculated as an index of the standardized effect of *i-Ready Personalized Instruction* for each subject. However, rather than standardizing to the pooled standard deviation (SD) of the outcome across groups, as is typical in a Cohen's *d* calculation, the index was standardized to the SD of the Diagnostic-only group (Evidence for ESSA Standards and Procedures, n.d.). The SD represents the raw SD and is not adjusted for covariates:

$$\frac{mean_{instruction} - mean_{Diagnostic\ Only}}{SD_{Diagnostic\ Only}}$$

Practical Significance of the Model

Practical significance of the impact model was evaluated using proportion reduction in variance statistics (Raudenbush & Bryk, 2002) that represent the proportion of variance reduced at the student or school level by the set of student and school covariates. The proportion of variance reduced was calculated using variance components from two models. Model 1 was an empty, random-intercepts-only model. Model 2 was the full impact model with student and school covariates as described above. To calculate the proportion of within-school student variability reduced, the following equation was used:

$$\frac{\sigma^2(model\ 1) - \sigma^2(model\ 2)}{\sigma^2(model\ 1)}$$

where σ^2 represents the variance attributed to students within schools. To calculate the proportion of between-school variability reduced, the following equation was used:

$$\frac{\tau_{00}(model\ 1) - \tau_{00}(model\ 2)}{\tau_{00}(model\ 1)}$$

where τ_{00} represents the variance attributed to schools. See the Appendix for the estimated student- and school-level variance components. It is important to note that because the proportion of variance reduced at either student or school levels can be affected by the addition of covariates at the other level (Raudenbush & Bryk, 2002), the proportion of variance reduced at either student or school level does not represent the unique effect of the student covariates in explaining within-school student variability, nor the unique effect of school covariates in explaining between-school variability. Rather, the proportion of variance reduced at either level represents the variance reduced by the *set* of student and school covariates. Although the unique effect of covariates at the student or school level could be calculated, an indication of how the overall model performed is more relevant given the proposed research questions. As such, any calculated proportion reduction in variance statistics represent the proportion of variability reduced at the student and school levels from the entire set of student and school covariates.

Results

What Is the Impact of *i-Ready Personalized Instruction* on MCAS Achievement for ELA and for Mathematics?

This study was designed to evaluate the impact of *i-Ready Personalized Instruction* on MCAS achievement for Grade 5 students. Two-level HLMs accounting for the student- and school-level covariates were used to estimate the impact of *i-Ready Personalized Instruction* on MCAS achievement. Both ELA and Mathematics models reduced a practically significant amount of variability at the student and school levels. See Table 6 for the proportion of variance reduced at the student and school levels by student and school covariates. Recall from ICC calculations above that 14% and 12% of the variability in MCAS scores for ELA and Mathematics, respectively, was due to between-school clustering. The proportion of variance reduced at the school level likely appears so large given the majority of variability in MCAS scores was within schools, not between. For the full table of variance estimates, see the Appendix.

Table 6. Proportion of Variance Explained at Each Level by the Set of Student and School Covariates

	Student	School
ELA	47.22%	72.54%
Mathematics	59.20%	77.85%

For both ELA and Mathematics, there was a statistically significant positive effect of *i-Ready Personalized Instruction* usage (see Table 7). On average, students using *i-Ready Personalized Instruction* for Reading performed 2.50 points higher ($p < .05$) on the ELA portion of the MCAS compared to a group of similar Diagnostic-only students after accounting for student and school characteristics. On average, students using *i-Ready Personalized Instruction* for Mathematics performed 4.28 points higher ($p < .001$) on the Mathematics portion of the MCAS compared to a group of similar Diagnostic-only students after accounting for student and school characteristics. For the full table of parameter estimates, see the Appendix.

Table 7. *i-Ready Personalized Instruction* Treatment Effect Information

	Treatment Effect			Diagnostic-Only SD	Standardized Effect Size
	Unstandardized Coefficient	Standard Error	Confidence Interval		
ELA	2.50*	.98	.57–4.43	20.13	.12
Mathematics	4.28***	1.09	2.14–6.41	18.68	.23

Note. *$p <= .05$, **$p <= .01$, ***$p <= .001$

Note. Confidence intervals are calculated using $\alpha = .05$.

Note. Standardized effect size is standardized to the SD of the Diagnostic-only group.

Exploratory Analyses

This section of the report includes methodology and results related to the exploratory research questions. The exploratory analyses were designed to answer the following questions for Grade 5 students:

1. To what extent were instruction students using *i-Ready Personalized Instruction* for Reading?
2. To what extent were instruction students using *i-Ready Personalized Instruction* for Mathematics?
3. What was the impact of *i-Ready Personalized Instruction* for Reading on MCAS ELA achievement for students who used *i-Ready Personalized Instruction* according to established usage recommendations?
4. What was the impact of *i-Ready Personalized Instruction* for Mathematics on MCAS Mathematics achievement for students who used *i-Ready Personalized Instruction* according to established usage recommendations?

Methodology

Unless otherwise noted, all analyses were completed using SAS 9.4M7 software with Enterprise Guide version 7.1 (SAS Institute, 2017).

Sample

The same pool of data was used for these analyses as above for the primary analyses. For these analyses, students were assigned to one of three groups in alignment with a treatment-on-the-treated quasiexperimental design. Within a treatment-on-the-treated framework, students are selected for analysis based on their exposure to a theoretically established "dosage." Estimated intervention effects from treatment-on-the-treated studies represent the effect of receiving an intervention as intended. For a given subject, Curriculum Associates recommends students consistently spend at least 30 minutes per week using *i-Ready Personalized Instruction* throughout the year and strive to maintain a lesson pass rate of at least 70%. These recommendations are provided to educators, and educators are encouraged to help their students meet these benchmarks. Because districts have varying testing schedules, school calendars, and seasonal breaks, consistent usage can be difficult to quantify. For most school districts, 18 weeks of *i-Ready Personalized Instruction* is feasible after accounting for seasonal breaks and Diagnostic administration. As such, a minimum of 18 weeks of *i-Ready Personalized Instruction* usage was used as a proxy for consistent usage.

Based on the above criteria, students were assigned to one of three groups. Students were assigned to the instruction guidance group if they used instruction and did meet usage recommendations described above, the instruction non-guidance group if they used instruction and did not meet usage recommendations described above, or the Diagnostic-only group if they did not complete an *i-Ready Personalized Instruction* lesson. Note the overall sample did not change from the primary analyses. The only difference between the primary analyses sample and this exploratory sample is that the instruction group was further split into the instruction guidance group and the instruction non-guidance group. See Table 8 for descriptive information of the sample.

Table 8. Final Sample Descriptive Information for *i-Ready Personalized Instruction* Users Who Met Usage Recommendations, *i-Ready Personalized Instruction* Users Who Did Not Meet Usage Recommendations, and Diagnostic-Only Groups

ELA									
	Student N	**School N**	**Mean Fall Diagnostic Score**	**% In-School Diagnostics**	**% Latino**	**% Black**	**% White**	**% Other Race**	**% Students with Disabilities**
***i-Ready Personalized Instruction*: Guidance**	427	33	551.66	6.09%	51.99%	24.82%	67.92%	7.26%	15.46%
***i-Ready Personalized Instruction*: Non-guidance**	907	43	532.38	9.04%	52.04%	21.94%	72.22%	5.84%	18.63%
Diagnostic Only	798	25	566.35	4.26%	10.27%	11.03%	79.45%	9.52%	17.17%
Mathematics									
	Student N	**School N**	**Mean Fall Diagnostic Score**	**% In-School Diagnostics**	**% Latino**	**% Black**	**% White**	**% Other Race**	**% Students with Disabilities**
***i-Ready Personalized Instruction*: Guidance**	665	33	461.08	9.92%	51.28%	20.45%	72.18%	7.37%	15.04%
***i-Ready Personalized Instruction*: Non-guidance**	699	41	450.31	3.43%	58.94%	22.46%	72.96%	4.58%	16.02%
Diagnostic Only	892	20	469.20	4.26%	8.07%	10.08%	79.48%	10.42%	17.15%

Note. The Other Race category combines the Massachusetts DESE's categories of Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, and all combinations of two or more races. Although combining these categories is not ideal or preferred, they had to be combined due to small sample sizes.
Note. The Students with Disabilities category was derived from an indicator referred to as "special education" provided by districts. "Students with Disabilities" is used here given it is preferred language to "special education," and the Massachusetts DESE (2021b) defines their special education variable in relation to the educational environment for a student with a disability.
Note. These values represent an un-weighted sample.

The primary comparison of these exploratory analyses is Diagnostic-only students and instruction guidance students. Because these results are exploratory and a follow-up to the primary analyses, re-matching for baseline equivalence between the Diagnostic-only and instruction guidance students was not completed. Moreover, because the CEM weights were computed based on the original comparison of Diagnostic-only and instruction groups, they were removed from all analyses in this section. All results represent un-weighted values.

To inform interpretation of the eventual impact results, baseline equivalence between the Diagnostic-only and instruction guidance groups on the fall *i-Ready Diagnostic* was evaluated according to the same baseline equivalence model as the primary analyses:

Level 1 (Student):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(Instruction_{ij}) + e_{ij}$$

Level 2 (School):

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

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

where terms have the same meaning as above, with the exception that β_{1j} is the adjusted mean difference in fall *i-Ready Diagnostic* score between instruction guidance and Diagnostic-only students, and $Instruction_{ij}$ is a dummy-coded variable where the instruction guidance students have a value of 1 and Diagnostic-only students have a value of 0.

Although the groups were not re-matched, the Diagnostic-only and instruction guidance groups were baseline equivalent on the fall *i-Ready Diagnostic* score (see Table 9; Evidence for ESSA Standards and Procedures, n.d.; What Works Clearinghouse, 2020).

Table 9. Baseline Equivalence on the Fall *i-Ready Diagnostic* for Diagnostic-Only Students and Instruction Guidance Students

	N		Adjusted Mean Difference		Diagnostic-Only SD	Effect Size
	Diagnostic-Only	Instruction Guidance	Estimate	Standard Error		
ELA	798	427	1.11	4.41	46.78	.02
Mathematics	892	665	-5.69	2.25	24.94	-.23

Note. These values represent an un-weighted sample. Mean difference is adjusted for school membership.

Impact Model

For consistency, the same impact model described above was used to evaluate the impact of *i-Ready Personalized Instruction* on MCAS achievement for instruction guidance students:

Level 1 (Student):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(Instruction_{ij}) + \beta_{2j}(Fall\,Score_{ij} - \overline{Fall\,Score}) + \beta_{3j}(Hispanic_{ij}) + \beta_{4j}(Disability_{ij}) + \beta_{5j}(Student\,of\,Color_{ij}) + \beta_{6j}(Fall\,Testing\,Location_{ij}) + e_{ij}$$

Level 2 (School):

$$\beta_{0j} = y_{00} + y_{01}(Percent\ White_j) + y_{02}(Student\text{–}Teacher\ Ratio_j) + u_{0j}$$

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

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

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

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

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

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

where all terms have the same meaning as the impact model above, with the exception that β_{1j} is now a vector representing the adjusted *i-Ready Personalized Instruction* treatment effect associated with using instruction according to guidance or not according to guidance. For the updated definitions and reference group of each variable, see Table 10. As above, given demographic covariates were not grand-mean centered, they retain interpretations in relation to the reference group, and β_{0j} retains an interpretation in relation to the average fall *i-Ready Diagnostic* score for reference groups of each student-level demographic covariate.

Table 10. Impact Model Variable Definitions and Reference Groups

Student Covariates		
Variable	**Definition**	**Reference**
Instruction	Instruction guidance, instruction non-guidance, or Diagnostic-only group membership	Diagnostic-only student
Fall Score	Fall *i-Ready Diagnostic* scale score	Continuous variable
Latino	Latino-identifying student	Non-Latino student
Disability	Disability status	Student without a disability
Student of Color	Race indicator representing White students and students of color	White student
Fall Testing Location	Testing location of student's fall *i-Ready Diagnostic*	In school
School Covariates		
Variable	**Definition**	**Reference**
Percent White	Percentage of White students within a given school	Continuous variable
Student–Teacher Ratio	Ratio of students to teachers within a given school	18 students per teacher

Note. The Students of Color Race category combines the Massachusetts DESE's categories of Asian, American Indian or Alaska Native, Black, Native Hawaiian or Other Pacific Islander, and all combinations of two or more races. Although combining these categories is not ideal or preferred, they had to be combined due to small sample sizes.

Note. The Students with Disabilities category was derived from an indicator referred to as "special education" provided by districts. "Students with Disabilities" is used here given it is preferred language to "special education," and the Massachusetts DESE (2021b) defines their special education variable in relation to the educational environment for a student with a disability.

An effect size similar in calculation to Cohen's *d* was calculated as an index of the standardized effect of *i-Ready Personalized Instruction* for each subject. However, rather than standardizing to the pooled SD of the outcome across groups, as is typical in a Cohen's *d* calculation, the index was standardized to the SD of the Diagnostic-only group (Evidence for ESSA Standards and Procedures, n.d.). The SD represents the raw SD and is not adjusted for covariates:

$$\frac{mean_{instruction} - mean_{Diagnostic\ Only}}{SD_{Diagnostic\ Only}}$$

Results

To What Extent Were Instruction Students Using *i-Ready Personalized Instruction*?

To contextualize the impact of *i-Ready Personalized Instruction* findings, metrics related to instruction usage and the extent to which students met usage recommendations were evaluated. As described above, the established usage recommendations for a given subject encourage students to consistently meet an average of 30 minutes per week and maintain an average pass rate of 70%. Consistent usage was operationalized as a minimum of 18 weeks of instruction.

As Table 11 demonstrates, the average and median values for most of the usage metrics were just above the minimum recommended usage guidance. Usage of *i-Ready Personalized Instruction* for Mathematics was higher, more consistent, and with a higher average pass rate compared to usage of *i-Ready Personalized Instruction* for Reading. Although there is no recommendation regarding the number of lessons students should complete in a school year, we also examined the average number of lessons completed as it is useful for contextualizing the progress students made through the content in each subject. The average number of lessons completed was similar for Reading and Mathematics.

Table 11. *i-Ready Personalized Instruction* Usage Metrics

		Lessons Completed		Lesson Pass Rate		Weeks of Usage		Minutes per Week	
	Student N	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Reading	1,334	30.95	27.00	69.70%	73.50%	21.24	24.00	36.66	35.81
Mathematics	1,364	31.09	27.00	78.29%	80.00%	24.05	26.50	40.54	38.16

Note. Values represent the instruction group only.

Table 12 displays the percentage of students who met each component of recommended usage guidance. The pattern of greater Mathematics usage compared to Reading usage is demonstrated for each individual component of recommended usage guidance as well as the overall percentage of students meeting usage recommendations. Regardless of subject, the percentage of students who met individual components of the usage recommendations is relatively high, but the percentage of students who met all three components of usage recommendations is relatively low, with only 32% of students meeting all recommendations in Reading and 49% of students meeting all three recommendations in Mathematics.

Table 12. Percentage of Students Who Met Recommended Usage Guidance

		Percentage of Students				
	Student N	>= 70% Lessons Passed	>= 18 Weeks of Instruction Usage	>= 30 Minutes per Week Average	Met Minutes per Week Usage and Weekly Usage	Met All Usage Guidance
Reading	1,334	58.92%	68.52%	66.42%	55.32%	32.01%
Mathematics	1,364	75.37%	77.20%	73.61%	63.86%	48.75%

Note. Values represent the instruction group only.

What Is the Impact of Using *i-Ready Personalized Instruction* according to Usage Recommendations on MCAS Achievement for ELA and for Mathematics?

For both ELA and Mathematics, there was a statistically significant positive effect of using *i-Ready Personalized Instruction* according to usage recommendations (see Table 13). On average, students who met usage recommendations for *i-Ready Personalized Instruction* for Reading performed 7.96 points higher ($p < .001$) on the ELA portion of the MCAS compared to similar Diagnostic-only students after accounting for student and school characteristics. On average, students who met usage recommendations for *i-Ready Personalized Instruction* for Mathematics performed 9.50 points higher ($p < .001$) on the Mathematics portion of the MCAS compared to similar Diagnostic-only students after accounting for student and school characteristics. For the full table of parameter estimates, see the Appendix.

Table 13. *i-Ready Personalized Instruction* Treatment Effect Information for Students Using *i-Ready Personalized Instruction* according to Usage Recommendations

	Treatment Effect			Diagnostic-Only SD	Standardized Effect Size
	Unstandardized Coefficient	Standard Error	Confidence Interval		
ELA	7.96***	1.58	4.85–11.07	20.13	.40
Mathematics	9.50***	1.59	6.36–12.65	18.68	.51

Note. *p <= .05, **p <= .01, ***p <= .001

Note. Confidence intervals are calculated using $\alpha = .05$.

Note. Standardized effect size is standardized to the SD of the Diagnostic-only group.

These results are sizeable and impressive. However, it is worth reiterating that these analyses were exploratory. Although instruction guidance and Diagnostic-only groups for this analysis were baseline equivalent on the fall *i-Ready Diagnostic* score, this sample was obtained by matching for the primary analyses. These exploratory results should be interpreted with this information in mind, and results should be replicated to provide additional evidence for the effect of *i-Ready Personalized Instruction* on MCAS achievement when instruction is used according to usage recommendations.

Conclusion

Since the onset of the COVID-19 pandemic, students, families, and educators have experienced new ways of teaching and learning. Although schools are more adjusted to protocols and creative solutions to manage the pandemic within their classrooms, teaching and learning disruptions are still prevalent. Now more than ever, students and educators need evidence-supported programs that facilitate student learning and development. This study was a rigorous, quasiexperimental study meeting ESSA Level 2 evidence standards. Results for two phases of the study were presented: 1) results for an intent-to-treat analysis where all students using *i-Ready Personalized Instruction* were included as instruction users, regardless of the extent to which they used the program and 2) results for follow-up exploratory analyses that evaluated the extent to which instruction students used *i-Ready Personalized Instruction* and the relationship between MCAS achievement and using *i-Ready Personalized Instruction* according to usage recommendations. Both sets of findings support that *i-Ready Personalized Instruction* is a helpful tool for students, especially as we continue to adjust to learning transitions and disruptions related to COVID-19.

Moreover, this study demonstrates the relationship between *i-Ready Personalized Instruction* usage and achievement on a statewide summative assessment. Summative assessment results have far-reaching implications for policy and practice, and understanding the relationship between instructional programs and summative assessments is imperative for educators. These findings establish not only a positive influence of *i-Ready Personalized Instruction* on MCAS achievement, but also suggest a magnitude of effects on par or larger than other educational programs on standardized assessment outcomes (Lipsey et al., 2012). In a sample with a wide range of *i-Ready Personalized Instruction* usage, the average achievement gains of 2.5 and 4 points for ELA and Mathematics, respectively, speak to the benefits of supplemental, personalized instruction.

Per the intent-to-treat design, inclusion in the *i-Ready Personalized Instruction* group did not require adherence to a recommended "dose." That is, many students in the instruction group had less than 30 minutes per subject per week with the program, and they may not have used the program consistently week over week or maintained average pass rate recommendations. For this reason, we explored the impact of *i-Ready Personalized Instruction* when used according to the recommended "dose" of consistent 30 minutes per subject per week usage, coupled with maintaining a lesson pass rate of 70%. The average MCAS achievement gains of 8 and 9.5 points for ELA and Mathematics, respectively, speak to the achievement benefits of supplemental, personalized instruction when used with fidelity. Although such results are worth celebrating, we would be remiss not to acknowledge that most students in this study did not meet usage recommendations. Only 55% of Reading instruction users and 64% of Mathematics instruction users in this study used *i-Ready Personalized Instruction* consistently according to Curriculum Associates' established weekly time recommendations. Taking lesson pass rate guidance into account, only 32% of Reading instruction users and 49% of Mathematics instruction users met the established time and pass rate recommendations. It is worth noting that in previous studies, large numbers of students and schools were able to meet instruction usage recommendations (Curriculum Associates, 2021a; Swain, Randel, & Dvorak, 2020). It is possible that for students and educators in this study, the challenges of teaching and learning during a pandemic took precedence over supporting optimal *i-Ready Personalized Instruction* usage.

Considering both the primary and exploratory results together speaks to the benefits of evaluating the effectiveness of educational programs from both an intent-to-treat and treatment-on-treated perspective. There are clear achievement benefits for students who used *i-Ready Personalized Instruction* according to recommended usage. However, only a subset of students in our sample met usage recommendations, and this subset was particularly small in Reading. These results naturally beg the question of what results look like in general practice in a sample where *i-Ready Personalized Instruction* was used to varying extents. By approaching the question of the impact of *i-Ready Personalized Instruction* on MCAS achievement from both a general usage and intended optimal usage perspective, we can understand the achievement effects of *i-Ready Personalized Instruction* more holistically than from only one perspective. These results are particularly relevant for educators who are able to directly influence the implementation of *i-Ready Personalized Instruction* in their schools or classrooms. For the average student, general usage of *i-Ready Personalized Instruction* is beneficial, but using *i-Ready Personalized Instruction* according to Curriculum Associates' usage recommendations is all the more beneficial for students.

Future Directions

As has been well-documented, the COVID-19 pandemic has disproportionately affected some groups of students (Dawson, 2021; Lewis et al., 2021; Rome & Lay, 2022), and it is important to evaluate the impact of *i-Ready Personalized Instruction* for students of all backgrounds separately, rather than in the aggregate alone. Although pre-COVID-19 research supports the benefits of *i-Ready Personalized Instruction* for striving learners (Randel, Swain, Dvorak, Spratto, & Prendez, 2020a; Randel, Swain, Dvorak, Spratto, & Prendez, 2020b), Black students, Latino students, English Learners, and students with disabilities (Randel, Swain, Dvorak, & Prendez, 2020; Swain, Randel, Dvorak, & Prendez, 2020), evaluating the impact of *i-Ready Personalized Instruction* for student demographic groups was not supported in this study due to limited sample size and availability of some key demographic variables. This leaves a gap in the research base relating *i-Ready Personalized Instruction* to state summative assessments for individual student populations and presents an opportunity for future research related to *i-Ready Personalized Instruction*.

Although we explored the impact of using *i-Ready Personalized Instruction* according to usage recommendations, it is worth reiterating the study was not primarily designed to address this question. We must acknowledge that although instruction guidance and Diagnostic-only groups were baseline equivalent on fall achievement, this sample was initially obtained via matching for the primary analyses. If instruction guidance and Diagnostic-only students were matched for the purpose of these exploratory analyses, a different sample may have been matched. Additionally, these groups were baseline equivalent on fall achievement, but demographically they were quite different. Future work should replicate the effect of using *i-Ready Personalized Instruction* with fidelity in a different sample that is intentionally designed to compare instruction guidance and Diagnostic-only groups. Relatedly, it would benefit educators to understand ways to help students meet usage recommendations. Understanding the factors that facilitate meeting usage recommendations is a research opportunity that could have far-reaching benefits for students.

Limitations

Limitations for this study primarily align with a lack of information to contextualize classroom experience. Effects of *i-Ready Personalized Instruction* are certainly related to classroom context as implementations may vary across classrooms. Ideally, students would analytically be nested within classrooms within schools. However, classroom membership information was unavailable for this study. Without this information, we acknowledge key implementation and contextual factors may have been unaccounted for in the findings.

Moreover, this study used data from the 2020–2021 school year, the first year in which students returned to school from spring 2020 school closures due to COVID-19 and a year in which students, families, and school personnel adjusted to unique learning environments. COVID-19-related factors undoubtedly had an impact on education throughout this school year, and we do not have adequate data to account for their full impact. For example, we know there were varying learning models for students in this study, including remote learning, in-school learning, and hybrid learning. As mentioned, effects of *i-Ready Personalized Instruction* are likely related to the context of the classroom, and learning models may relate to—or interact with—the impact of *i-Ready Personalized instruction* on MCAS achievement. However, in this study, we did not have reliable data to account for these learning models or how they changed throughout the year.

Limitations related to testing location are worth noting for both the fall *i-Ready Diagnostic* and MCAS assessments. The vast majority of the sample in this study completed their fall *i-Ready Diagnostic* outside of school, and out-of-school *i-Ready Diagnostic* scores may be less reliable than those completed in school (Huff, 2020). Only about 4% of the *i-Ready Diagnostics* for either subject were completed in school. Although the inclusion of testing location in our matching and impact models may have mitigated some of the limitations associated with out-of-school *i-Ready Diagnostic* scores, we did not have a sufficient sample of in-school *i-Ready Diagnostics* to examine the results of in-school testers alone. Relatedly, some students completed their MCAS assessments remotely, and we were not able to attain information about students' testing location. Although no major differences at the state level were observed for most MCAS items between in-school and remote testers (Massachusetts Department of Elementary and Secondary Education, 2021a), this finding may or may not hold within the sample for this study.

References

Bell, B. A., Schoeneberger, J. A., Morgan, G. B., Kromrey, J. D., & Ferron, J. M. (2010). *Fundamental diagnostics for two-level mixed models: The SAS® macro MIXED_DX*. SAS Global Forum 2010 Proceedings. SAS Institute Inc. https://support.sas.com/resources/papers/proceedings10/201-2010.pdf

Curriculum Associates. (2021a). *Impact of* i-Ready Personalized Instruction *time and lesson pass rates on student learning gains*. Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-instruction-dosage-study-2021.pdf

Curriculum Associates. (2021b). *Understanding student learning: Insights from fall 2021*. Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-understanding-student-learning-paper-fall-results-2021.pdf

Curriculum Associates. (in press). *Impact of* i-Ready Personalized Instruction *time and lesson pass rates on student learning gains: Results from the 2018–2019 school year*. Curriculum Associates.

Dawson, M. (2021). *The impact of COVID-19 on student academic growth in 2020–2021*. (Report No. 19). Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-covid-growth-research-paper-2021.pdf

Evidence for ESSA Standards and Procedures. (n.d.). *Frequently asked questions*. Evidence for ESSA. https://www.evidenceforessa.org/page/frequently-asked-questions

Globe Newswire. (2019). *Online education market study 2019. World market projected to reach $350 billion by 2023, dominated by the United States and China*. Globe Newswire. https://www.globenewswire.com/news-release/2019/12/17/1961785/0/en/Online-Education-Market-Study-2019-World-Market-Projected-to-Reach-350-Billion-by-2025-Dominated-by-the-United-States-and-China.html

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

Huff, K. (2020). *National data quantifies impact of COVID learning loss; Raises questions about at-home testing*. PR Newswire. https://www.prnewswire.com/news-releases/new-data-from-curriculum-associates-quantifies-impact-of-covid-learning-loss-raises-questions-about-at-home-testing-301145294.html

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*(1), 1–24. https://doi.org/10.1093/pan/mpr013

Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020). *Learning during COVID-19: Initial findings on students' reading and math achievement and growth*. NWEA Collaborative for Student Growth. NWEA. https://www.nwea.org/content/uploads/2020/11/Collaborative-brief-Learning-during-COVID-19.NOV2020.pdf

Lewis, K., Kuhfeld, M., Ruzek, E., & McEachin, A. (2021). *Learning during COVID-19: Reading and math achievement in the 2020–21 school year*. NWEA Center for School and Student Progress. https://www.nwea.org/content/uploads/2021/07/Learning-during-COVID-19-Reading-and-math-achievement-in-the-2020-2021-school-year.research-brief-1.pdf

Li, C., & Lalani, F. (2020). *The COVID-19 pandemic has changed education forever. This is how*. World Economic Forum. https://www.weforum.org/agenda/2020/04/coronavirus-education-global-covid19-online-digital-learning

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. (2021a). *Frequently asked questions: Assessment and accountability: September 2021*. Massachusetts Department of Elementary and Secondary Education. https://www.doe.mass.edu/mcas/results.html

Massachusetts Department of Elementary and Secondary Education. (2021b). *SIMS data handbook version 20.6*. Massachusetts Department of Elementary and Secondary Education. https://www.doe.mass.edu/infoservices/data/sims

Massachusetts Department of Elementary and Secondary Education. (2022). *Massachusetts Comprehensive Assessment System.* Massachusetts Department of Elementary and Secondary Education. https://www.doe.mass.edu/mcas

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review, 28,* 295–314. https://doi.org/10.1007/s10648-014-9287-x

O'Connell, A. A., & McCoach, D. B. (2008). *Multilevel modeling of educational data.* Information Age Publishing.

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria.

Randel, B., Swain, M., Dvorak, R. N., & Prendez, J. Y. (2020). *Effectiveness of* i-Ready Instruction *in elementary reading for traditionally disadvantaged groups*. Human Resources Research Organization.

Randel, B., Swain, M., Dvorak, R. N., Spratto, E., & Prendez, J. Y. (2020a). *Impact evaluation of mathematics "*i-Ready" *for striving learners using 2018–19 data.* (Report No. 048). Human Resources Research Organization. https://eric.ed.gov/?q=%22striving+learners%22&id=ED610444

Randel, B., Swain, M., Dvorak, R. N., Spratto, E., & Prendez, J. Y. (2020b). *Impact evaluation of reading "*i-Ready" *for striving learners using 2018–19 data.* (Report No. 053). Human Resources Research Organization. https://eric.ed.gov/?q=%22striving+learners%22&id=ED610441

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Sage Publications.

Richards, E., Aspegren, E., & Mansfield, E. (2021). *A year into the pandemic, thousands of students still can't get reliable wifi for school. The digital divide remains worse than ever.* USA Today. https://www.usatoday.com/story/news/education/2021/02/04/ covid-online-school-broadband-internet-laptops/3930744001/

Rome, L., & Lay, L. (2022). *Revisiting the digital divide: Digital learning during the pandemic.* (Report No. 01). Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-digital-divide-report-2022.pdf

SAS Institute. (2017). *SAS/STAT 15.1 User's Guide.* SAS Institute Inc. https://go.documentation.sas.com/doc/en/pgmsascdc/9.4_3.4/statug/titlepage.htm

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464. http://doi.org/10.1214/aos/1176344136

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Houghton Mifflin Company.

Swain, M., Randel, B., & Dvorak, R. N. (2020). *Impact evaluation of reading* i-Ready Instruction *for elementary grades using 2018–19 data*: *Final report.* (Report No. 107). Human Resources Research Organization. https://files.eric.ed.gov/fulltext/ED604746.pdf

Swain, M., Randel, B., Dvorak, R. N., & Prendez, J. Y. (2020). *Effectiveness of* i-Ready Instruction *in elementary mathematics for traditionally disadvantaged groups.* Human Resources Research Organization.

The Pew Charitable Trusts. (2020). *What COVID-19 underscores about how broadband connectivity affects educational attainment.* The Pew Charitable Trusts. https://www.pewtrusts.org/en/research-and-analysis/articles/2020/12/08/what-covid-19-underscores-about-how-broadband-connectivity-affects-educational-attainment

Wells, A. R., Hamar, B., Bradley, C., Gandy, W. M., Harrison, P. L., Sidney, J. A., Coberley, C. R., Rula, E. Y., & Pope, J. E. (2013). Exploring robust methods for evaluating treatment and comparison groups in chronic care management programs. *Population Health Management, 16*(1), 35–45. https://doi.org/10.1089/pop.2011.0104

Whittaker, T. A., & Furlow, C. F. (2009). The comparison of model selection criteria when selecting among competing hierarchical linear models. *Journal of Modern Applied Statistical Methods, 8*(1). https://doi.org/10.22237/jmasm/1241136840

What Works Clearinghouse. (2020). *Standards handbook: Version 4.1*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Appendix

Table A1. Variance Estimates for the Empty and Impact HLMs

ELA			
Empty Model		Impact Model	
σ^2	τ_{00}	σ^2	τ_{00}
380.50	62.40	200.83	17.13
Mathematics			
Empty Model		**Impact Model**	
σ^2	τ_{00}	σ^2	τ_{00}
368.96	49.48	150.53	10.96

Table A2. Parameter Estimates for the Impact of *i-Ready Personalized Instruction*: ELA

Level	Variable	Category	Unstandardized Estimate	Standard Error	*p*	Confidence Interval
Student	Intercept		486.95	5.53	< .001	475.79–498.12
	Fall Score		.28	.01	< .001	.26–.30
	Instruction (reference = Diagnostic only)		2.50	.98	.01	.57–4.43
	Latino (reference = non-Latino)		-.35	1.08	.74	-2.46–1.76
	Disability (reference = no disability)		-4.59	.96	< .001	-6.47–-2.71
	Student of Color (reference = White)		.82	.85	.34	-.85–2.49
	Fall Testing Location (reference = in school)	Hybrid	-3.48	2.01	.08	-7.41–.46
		Out of School	-.97	1.83	.60	-4.56–2.63
		Unknown	-11.41	2.91	< .001	-17.12–-5.70
School	Percent White		.06	.03	.07	-.01–.12
	Student–Teacher Ratio (reference = 18)	3	5.93	9.90	.55	-13.68–25.54
		8	-10.32	11.78	.38	-33.53–12.89
		10	1.18	5.31	.83	-9.64–12.00
		11	3.57	5.34	.51	-7.30 –14.43
		12	7.66	5.24	.15	-3.02–18.34
		13	1.79	5.40	.74	-9.21–12.79
		14	.98	6.04	.87	-11.36–13.32
		15	2.70	6.77	.69	-11.21–16.61
		16	5.74	7.03	.42	-8.64–20.13
		17	4.06	7.02	.57	-10.31–18.43

Note. Confidence intervals are calculated using $\alpha = .05$.

Table A3. Parameter Estimates for the Impact of *i-Ready Personalized Instruction*: Mathematics

Level	Variable	Category	Unstandardized Estimate	Standard Error	*p*	Confidence Interval
Student	Intercept		473.28	4.52	< .001	464.14–482.42
	Fall Score		.57	.01	< .001	.55–.60
	Instruction (reference = Diagnostic only)		4.28	1.09	< .001	2.14–6.41
	Latino (reference = non-Latino)		-2.63	.96	.01	-4.51–-0.74
	Disability (reference = no disability)		-4.40	.80	< .001	-5.97–-2.84
	White (reference = White)		-.07	.70	.92	-1.45–1.30
	Fall Testing Location (reference = in school)	Hybrid	2.21	1.69	.19	-1.11–5.52
		Out of School	.27	1.46	.85	-2.60–3.14
		Unknown	2.97	1.68	.08	-.33–6.27
School	Percent White		.10	.03	< .001	.05–.15
	Student–Teacher Ratio (reference = 18)	3	4.61	7.58	.54	-10.44–19.67
		10	4.54	4.06	.27	-3.80–12.89
		11	9.90	4.15	.02	1.36–18.43
		12	7.63	4.00	.07	-.58–15.85
		13	8.16	4.03	.05	-.11–16.43
		14	5.30	4.65	.26	-4.23–14.83
		15	5.17	5.23	.33	-5.60–15.95
		16	4.30	5.39	.43	-6.73–15.34
		17	3.79	5.37	.49	-7.21–14.80

Note. Confidence intervals are calculated using $\alpha = .05$.

Table A4. Parameter Estimates for the Impact of Using *i-Ready Personalized Instruction* according to Usage Recommendations: ELA

Level	Variable	Category	Unstandardized Estimate	Standard Error	*p*	Confidence Interval
Student	Intercept		491.11	5.45	< .001	480.22–501.99
	Fall Score		.27	.01	< .001	.25–.28
	Instruction (reference = Diagnostic only)	Nonguidance	1.57	1.35	.25	-1.08–4.22
		Guidance	7.96	1.58	< .001	4.85–11.07
	Latino (reference = non-Latino)		.03	.97	.98	-1.89–1.92
	Disability (reference = no disability)		-3.37	.99	< .001	-5.31–-1.42
	White (reference = White)		-.32	.92	.73	-2.11–1.48
	Fall Testing Location (reference = in school)	Hybrid	-4.61	2.13	.03	-8.79–-.42
		Out of School	-4.28	1.85	.02	-7.91–-.65
		Unknown	-6.05	3.28	.07	-12.48–.37
School	Percent White		.08	.03	.02	.01–.14
	Student–Teacher Ratio (reference = 18)	3	3.42	8.77	.70	-13.94–20.77
		8	-12.27	12.29	.32	-36.43–11.89
		10	.11	5.14	.98	-10.23–10.45
		11	1.40	5.16	.79	-8.99–11.78
		12	2.79	5.09	.59	-7.44–13.03
		13	-1.22	5.26	.82	-11.78–9.35
		14	-4.39	5.83	.46	-16.14–7.37
		15	.42	6.42	.95	-12.60–13.45
		16	2.49	6.77	.72	-11.17–16.15
		17	1.44	6.73	.83	-12.16–15.04

Note. Confidence intervals are calculated using $\alpha = .05$.

Table A5. Parameter Estimates for the Impact of Using *i-Ready Personalized Instruction* according to Usage Recommendations: Mathematics

Level	Variable	Category	Unstandardized Estimate	Standard Error	*p*	Confidence Interval
Student	Intercept		474.50	4.34	< .001	465.79–483.21
	Fall Score		.48	.01	< .001	.46–.51
	Instruction: Nonguidance (reference = Diagnostic only)		3.63	1.53	.02	.60–6.66
	Instruction: Guidance		9.51	1.59	< .001	6.36–12.65
	Latino (reference = non-Latino)		-1.34	.83	.11	-2.98–.29
	Disability (reference = no disability)		-1.58	.86	.07	-3.26–.10
	White (reference = White)		-.09	.80	.91	-1.65–1.47
	Fall Testing Location (reference = in school)	Hybrid	1.06	1.77	.55	-2.41–4.54
		Out of School	1.35	1.54	.38	-1.68–4.37
		Unknown	2.50	2.14	.24	-1.70–6.70
School	Percent White		.14	.03	< .001	.09–.20
	Student–Teacher Ratio (reference = 18)	3	9.94	7.35	.18	-4.583–24.46
		10	.28	3.64	.94	-7.17–7.72
		11	2.39	3.73	.53	-5.23–10.01
		12	1.65	3.62	.65	-5.73–9.03
		13	1.83	3.67	.62	-5.64–9.30
		14	-2.20	4.16	.60	-10.68–6.28
		15	1.19	4.59	.80	-8.24–10.61
		16	-.67	4.80	.89	-10.46–9.11
		17	-.65	4.78	.89	-10.39–9.10

Note. Confidence intervals are calculated using α = .05.