Curriculum Associates RESEARCH

Impact of *i-Ready Personalized Instruction* on Massachusetts Comprehensive Assessment System Scores in Middle School

Reading and Mathematics

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

Research Report, November 2023

i-Ready

Curriculum Associates RESEARCH

Impact of *i-Ready Personalized Instruction* on Massachusetts Comprehensive Assessment System Scores in Middle School

Reading and Mathematics

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 6–8 students' scores on the Massachusetts Comprehensive Assessment System (MCAS) during the 2021–2022 school year. Students who used *i-Ready Personalized Instruction* for an average of 30–49 minutes per week for at least 18 weeks and who ended the year with at least a 70% lesson pass rate were matched to students who had not used *i-Ready Personalized Instruction* and were similar based on a pretest measure of achievement in the relevant subject and other important covariates. Students who used *i-Ready Personalized Instruction* scored, on average, about five points higher on the MCAS for English language arts (ELA) or MCAS for mathematics. This study was designed to meet the standards for ESSA Level 2 evidence.

i-Ready

Introduction

During the 2021–2022 school year, Curriculum Associates conducted research about the efficacy of *i-Ready Personalized Instruction* in middle school. *i-Ready Personalized Instruction* is a digital instructional supplement available for mathematics and reading[1] in Grades K–8. The *i-Ready* program offers its own assessment solutions, including the *i-Ready Diagnostic*, an adaptive assessment that provides information about each student's strengths and areas for growth in four Mathematics domains and up to six Reading domains. A multitude of research connects the use of *i-Ready Personalized Instruction* with better outcomes on the *i-Ready Diagnostic* (Curriculum Associates, 2020; Swain et al., 2020a; Swain et al., 2020b). However, this study examines the impact of the use of *i-Ready Personalized Instruction* according to guidance on an outcome that is top of mind for most educators: state test scores.

Study Purpose and Research Questions

This study was designed to examine the efficacy of *i-Ready Personalized Instruction* for improving scores on the MCAS. The following research questions were addressed:

1. What was the impact of using *i-Ready Personalized Instruction* for Reading according to Curriculum Associates' guidance on MCAS ELA scores?
2. What was the impact of using *i-Ready Personalized Instruction* for Mathematics according to Curriculum Associates' guidance on MCAS mathematics scores?

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.

[1]The *i-Ready Diagnostic* and *i-Ready Personalized Instruction* assess and teach reading skills, and they do not include assessments of writing. However, the MCAS assessment used as one of the outcome measures in this study is a full ELA assessment that includes writing and other skills. Throughout this paper, we use the term "reading" when talking about the *i-Ready* products alone, and "ELA" when talking about the MCAS assessment and the analyses in this subject.

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. In middle school grades, correlations of spring *i-Ready Diagnostic* for Reading scores with MCAS ELA scores range from .80 to .82 and correlations of spring *i-Ready Diagnostic* for Mathematics scores with MCAS mathematics scores range from .86 to .88 (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.

Student's MCAS ELA and mathematics scores from spring 2022 were the outcome measures 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 both subjects, 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, 2023a). In Grades 6–8, the MCAS for mathematics includes different domains by grade level, including ratios and proportional relationships, the number system, expressions and equations, functions, geometry, and statistics and probability (Massachusetts Department of Elementary and Secondary Education, 2023b).

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

The purpose of this study was to estimate the impact of usage of *i-Ready Personalized Instruction* when it is used according to Curriculum Associates' guidance. 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 in Curriculum Associates' research as a minimum cutoff (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 into the program during at least 18 calendar weeks, and used the program for an average of 30–49 minutes during those weeks. The comparison group was selected from the pool of students who had not completed any *i-Ready Personalized Instruction* lessons. Students who did not fall into either of these groups (e.g., a student who used the program for only two weeks) were not included. A student's use of *i-Ready Personalized Instruction* in one subject did not count toward their usage in the analysis of the other subject. A single student may have been in the treatment group for one subject and in the comparison group for the other subject.

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. Sample sizes in some groups were too low to permit individual analyses by grade level, so grade levels were combined within each subject for matching and analysis.

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

Grade Level	Student Group	Number of Students	Mean Fall *i-Ready Diagnostic* Score	Standard Deviation of Fall *i-Ready Diagnostic* Scores
ELA				
Grade 6	No *i-Ready Personalized Instruction* Usage	1,100	574.16	48.50
	i-Ready Personalized Instruction Usage According to Guidance	216	582.46	56.27
Grade 7	No *i-Ready Personalized Instruction* Usage	1,159	598.42	49.71
	i-Ready Personalized Instruction Usage According to Guidance	130	588.14	67.60
Grade 8	No *i-Ready Personalized Instruction* Usage	1,334	608.70	49.51
	i-Ready Personalized Instruction Usage According to Guidance	56	585.93	75.58
Mathematics				
Grade 6	No *i-Ready Personalized Instruction* Usage	1,168	477.93	27.02
	i-Ready Personalized Instruction Usage According to Guidance	240	466.06	33.63
Grade 7	No *i-Ready Personalized Instruction* Usage	1,093	489.35	29.48
	i-Ready Personalized Instruction Usage According to Guidance	143	482.20	41.60
Grade 8	No *i-Ready Personalized Instruction* Usage	1,225	499.12	35.30
	i-Ready Personalized Instruction Usage According to Guidance	155	487.66	37.47

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 subjects. 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. All models required an exact match on student grade level. 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 before the pretest measure. The final 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 the fall *i-Ready Diagnostic* score, which was centered at the mean score from the original dataset that included unmatched students. Additionally, the matching model used a .20 SD caliper, and the treatment student with the largest propensity score received the first match. Information about sample size and baseline equivalence on the fall *i-Ready Diagnostic* is presented in Table 3. Further information about sample demographics is presented in the Appendix—Table A1.

Table 3. Descriptive Statistics for Matched Samples

Subject	Total Students	Total Schools	*i-Ready Personalized Instruction* Group Mean Fall Diagnostic Score (SD)	Comparison Group Mean Fall Diagnostic Score (SD)	Glass's Delta for Fall Diagnostic Scores
ELA	758	25	592.61 (54.84)	590.83 (53.90)	.03
Mathematics	1,018	27	479.95 (35.95)	479.77 (31.79)	.01

Impact Model

To account for the clustered nature of the data, we used hierarchical linear models with students clustered within schools to estimate the impact of *i-Ready Personalized Instruction* on the MCAS (Raudenbush & Bryk, 2002). We first calculated the intraclass correlation coefficient (ICC) to confirm the magnitude of clustering by fitting an unconditional random-intercepts-only model for each subject. Each model predicted student-level MCAS 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. Variance at Each Level

Subject	Between Schools Variance	Within Schools Variance	ICC
ELA	101.18	481.07	.17
Mathematics	93.30	389.75	.19

After estimating the unconditional models, all covariates were entered as a block for the final impact models. Each of the two models took the following form:

Level 1 (Student):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(Usage\ of\ \text{i-Ready Personalized Instruction}_{ij}) + \beta_{2j}(Centered\ fall\ \text{i-Ready Diagnostic}\ score\ ) + \sum\beta_{3j}(Race/ethnicity_{ij}) + \beta_{4j}(Low\ income\ status_{ij}) + \beta_{5j}(Special\ Education\ status_{ij}) + \beta_{6j}(English\ learner\ status_{ij}) + \sum\beta_{7j}(Grade\ level_{ij}) + \sum\beta_{8j}(Grade\ level_{ij} \times centered\ fall\ \text{i-Ready Diagnostic}\ score_{ij}) + e_{ij}$$

Level 2 (School):

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

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

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

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

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

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

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

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{80}$$

where Y_{ij} represents the expected MCAS score for student *i* in school *j*; β_{0j} represents the school-level intercept, that is, the average MCAS score for a student in school *j* for whom all predictors are zero or the centered value; β_{1j} represents the difference in MCAS score associated with usage of *i-Ready Personalized Instruction*; β_{2j} represents the difference in MCAS score associated with a one-point increase in fall *i-Ready Diagnostic* score; β_{3j} is a vector of values that represent the difference in MCAS scores associated with the race/ethnicity reported for the student; β_{4j} through β_{6j} represent the differences in MCAS score associated with the binary variables low-income status, special education status, and English Learner status, respectively; β_{7j} is a vector of values that represent difference in MCAS score associated with the student's grade level, and β_{8j} is a vector of values that represent the interaction between student grade level and fall *i-Ready Diagnostic* score. It was necessary to include a term that captured the interaction between grade level and fall *i-Ready*

Diagnostic score because the *i-Ready Diagnostic* is vertically scaled (that is, the average score for each grade level is different), whereas MCAS scores are not vertically scaled. The interaction term accounts for these differences in scaling. Information about the values of these variables is available in Table 5.

Table 5. Values of Student-Level Variables

Variable	Values
Instruction User	Student did not use *i-Ready Personalized Instruction* (reference value), OR student used *i-Ready Personalized Instruction.*
Fall *i-Ready Diagnostic* Score	This value was centered at the mean for the full pre-matched sample for the relevant subject, that is, 594.40 for the ELA analysis and 487.46 for the mathematics analysis.
Race/Ethnicity	Black (reference value) OR Hispanic OR White OR Other race/ethnicity
Special Education Status	Student is not a special education student (reference value), OR student is a special education student.
Low-Income Status	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
English Learner Status	Student is not an English Learner (reference value), OR student is an English Learner.
Grade Level	Grade 6 (reference value) OR Grade 7 OR Grade 8

Note: The variable names and values used here and throughout the 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.

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, 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 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 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).

The second metric that was calculated is the improvement index, which is 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 percentile that could be expected by a median student from the comparison group if that student had received the intervention.

Lastly, we also calculated the number of additional comparison group students that 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

The outcome models indicated that usage of *i-Ready Personalized Instruction* is associated with statistically significantly higher MCAS scores after controlling for baseline achievement and other student characteristics. On average, students who used *i-Ready Personalized Instruction* for Reading according to Curriculum Associates' guidance scored 5.13 points higher ($p < .01$) on the MCAS ELA assessment than similar students who did not use *i-Ready Personalized Instruction* for Reading. This corresponds to a standardized effect size of .23 and an improvement index of 9.00. If each comparison group student had scored 5.13 points higher as a result of using *i-Ready Personalized Instruction*, an additional 7% of the comparison group would have been proficient. On average, students who used *i-Ready Personalized Instruction* for Mathematics according to Curriculum Associates' guidance scored 5.05 points higher on the MCAS mathematics assessment than similar students who did not use *i-Ready Personalized Instruction* for Mathematics. This corresponds to a standardized effect size of .25 and an improvement index of 9.86. If each comparison group student had scored 5.05 points higher as a result of using *i-Ready Personalized Instruction*, an additional 9% of the comparison group would have been proficient. The full parameter estimates from these models are presented in the Appendix—Tables A2 and A3.

Table 6. Impact of Usage of *i-Ready Personalized Instruction* According to Guidance

Subject	MCAS ELA Score Difference	*p*	Confidence Interval	Comparison Group SD	Glass's Delta	Improvement Index	Additional Proficient Comparison Group Students
ELA	5.13	< .01	2.08	8.22	.23	9.00	6.60%
Mathematics	5.05	< .01	2.90	7.20	.25	9.86	9.43%

Discussion

This study demonstrated evidence of the positive impact of *i-Ready Personalized Instruction* when used according to Curriculum Associates' guidance on a statewide comprehensive exam. The medium (Kraft, 2020) effect sizes of .23 (ELA) and .25 (mathematics) are particularly remarkable given the fact that the intervention is only an instructional supplement recommended to be used between 30 and 49 minutes per week. *i-Ready Personalized Instruction* uses such a small amount of instructional time that it may be used alongside other evidence-based products and practices to improve student learning even more.

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 the other methods of instruction in ELA and mathematics. 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 by illuminating what the comparison group was doing that was less helpful than *i-Ready Personalized Instruction*. 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* according to Curriculum Associates' guidance performed better on a statewide end-of-year exam compared to similar students who did not use *i-Ready Personalized Instruction*. 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, a school, or a district, the value of *i-Ready Personalized Instruction* is evidenced in the 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

Curriculum Associates. (2020). *Efficacy research for* i-Ready Personalized Instruction: *Reading and mathematics.* https://www.curriculumassociates.com/-/media/mainsite/files/i-Ready/iready-essa-l3-research-brief-2020.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

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

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, E. (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

Ho, D. E., 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. (2023a). *Massachusetts Comprehensive Assessment System: English language arts test design and development.* DESE. https://www.doe.mass.edu/mcas/tdd/ela.html

Massachusetts Department of Elementary and Secondary Education. (2023b). *Massachusetts Comprehensive Assessment System: Mathematics test design and development*. DESE. https://www.doe.mass.edu/mcas/tdd/math.html

R Core Team. (2022). *R (version 4.1.3).* The R Project for Statistical Computing. https://www.R-project.org

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 (2nd ed.).* Cengage Learning.

Swain, M., Randel, B., & Dvorak, R. N. (2020a). *Impact evaluation of mathematics* i-Ready *instruction for middle school grades using 2018-19 data.* https://files.eric.ed.gov/fulltext/ED604747.pdf

Swain, M., Randel, B., & Dvorak, R. N. (2020b). *Impact evaluation of reading* i-Ready *instruction for middle school grades using 2018-19 data.* https://files.eric.ed.gov/fulltext/ED604745.pdf

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

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. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Appendix

Table A1. Information about Student Samples after Matching

		Percentage:						
Group	**Total Students**	**Black**	**Hispanic**	**White**	**Other Race/ Ethnicity**	**Students with Disabilities**	**Low-Income Students**	**English Learners**
ELA								
i-Ready Personalized Instruction	379	10.03%	29.29%	50.13%	10.55%	13.19%	49.08%	5.80%
No *i-Ready Personalized Instruction* Usage	379	9.76%	28.76%	50.66%	10.82%	13.46%	47.23%	6.60%
Mathematics								
i-Ready Personalized Instruction	509	11.20%	35.56%	41.65%	11.59%	17.29%	56.19%	7.86%
No *i-Ready Personalized Instruction* Usage	509	9.82%	36.15%	44.01%	10.02%	13.95%	51.67%	7.27%

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. Fixed Effects Estimates for ELA Model

Subject	Estimate	Category	Unstandardized Estimate	Standard Error	*t*	*p*
ELA	(Intercept)		499.20	2.30	216.99	< 0.01
	i-Ready Personalized Instruction Usage		5.13	1.53	3.35	< 0.01
	Fall *i-Ready Diagnostic* Score		.36	.02	22.93	< 0.01
	Race/Ethnicity	Hispanic	2.41	1.96	1.23	.22
		White	1.71	1.99	.86	.39
		Other	3.15	2.28	1.38	.17
	Low-Income Status		-1.83	1.16	-1.59	.11
	Disability Status		-3.93	1.67	-2.35	.02
	English Learner Status		-1.76	2.30	-.77	.44
	Grade Level	Grade 7	-4.94	1.16	-4.24	< 0.01
		Grade 8	-10.06	1.62	-6.23	< 0.01
	Grade Level x Fall *i-Ready Diagnostic* Score Interaction	Grade 7 x Score	-.06	.02	-2.88	< 0.01
		Grade 8 x Score	-.06	.03	-2.16	.03

Table A3. Fixed Effects Estimates for Mathematics Model

Subject	Estimate	Category	Unstandardized Estimate	Standard Error	*t*	*p*
Mathematics	(Intercept)		498.58	1.66	300.65	< .01
	i-Ready Personalized Instruction Usage		5.05	1.09	4.64	< .01
	Fall *i-Ready Diagnostic* Score		.50	.02	24.97	< .01
	Race/Ethnicity	Hispanic	.29	1.37	.21	.83
		White	2.15	1.41	1.53	.13
		Other	3.71	1.63	2.27	.02
	Low-Income Status		-1.69	.87	-1.94	.05
	Disability Status		-3.09	1.15	-2.68	.01
	English Learner Status		-1.41	1.47	-.96	.34
	Grade Level	Grade 7	-8.39	.97	-8.65	< .01
		Grade 8	-11.88	1.00	-11.90	< .01
	Grade Level x Fall *i-Ready Diagnostic* Score Interaction	Grade 7 x Score	.04	.03	1.37	.17
		Grade 8 x score	-.06	.03	-2.16	.03