REVIEWED BY
EVIDENCE for ESSA
MEETS EVIDENCE STANDARDS

Curriculum Associates RESEARCH

Impact of *Magnetic Reading* in Iowa Schools: Evidence from the 2021–2022 School Year

Reading

Research Technical Report, November 2023

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

i-Ready Learning
Magnetic Reading™

REVIEWED BY EVIDENCE for ESSA MEETS EVIDENCE STANDARDS

Curriculum Associates RESEARCH

Impact of *Magnetic Reading* in Iowa Schools: Evidence from the 2021–2022 School Year

Reading

Research Technical Report, November 2023

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

Summary

This study evaluated the impact of *Magnetic Reading* on student reading achievement for students in Grades 3–5 in seven elementary schools in Iowa. The analysis was conducted using a sample of *Magnetic Reading* and comparison students who were from similar schools and who were baseline equivalent on their fall *i-Ready Diagnostic* for Reading. Findings suggest that *Magnetic Reading* usage is associated with an 11-point advantage on the Diagnostic and supports students' progress toward meeting their Stretch Growth® targets and grade-level proficiency. This study's design aligns with the Every Student Succeeds Act (ESSA)'s Moderate design principles, meets Evidence for ESSA's principles for "promising" evidence, and provides evidence that *Magnetic Reading* has a positive and significant impact on Grades 3–5 students' reading achievement and growth.

i-Ready Learning Magnetic Reading™

Introduction

In fall 2021, Curriculum Associates launched *Magnetic Reading*, a print-based, teacher-led reading program designed to support students' grade-level comprehension skills development in Grades 3–5. The program's foundation comprises texts that are written to be high-interest, culturally relevant, and informational. The program's purposes are to support students' comprehension skills, build knowledge that fosters deeper learning and connection with the texts, and nurture a love of reading. For educators, *Magnetic Reading* offers a digital library of resources for whole class grade-level instruction and differentiated learning, as well as scaffolds and protocols with which all students can relate and participate. Students are intended to use the program daily for 30–45 minutes as one component of their longer and more robust reading block.

To understand how *Magnetic Reading* contributes to students' reading skills and knowledge, we conducted a research study designed to answer the question: What is the effect of *Magnetic Reading* on students' reading achievement? Reading achievement was measured by students' *i-Ready Diagnostic* for Reading (i.e., Diagnostic) assessment scores. Although it is useful to understand how the benefits of *Magnetic Reading* will influence students' scale scores, we know that scale scores do not tell the whole story. We know that readiness for grade-level material is also important—perhaps more important—than a change in scale score points alone. As such, we also evaluated the impact of *Magnetic Reading* from the lens of how the program influences students' journeys toward grade-level knowledge and skills through the question: Do more students meet their Typical Growth and Stretch Growth targets when using *Magnetic Reading*?

Specifically, this study answered these questions for students in small, rural, Title I-eligible schools in Iowa. All schools were in their first full year of *Magnetic Reading* implementation, during which educators and students were learning and adjusting to the new program.

Methodology

Students from seven schools in Iowa are represented in this study. The sample was identified by first selecting eligible schools, followed by selecting eligible students within those schools. All data management and analysis needed for sample identification and outcomes analyses were conducted in *R* version 4.1.3 (R Core Team, 2021).

School Identification

Three schools in Iowa who were known to use *Magnetic Reading* as their supplemental comprehension reading program were identified as treatment schools. These schools implemented *Magnetic Reading* in Grades 3, 4, and 5 and reported using *Magnetic Reading* as their primary comprehension program for 30–45 minutes daily in their reading block for the full 2021–2022 academic year. According to the National Center for Education Statistics (NCES)'s Common Core of Data (CCD; US Department of Education, 2022), treatment schools were small, rural, Title I eligible, and composed of 90% or more White students.

Due to limited sample size and school-level information, a statistical matching process to identify comparison schools was not feasible. Instead, NCES CCD data were used to identify comparison elementary schools in Iowa who were also small, rural, Title I eligible, and composed of 90% or more White students but did not use *Magnetic Reading* or *Ready® Reading*.[1]

All treatment and possible comparison schools used *i-Ready Personalized Instruction*. Because *i-Ready* usage is known to have a positive effect on students' achievement (Curriculum Associates, 2021a; Randel, Swain, Dvorak, & Prendez, 2020; Randel, Swain, Dvorak, Spratto, & Prendez, 2020; Swain et al., 2019), it was important to take *i-Ready* usage at the treatment and comparison schools into account to obtain an accurate estimate of the impact of *Magnetic Reading* on student achievement. Accordingly, to ensure that comparison students had similar access to *i-Ready*, comparison schools: 1) were required to use *i-Ready* with at least 85% of their students in each Grade 3, 4, and 5 and 2) whose students used *i-Ready* for an average of at least 10 weeks in each Grade 3, 4, and 5. The final school sample consisted of three treatment schools and four comparison schools.

Student Identification

To be eligible for inclusion in the student sample, students in the identified schools were required to: 1) complete a fall and spring *i-Ready Diagnostic* for Reading in school and 2) have used *i-Ready Personalized Instruction* for Reading. Some students in the sample had extreme *i-Ready* usage compared to the rest of the sample. Because such high *i-Ready* usage may represent a different instructional track than peers within the classroom, students who completed 87 or more lessons and spent 1,533 minutes or more on *i-Ready* (i.e., students who were in the top 10% of *i-Ready* usage) were removed.

Using the MatchIt package in *R*, (Ho et al., 2011), propensity score matching was used to identify a sample of Grades 3–5 treatment and comparison students who were similar to one another within their grade on fall Diagnostic for Reading placement. By accounting for fall achievement, we identified a sample of treatment and comparison students who were similar in fall achievement, which is known to be highly correlated to end-of-year achievement. A less-biased estimate of the effect of *Magnetic Reading* was obtained via propensity score matching (Rosenbaum & Rubin, 1983).

An iterative process was used to procure the matched sample. First, a model employing nearest neighbor matching without a caliper was used. Next, separate models with the addition of various calipers were evaluated, with the goal of identifying the model in which covariate balance was maximized and loss of *Magnetic Reading* students was minimized (Harris & Horst, 2016). The sample was considered to have adequate balance if the standardized mean difference between *Magnetic Reading* and comparison groups after matching was < |.25| on all covariates (What Works

[1]*Ready Reading* is a Curriculum Associates teacher-led reading program. It is the legacy product prior to *Magnetic Reading* and, like *Magnetic Reading*, is intended to serve as the primary comprehension program in students' daily reading blocks. Although *Ready Reading* and *Magnetic Reading* are distinct products, some content, features, and strategies are common between the programs. For this reason, schools using *Ready Reading* were removed from the sample of comparison schools.

Clearinghouse, 2022). The final matching model used nearest neighbor matching without a caliper and predicted *Magnetic Reading* assignment from fall Diagnostic for Reading score, an exact match on grade, and an exact match on relative-to-grade-level placement. See Table 1 for sample balance before and after matching.

Table 1. Covariate Balance before and after Nearest Neighbor Matching

		Before Matching			After Matching		
		Magnetic Reading	**Comparison**	**|Standardized Mean Difference|**	***Magnetic Reading***	**Comparison**	**|Standardized Mean Difference|**
N		223	339	—	214	214	—
Fall Diagnostic Score		522.42	522.49	.00	523.60	524.63	.02
Fall Relative-to-Grade-Level Placement	On or Above Grade Level	.29	.27	.05	.30	.30	.00
	One Grade Level Below	.38	.33	.09	.36	.36	.00
	Two Grade Levels Below	.23	.24	.01	.23	.23	.00
	Three or More Grade Levels Below	.10	.16	.21	.10	.10	.00
Grade	Grade 3	.37	.23	.29	.35	.35	.00
	Grade 4	.27	.33	.13	.29	.29	.00
	Grade 5	.35	.44	.17	.37	.37	.00

Impact Model

The first research question sought to evaluate the effect of *Magnetic Reading* on students' reading achievement. The impact of *Magnetic Reading* on spring reading achievement was evaluated via a linear regression analysis on the matched Grades 3–5 sample. Using the stats package in *R* (R Core Team, 2021), a series of nested and non-nested models that included fall Diagnostic for Reading scores, grade, *Magnetic Reading* usage status, and various *i-Ready Personalized Instruction* metrics were compared to one another to determine the best combination of predictors to evaluate the impact of *Magnetic Reading* on spring Diagnostic for Reading scores. Although the final model was identified empirically, all tested covariates and covariate combinations were selected based on theoretical importance to the outcome. The final model predicted students' spring Diagnostic scores from their fall Diagnostic score, their grade, their average weekly time spent using *i-Ready,* number of weeks using *i-Ready* in the school year, interaction of their grade and fall Diagnostic score, and *Magnetic Reading* usage. To accommodate interpretation of the eventual model coefficients, fall Diagnostic score and average weekly time spent using *i-Ready,* and number of weeks using *i-Ready* in the school year were centered. Fall Diagnostic score was centered at 500

because 500 is a plausible score for each Grade 3, 4, or 5. Average weekly time spent using *i-Ready* was centered at 30 minutes, and number of weeks using *i-Ready* was centered at 18 weeks because together they represent the lower end of Curriculum Associates' suggested usage guidance of at least 30 minutes of *i-Ready* usage per week for at least 18 weeks in the school year. The final model explained 73.9% of variability in spring Diagnostic scores. Assumption checks were conducted, and no major violations were identified.

Given the data were clustered within schools, cluster-robust standard errors were employed. Although hierarchical linear modeling (HLM) is the more common technique used to account for clustered data in education, cluster-robust standard errors were employed as an alternative to HLM because the sample size within schools was small, there were few schools in the sample, and the school effect was not of substantive interest for our research questions (McNeish et al., 2017). Cluster-robust standard errors were obtained using the miceadds package in *R* (Robitzsch & Grund, 2022).

After estimating the regression coefficients, Glass's Delta was calculated as a standardized effect size for the impact of *Magnetic Reading*. To avoid inflation of the standardized effect due to any change in spring Diagnostic score variability as an outcome of using *Magnetic Reading,* Glass's Delta was standardized to the standard deviation of the raw spring Diagnostic score of the comparison group (Glass et al., 1981). To contextualize the range of the plausible effect of *Magnetic Reading* on spring achievement, 95% confidence intervals were calculated around the *Magnetic Reading* impact estimate using the cluster-robust standard errors.

Grade-Level Placement Changes

To contextualize the effect of *Magnetic Reading* on students' reading achievement, placement changes from fall to spring were evaluated using the matched sample. At fall and spring, students' placements were combined into three placements relative to their grade level: Two or More Grade Levels Below, One Grade Level Below, and Early On Grade Level or Above. Placements were defined by Curriculum Associates' "Standard View"[2] at both fall and spring. Given the propensity score matching model included an exact match on fall placement, the fall relative to-grade-level placement counts and percentages for the matched sample were the same for both the *Magnetic Reading* and comparison groups. In the final sample, 71 students (i.e., 33%) placed Two or More Grade Levels Below, 78 students (i.e., 36%) placed One Grade Level Below, and 65 students (i.e., 30%) of students placed On or Above Grade Level in each group in the fall.

[2]Per the "Standard View," students who placed Early On Grade Level, Mid On Grade Level, or above their chronological grade were considered to be Early On Grade Level or Above.

Growth Targets

The second research question sought to evaluate whether more students met their Typical Growth and Stretch Growth targets if they used *Magnetic Reading*. Curriculum Associates' Typical Growth and Stretch Growth are used to set goals for individual students or groups of students. Growth targets are determined based on students' grades and fall placements on the Diagnostic for a given subject. Typical Growth is the fall-to-spring growth expected for the average student in a given grade, grade-level placement, and subject on their Diagnostic. For many students, meeting Typical Growth will likely help them maintain their grade-level placement. This means that for students who are below grade level, meeting Typical Growth is insufficient for achieving grade-level proficiency (Curriculum Associates, 2023). Stretch Growth is an ambitious yet attainable goal representing the fall-to-spring growth a student should aim for to be on a path toward grade-level proficiency. Students' Typical Growth and Stretch Growth goals are provided to educators in *i-Ready* reports and are suggested focus areas when setting goals with individual students or groups of students.

Students were considered to have met their Typical Growth or Stretch Growth targets if their spring Diagnostic score was at or above their respective growth target. The percentage of students in the matched sample who met their Typical Growth or Stretch Growth targets was calculated for the *Magnetic Reading* and comparison groups. Using the gmodels package (Warnes et al., 2022) in *R*, a chi-square test of independence was used to evaluate whether there was an association between *Magnetic Reading* usage and meeting Typical Growth or Stretch Growth targets. Phi (φ) is provided as the effect size. Phi ranges from 0 to 1, with values closer to 1 representing a stronger relationship between *Magnetic Reading* usage and meeting the respective growth targets.

To further interpret the influence of *Magnetic Reading* on Typical Growth or Stretch Growth goal attainment, the relative benefit[3] associated with *Magnetic Reading* usage and meeting Typical Growth or Stretch Growth was calculated. The relative benefit represents how much more likely students are to meet their respective growth target if they use *Magnetic Reading*.[3]

Findings

What is the effect of *Magnetic Reading* on students' reading achievement?

In small, rural schools in Iowa, Grades 3–5 students using *Magnetic Reading* scored, on average, nearly 11 points higher on their spring Diagnostic compared to similar students who did not use *Magnetic Reading* (see Table 2). Considering we expect the impact of *Magnetic Reading* to vary by student, the confidence interval of the program's effect is a useful metric to consider. The confidence interval represents the plausible range of scale score points *Magnetic Reading* students should expect to score higher than their similar peers who do not use *Magnetic Reading*. In this sample and context, the impact of *Magnetic Reading* may range from a nearly five-point to 16.5-

[3]Although "relative risk" is typically the language used in the literature, in this case the language of relative risk is non-intuitive because we hope students do achieve their growth targets. Instead, we interpret findings as the "relative benefit" of *Magnetic Reading* for meeting Typical Growth or Stretch Growth targets.

point advantage on students' spring Diagnostic scores. This finding accounts for students' fall Diagnostic scores, average weekly time spent using *i-Ready* in the academic year, number of weeks using *i-Ready* in the academic year, grade, and the relationship between grade and fall Diagnostic scores.

The Glass's Delta, or standardized effect size, associated with the 11-point scale score benefit of *Magnetic Reading* is .23 standard deviations. Based on Lipsey and colleagues' (2012) and Kraft's (2020) research on the effects of educational interventions, the standardized effect size of *Magnetic Reading* is considered moderate and practically meaningful. See the Appendix for the full table of results from the impact model.

Table 2. Impact of *Magnetic Reading* on Spring Reading Achievement

	Diagnostic Scale Score Difference	Standard Error	Confidence Interval	Glass's Delta
Magnetic Reading	10.59***	3.01	4.68–16.49	.23

Note: Standard error represents the cluster-robust standard error.
Note: ***p <= .001.
Note: Confidence interval is calculated using α = .05.

To further contextualize what the difference in average scores between groups meant for students, placement change from fall to spring was also examined. Notably, 71% of *Magnetic Reading* students who placed One Grade Level Below in the fall placed on or above grade level in the spring. This is compared to 56% of comparison students who ended On or Above Grade Level when starting the year One Grade Level Below. Additionally, more than two-thirds (i.e., 68%) of the *Magnetic Reading* students who placed Two or More Grade Levels Below in the fall ended the year either One Grade Level Below or On or Above Grade Level. This is compared to just more than half (i.e., 53%) of comparison students with the same change in placement. Although sample sizes are small and percentages should be interpreted with caution, these results are promising and suggest *Magnetic Reading* may help Grades 3–5 students approach or reach grade-level expectations. See Table 3 for the placement transition table. Percentages represent the percentage of students in either the *Magnetic Reading* or comparison group in each placement category.

Table 3. Placement Transition from Fall to Spring by Group

Magnetic Reading **Spring Diagnostic Placement**					
		n	Early On Grade Level or Above	One Grade Level Below	Two or More Grade Levels Below
Fall Diagnostic Placement	Early On Grade Level or Above	65	94%	6%	0%
	One Grade Level Below	78	71%	27%	3%
	Two or More Grade Levels Below	71	20%	48%	32%
Comparison Spring Diagnostic Placement					
		n	Early On Grade Level or Above	One Grade Level Below	Two or More Grade Levels Below
Fall Diagnostic Placement	Early On Grade Level or Above	65	95%	5%	0%
	One Grade Level Below	78	56%	40%	4%
	Two or More Grade Levels Below	71	18%	35%	47%

Note: Percentages represent the percentage of students within a *Magnetic Reading* or comparison group who began the year in a given placement and ended in a given placement. Sample interpretation: Of the *Magnetic Reading* students who started Two or More Grade Levels Below in the fall, 20% experienced growth that resulted in on-grade level placement in the spring.

Do more students meet their Typical Growth and Stretch Growth targets when using *Magnetic Reading*?

Significantly more *Magnetic Reading* learners met their Typical Growth and Stretch Growth targets than similar comparison students. Three-fourths of *Magnetic Reading* students met their Typical Growth goals compared to only 61% of comparison students. 41% of *Magnetic Reading* students met their Stretch Growth goals compared to only 29% of comparison students. This means that of 214 students in each group, 31 more *Magnetic Reading* students met their Typical Growth targets, and 26 more *Magnetic Reading* students met their Stretch Growth targets compared to their non-*Magnetic Reading* peers. Furthermore, students using *Magnetic Reading* were 1.2 times more likely to meet their Typical Growth targets and 1.4 times more likely to meet their Stretch Growth targets than their similar peers. See Table 4 for the counts and percentages of students who met their Typical Growth and Stretch Growth targets. These results suggest that *Magnetic Reading* may support students in meeting their growth targets as well as support students' growth toward grade-level proficiency.

Table 4. Percentage of Students Who Met Typical or Stretch Growth Targets by Group

	Typical Growth		Stretch Growth	
	Did Not Meet Typical Growth Target	Met Typical Growth Target	Did Not Meet Stretch Growth Target	Met Stretch Growth Target
Magnetic Reading	53 (25%)	161 (75%)	127 (59%)	87 (41%)
Comparison	84 (39%)	130 (61%)	153 (71%)	61 (29%)

Note: Typical Growth [$\chi 2(1, N = 428) = 10.32, p = .001, \varphi = 0.16$]; Stretch Growth [$\chi 2(1, N = 428) = 6.98, p = .008, \varphi = 0.13$]

Discussion

Acceleration toward Grade-Level Expectations

The findings from this study are promising and reflect that *Magnetic Reading* has a positive impact on students' reading achievement. In particular, these findings suggest that *Magnetic Reading* may help more students move toward or achieve grade-level proficiency. More *Magnetic Reading* than comparison students increased in grade-level placement from fall to spring. A similar proportion of comparison students advanced one grade-level placement as we observed nationally in the 2021-2022 academic year. That is, more *Magnetic Reading* students advanced in one grade-level placement than we would expect based on the national *i-Ready* population, further highlighting the benefits of *Magnetic Reading* in supporting students in their growth.

In addition, findings suggest that *Magnetic Reading* supports students in meeting their growth target goals. These findings are particularly relevant as educators, students, and their caregivers continue to navigate teaching and learning after disruptions related to the COVID-19 pandemic. Data for this study were from the 2021–2022 school year, just one school year after teaching and learning disruptions due to the COVID-19 pandemic in the 2019-2020 and 2020-2021 school years. We know that the teaching and learning disruptions—as well as the mental, emotional, and financial toll of the pandemic—negatively impacted students' educational experiences and resulted in unfinished learning that is unprecedented in our modern education system. Moreover, we know the negative effects of the pandemic on education were more severe for marginalized groups of students, namely students in schools who serve primarily Latino, Black, and socioeconomically disadvantaged students (Curriculum Associates, 2021b) as well as students who are below grade level (Dawson, 2022; Lewis et al., 2022). Consequently, it is more important than ever to ensure learners have effective instructional materials that not only support their knowledge and skills development but also put them on the path toward grade-level proficiency and support their acceleration of gaining grade-level skills and knowledge.

Curriculum Associates' Stretch Growth makes tangible for educators, students, and their caregivers what goals students must meet to achieve grade-level skills and knowledge. Stretch Growth targets are ambitious, and meeting Stretch Growth puts students on the path toward grade-level proficiency. In fact, for Grades 3–5 students who meet Stretch Growth two years in a row, the vast majority reach grade-level proficiency at the end of the second year (Rome & Daisher, 2022). While Stretch Growth targets are attainable with the appropriate scaffolds and support, few students meet Stretch Growth in one school year, let alone in two consecutive years. As such, we need to understand how to better support students in meeting their Stretch Growth targets. This study demonstrates that *Magnetic Reading* may help more students meet their Stretch Growth targets and move toward or achieve grade-level proficiency. In the 2017–2018 and 2018–2019 school years, 28% to 34% of students in Grades 3–5 met their Reading Stretch Growth targets in a single year (Rome & Daisher, 2023). In this study, 41% of students who used *Magnetic Reading* met their Stretch Growth targets. Their peers who did not use *Magnetic Reading* met Stretch Growth at levels on par with those observed historically. Moreover, an additional 26 students who used *Magnetic Reading* met Stretch Growth compared to the number of comparison students who met Stretch Growth. Twenty-six students is more than one classroom of students who are now on track toward grade-level proficiency and who may not have been without the opportunity to use *Magnetic Reading*. *Magnetic Reading* provides an opportunity for educators to support students' Stretch Growth attainment and thus an opportunity for students to achieve the educational outcomes to which they aspire.

Limitations and Future Directions

This research provides encouraging evidence for *Magnetic Reading* and the academic benefits of the program for students in Grades 3–5. Yet, it is important to acknowledge that this study used a small and specific sample relative to the national population of Grades 3–5 students in the United States. Students in this sample were primarily White and attended small, rural, Title I-eligible schools in Iowa. Future research will focus on broadening our evaluation of *Magnetic Reading* by including students and schools with a wider range of diverse characteristics and identities.

We know that implementations of *Magnetic Reading* vary by schools and classrooms. Moreover, we know the other curricular materials used in addition to *Magnetic Reading* vary by schools and classrooms. In this study, we did not have information regarding classrooms' implementation of *Magnetic Reading* nor about their other programs or materials they used alongside of *Magnetic Reading*. Because we could not take classroom-level information into account, it is possible the effects of *Magnetic Reading* presented in this paper could be misattributed to unknown and unmeasured factors in the classroom.

Conclusion

This study answered two questions related to the efficacy of *Magnetic Reading*, a print-based reading comprehension program for students in Grades 3–5:

1. What is the effect of *Magnetic Reading* on Grade 3–5 students' reading achievement?
2. Do more students meet their Typical Growth and Stretch Growth targets when using *Magnetic Reading*?

The results presented in this paper illustrate that *Magnetic Reading* has a positive and significant impact on Grades 3–5 students' reading achievement and growth. For students in small, rural, Title I-eligible schools in Iowa, *Magnetic Reading* had a positive influence on their *i-Ready Diagnostic* for Reading achievement and their growth toward meeting grade-level expectations. Students who used *Magnetic Reading* scored nearly 11 points higher on their spring Diagnostic, more *Magnetic Reading* students met or progressed toward grade-level expectations, and more *Magnetic Reading* students met their Typical Growth and Stretch Growth targets in Reading compared to similar peers who did not use *Magnetic Reading*. Notably, these results represent the impact of *Magnetic Reading* above and beyond the benefits students likely already experience from *i-Ready* usage.

Although this was a small study, the findings are promising and suggest that *Magnetic Reading* supports students in achieving their educational goals. Future research will commit to broadening our understanding of the impact of *Magnetic Reading* for all students across educational contexts, school environments, and student identities.

References

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

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

Curriculum Associates. (2023). *Variation in* i-Ready *growth: Is Typical Growth enough for students who have fallen behind?* Curriculum Associates. https://www.curriculumassociates.com/research-and-efficacy/i-ready-stretch-growth-two-grades-below-trends

Dawson, M. (2022). *Student growth during COVID-19: Grade-level readiness matters*. Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/student-growth-during-covid-19.pdf

Glass G. V., McGaw B., Smith M. L. (1981). *Meta-analysis in social research*. Sage.

Harris, H. & Horst, S. J. (2016). A brief guide to decisions at each step of the propensity score matching process. *Practical Assessment, Research, and Evaluation, 21*(4). https://doi.org/10.7275/yq7r-4820

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

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241–253. https://journals.sagepub.com/doi/full/10.3102/0013189X20912798

Lewis, K., Kuhfeld, M., Langi, M., Peters, S., & Fahle, E. (2022). *The widening achievement divide during COVID-19*. NWEA Center for School and Student Progress. https://www.nwea.org/uploads/2022/11/CSSP-Brief_Widening-achievement-divide-COVID-19.pdf

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

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114–140. https://psycnet.apa.org/doiLanding?doi=10.1037%2Fmet0000078

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

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. (2020). *Impact evaluation of reading.* i-Ready *for striving learners using 2018–19 data: Final report*. (Report No. 053). Human Resources Research Organization. https://files.eric.ed.gov/fulltext/ED610441.pdf

Robitzsch, A., & Grund, S. (2022). *miceadds: Some additional multiple imputation functions, especially for 'mice.'* R package version 3.15-21. https://CRAN.R-project.org/package=miceadds

Rome, L. & Daisher, T. (2022). i-Ready Stretch Growth *as a path toward proficiency*. Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/ireadytwoyearstretchgrowthresearchtechnicalreview20220913-1.pdf

Rome, L. & Daisher, T. (2023). i-Ready Stretch Growth *in the pandemic context*. Curriculum Associates. https://www.curriculumassociates.com/research-and-efficacy/i-ready-stretch-growth-post-pandemic

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Swain, M., Randel, B., & Dvorak, R. N. (2019). *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

US Department of Education. (2022). Institute of Education Sciences, National Center for Education Statistics. https://nces.ed.gov/ccd/ccddata.asp

Warnes, G. R., Bolker, B., Lumley, T., & Johnson, R. C. (2022). *Gmodels: Various R programming tools for model fitting*. R package version 2.18.1.1. https://CRAN.R-project.org/package=gmodels

What Works Clearinghouse. (2022). *Procedures and standards handbook, version 5.0*. US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5_0-0-508.pdf

Appendix

Table A.1. Impact Model Parameter Estimates

	Category	Unstandardized Estimate	Standard Error	*p*
Intercept		523.98	3.27	<.001
Fall Diagnostic Score		.72	.03	<.001
***Magnetic Reading* Usage**		10.59	3.01	<.001
Grade (Reference = Grade 3)	Grade 4	7.77	6.45	.23
	Grade 5	-3.15	5.24	.55
Average Weekly Time Spent on *i-Ready* (Minutes)		.23	.26	.37
Number of Weeks Using *i-Ready*		.52	.26	.05
Fall Diagnostic Score by Grade Interaction (Reference = Grade 3)	Grade 4	-.02	.06	.78
	Grade 5	.09	.06	.13

Note: Fall Diagnostic score was centered at 500, average weekly time spent on *i-Ready* was centered at 30 minutes, and number of weeks using *i-Ready* was centered at 18 minutes.

Note: Standard errors represent the cluster-robust standard errors.