

Really Great Reading: Examining Impacts among Urban Charter School Students

Study Type: ESSA Evidence Level II

Prepared for:
Really Great Reading

Prepared by LearnPlatform:
Alexandra Lee, Ph.D., Researcher, LearnPlatform
Meetal Shah, Ph.D., Senior Researcher, LearnPlatform

June 27, 2023



EXECUTIVE SUMMARY

Really Great Reading (RGR) contracted with LearnPlatform by Instructure, a third-party edtech research company, to examine the impact of usage of its reading program on student literacy outcomes. LearnPlatform designed the study to satisfy Level II requirements (Moderate Evidence) according to the Every Student Succeeds Act (ESSA).

Study Sample, Measures, and Methods

This study occurred during the first two school years of RGR implementation: 2021-22 (Year 1) and 2022-23 (Year 2). The sample included 1,117 students who were in Grades 1-3 in the 2021-22 school year (Year 1) from across 13 schools in one urban charter school network in the eastern U.S. In terms of demographics, the charter school network was predominantly Black (99%). Thirteen percent of the students had individualized educational programs (IEPs) and one percent were English language learners (ELLs).

Researchers identified students¹ who received RGR instruction (i.e., an evidence-based foundational science of reading phonics and literacy curriculum) by verifying that their primary teacher completed the requisite RGR professional development and logged into the RGR platform. Reading achievement was measured using NWEA MAP® Reading Growth scores. Taken together, these measures allowed researchers to investigate patterns in RGR implementation and potential impacts of program use on students' reading skills².

Researchers used a variety of quantitative analytic approaches to answer the research questions. First, researchers used descriptive statistics to examine participant characteristics and implementation of the program. Researchers then conducted partial correlations to examine whether RGR usage was significantly associated with reading scores for both years of program implementation examined in this study. Next, a series of linear regressions were used to examine whether Year 1 RGR use (i.e., implemented in fall 2021) related to significant differences in student reading scores in spring 2022, fall 2022, and winter 2023 and whether Year 2 RGR use (i.e., continued implementation in fall 2022) related to significant differences in winter 2023 reading scores. The partial correlation and regression analyses included student-level covariates to control for potential selection bias (i.e., baseline test scores, gender, special education status). In addition, researchers calculated standardized effect sizes (i.e., Cohen's *d* and Hedge's *g*) to determine the magnitude of changes in student outcomes.

¹ Students were also considered to be “users” of RGR if their teacher met these criteria and therefore, were deemed users of the program.

² The MAP Reading Growth assessment measured students' foundational reading skills broadly and included three distinct domains: phonological awareness, phonics and word recognition, and language comprehension.

Student Outcomes



Grade 1 students who received Really Great Reading (RGR) instruction had significantly higher NWEA MAP® Reading scores at the end-of-the-year than students who did not receive RGR-aligned instruction in 2021-2022.



Grade 2 – 3 students who received RGR instruction in the prior school year (2021-2022) had significantly higher NWEA MAP® Reading scores at the beginning of the following school year (fall 2022) than students who did not receive RGR instruction in 2021-2022. Notably, Grade 2 students who received RGR instruction did not have the expected summer slide in test scores.



Grade 2 students who received RGR instruction in the prior school year (2021-2022) also had significantly higher NWEA MAP® Reading scores at mid-year (winter 2023) than students who did not receive RGR instruction in 2021-2023.



Teachers' use of RGR instruction in the 2021-2022 school year was positively associated with reading outcomes for Grade 1 – 2 students' reading scores in spring 2022.

Conclusions

This study provides results to satisfy ESSA evidence requirements for Level II (Moderate Evidence) given the quasi-experimental study design, positive statistically significant findings, and large sample size.

TABLE OF CONTENTS

Introduction	3
Methods	4
Program Implementation Findings	7
Program Effectiveness Findings	9
Conclusions and Recommendations	13
References	14
Appendix A. Really Great Reading Logic Model	15
Appendix B. Additional Information on Study Design and Methods	16
Appendix C. Additional Information on K – 3 Outcome Findings	20

Introduction

Recognizing that 65% of students cannot read proficiently by 4th grade (McFarland et al., 2019), which is due in part to the enduring research-practice gap (Schneider, 2018), Really Great Reading (RGR) provides teachers with the tools to implement evidence-based reading instruction rooted in the science of reading to help students develop foundational literacy skills including phonics, phonemic awareness, and orthographic mapping.

As part of their ongoing efforts to demonstrate the efficacy of its literacy programs, RGR contracted with LearnPlatform by Instructure, a third-party edtech research company, to examine the relationship between usage of its program and student outcomes. After collaborating on the development of an updated logic model (Appendix A) for RGR (Lee et al., 2023), LearnPlatform designed a study to satisfy ESSA Level II requirements (Moderate Evidence) with the following research questions.

Program Implementation Research Questions

1. Among teachers, what were the usage patterns of RGR resources?
 - a. Did teachers complete professional development on implementation of RGR programs?
 - b. Did teachers log in to online teacher tools?

Effectiveness Research Questions

2. After controlling for students' prior literacy levels, how was teachers' use of RGR related to student performance on standardized literacy assessments?
3. After controlling for students' prior literacy levels, what is the magnitude of observed differences of students who received RGR instruction compared to students who did not receive RGR instruction on standardized literacy assessments?

Methods

This section of the report briefly describes the setting, participants, measures, and analysis methods.

Setting

The study included one charter school network in a large city in the eastern U.S. during the first two school years of RGR implementation: 2021-22 (Year 1) and 2022-23 (Year 2). The sample included 1,117 students who were in Grades 1-3 in the 2021-22 school year (Year 1) from across 13 schools.

Participants

Seventy-nine percent ($n = 886$) of the students in the sample ($n = 1,117$) received literacy instruction by a teacher who completed RGR professional development and logged into the RGR platform; these students were considered to be in the intervention group for the purpose of the comparative analyses. The remaining 21% ($n = 231$) students were considered to be in the control group as they received literacy instruction from a teacher who had never completed RGR professional development or logged into the RGR platform.

There were no significant differences in terms of demographics between students who used RGR (i.e., intervention group) and those who did not use RGR (i.e., control group)³. The grade levels were similar across both conditions. Specifically, among students in the intervention group, 36% were in Grade 1, 32% were in Grade 2, and 32% were in Grade 3; and, among students in the control group, 32% were in Grade 1, 31% were in Grade 2, and 37% were in Grade 3 in Year 1 (2021-22). These same students were followed up with in Year 2 (2022-23) when they were in Grades 2 – 4. In terms of demographics, the charter school network was predominantly Black (99%). Thirteen percent of the students had individualized educational programs (IEPs) and one percent were English language learners (ELL).

Measures

This study included the following measures to provide insights into Really Great Reading (RGR) implementation and evidence about the potential impacts of the program on student outcomes.

RGR Usage Metrics. Researchers utilized 2021-22 and 2022-23 teacher-level usage data (i.e., RGR professional development completion and RGR platform logins) to determine which teachers used RGR in their literacy instruction, and therefore, identify which students were RGR users and which students were not. According to RGR, professional development includes online training that prepares teachers to successfully implement the programs for different grade levels [i.e., the programs are named “Countdown” (Kindergarten), “Blast” (Grade 1), and “HD Word” (Grades 2-3)]. Each program-

³ According to chi-square (χ^2) tests performed comparing demographics by condition (see Baseline Equivalence below).

specific training includes five self-paced, interactive courses, each with multiple modules of content. The first two courses of the training are designed to be completed prior to beginning RGR instruction and are comparable to a full day of face-to-face training. The remaining three courses are designed to provide teachers with incremental professional development related to practices for teaching reading effectively and are often completed after instruction has begun. Given that the completion of the first two courses is necessary for successful implementation of RGR, it was used as one of the criteria to classify teachers as users and non-users of the program.

RGR platform logins were used by researchers as the second criterion to classify teachers as users and non-users of RGR. A login indicates that the teacher used the online presentation tool which is a required component of RGR's Lesson Plan Teacher Guide Set and helps teachers deliver instruction with minimal preparation.

Standardized Student Assessments. Researchers used NWEA MAP® Reading Growth scores, a standardized research-based assessment that reliably measures reading ability and progress from Kindergarten through Grade 12. The MAP® Reading Growth assessment measured students' foundational reading skills broadly and included three distinct domains: phonological awareness, phonics and word recognition, and language comprehension. The concepts assessed by the MAP® Reading Growth assessment are standards-aligned and change for each year. Specifically, the assessment used for Kindergarten – Grade 2 students has an emphasis on phonological awareness and phonics, whereas the assessment used for Grade 3 – Grade 4 students has an emphasis on comprehension and literary and informational concepts. In other words, the assessment used for this study measured broader literacy skills than those specifically targeted by RGR's curriculum for all grades and the emphasis on concepts not covered by RGR increased as students progressed through school.

Data Analysis

Researchers used a variety of quantitative analytic approaches to answer the research questions. First, researchers used descriptive statistics to examine student characteristics and implementation of the program. Researchers then conducted partial correlations to examine whether RGR usage was significantly associated with reading scores for both years of program implementation. Next, a series of linear regressions were used to examine whether Year 1 RGR use (i.e., implemented in fall 2021) related to significant differences in student reading scores in spring 2022, fall 2022, and winter 2023 and whether Year 2 RGR use (i.e., continued implementation in fall 2022) related to significant differences in winter 2023 reading scores. The partial correlation and regression analyses included student-level covariates to control for potential selection bias (i.e., baseline test scores, gender, special education status). In addition, researchers calculated standardized effect sizes (Cohen's *d* and Hedge's *g*) to determine the magnitude of changes in student outcomes.

Baseline Equivalence

To ensure the validity of the study's findings and to adhere to ESSA Level II standards, the researchers assessed the equivalence of student demographic characteristics (i.e., ethnicity, gender, ELL status,

and IEP status) and standardized assessment scores at baseline between intervention and control groups. Grades 1 – 3 students who used RGR were not statistically significantly different from students who did not use RGR in terms of ethnicity ($\chi^2 = 4.35, p = .630$), gender ($\chi^2 = 3.43, p = .064$), ELL status ($\chi^2 = 1.64, p = .200$), or IEP status ($\chi^2 = 0.36, p = .547$).

Fall 2021 NWEA MAP® Reading Growth scores were examined separately for each grade level to verify that there were no significant differences between students who used RGR and those who did not prior to the first use of RGR for literacy instruction by the charter school network. The results of ANOVA tests indicated that the differences at baseline in terms of NWEA MAP® Reading Growth scores were within the boundary for statistical adjustment according to the What Works Clearinghouse (WWC) criteria (WWC, 2022). Specifically, statistically significant baseline differences with an effect size (measured by Hedge's g) between 0.05 and 0.25 are acceptable with statistical adjustment, which was demonstrated by the ANOVA tests (Grade 1: Hedge's $g = .21, p = .047$; Grade 2: Hedge's $g = .12, p = .275$; Grade 3: Hedge's $g = .09, p = .531$). Due to these results, fall 2021 reading scores were statistically controlled for in the final models. See Appendix B for more details regarding baseline equivalence.

Program Implementation Findings

RGR program implementation involved teachers completing professional development (PD) prior to initial use of the program and then logging into the RGR platform. Overall, Grade 1-3 teachers completed an average of 2.5 RGR PD courses (SD = 2.0) over the school year, with some teachers completing no courses and others completing as many as 9 courses (across various programs and all years included in the study). Most teachers completed the requisite PD in the first year of RGR program implementation (77% in Year 1, Figure 1). Additionally, a majority of teachers in the charter school network logged into the RGR platform during program implementation (79% in Year 1). The students of these teachers composed the intervention group.

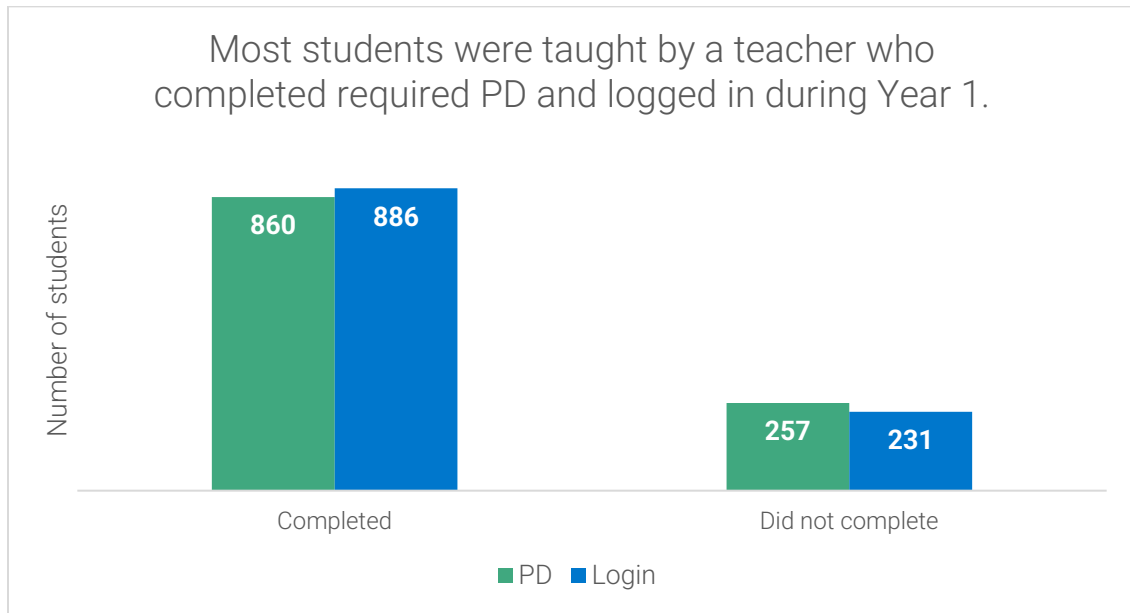


Figure 1. Overall distribution of RGR professional development completion and logins by Grade 1-3 teachers for first year of implementation (2021-2022).

During Year 1, the 13 schools in the sample had different usage of RGR, with three schools having a mixture of students who received RGR instruction and students who did not, one school where no students received RGR instruction, and nine schools where all students received RGR instruction. (Figure 2).

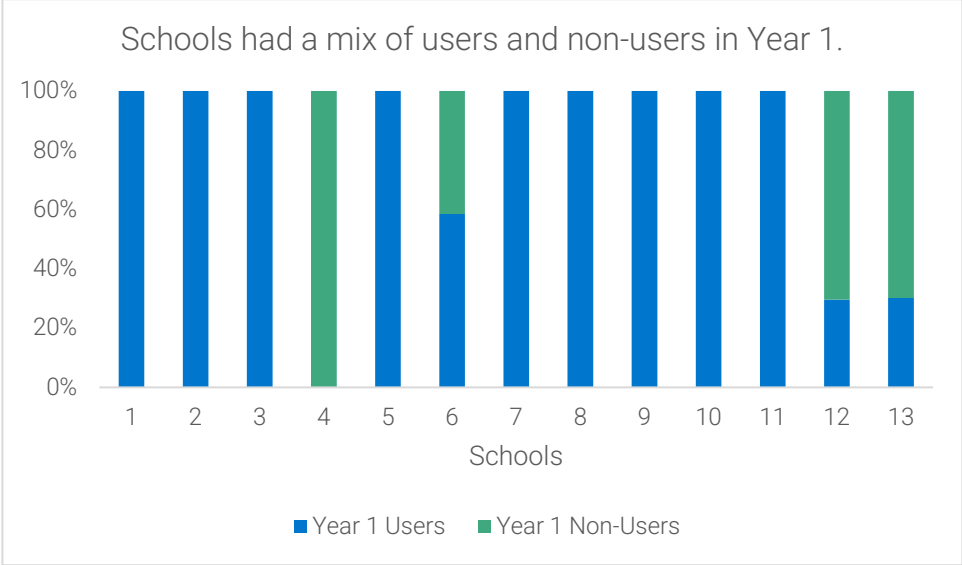


Figure 2. Overall distribution of RGR users (i.e., Grade 1-3 students who had a teacher that completed PD and logged in) by school for first year of implementation.

Program Effectiveness Findings

To answer the remaining study research questions, researchers used descriptive statistics, partial correlations, and regressions. In addition to examining the statistical significance of the tests used, researchers calculated standardized effect sizes to determine the magnitude of the relationship between RGR usage and student literacy outcomes. The key study findings are included below, and the full set of results can be found in Appendix C.

Relationship Between Receiving RGR Instruction in Grades 1 – 3 with Students’ Outcomes on NWEA MAP® Reading Growth Assessment

Researchers used partial correlation tests to examine whether there was a significant relationship between RGR usage and students’ reading outcomes during the first two years of implementation. The partial correlations controlled for students’ fall 2021 reading scores (baseline), IEP-status, and gender. Including these covariates allowed for increased statistical precision and interpretation of the relationship of RGR with reading scores above and beyond students’ prior achievement (Fortson et al., 2014). Overall, the results of the partial correlations demonstrated a significant positive relationship between receiving RGR instruction and students’ reading achievement as measured by the NWEA MAP® Reading scores.

Key Finding 1. RGR usage among **Grade 1** students in Year 1 (2021-22) was significantly associated with higher NWEA MAP® Reading Growth scores for the subsequent three semesters. In other words, receiving RGR-aligned instruction in Grade 1 was significantly related to higher reading scores not only in Grade 1, but also in Grade 2.

Among Grade 1 students, receiving RGR instruction was significantly related to higher reading scores	
Timepoint	Partial Correlation Coefficient
Spring 2021	.19
Fall 2022	.19
Winter 2023	.18

Note. The Pearson’s r correlation coefficient indicates the degree to which there is a linear relationship between RGR instruction and reading scores after controlling for baseline reading scores, IEP-status, and gender on a scale of -1 to 1. Green indicates the partial correlation test was statistically significant.

Key Finding 2 RGR usage among **Grade 2** students in Year 1 (2021-22) was significantly associated with higher NWEA MAP® Reading Growth scores for the subsequent two semesters. In other words, receiving RGR-aligned instruction in Grade 2 was significantly related to higher reading scores not only in Grade 2 (i.e., Spring 2021), but also in **Grade 3** (i.e., Fall 2022).

Among Grade 2 students, receiving RGR instruction was significantly related to higher reading scores	
Timepoint	Partial Correlation Coefficient
Spring 2021	.12
Fall 2022	.19
Winter 2023	.08

Note. The Pearson's *r* correlation coefficient indicates the degree to which there is a linear relationship between RGR instruction and reading scores after controlling for baseline reading scores, IEP-status, and gender on a scale of -1 to 1. Green indicates the partial correlation test was statistically significant.

Differences in Literacy Outcomes Among Grade 1 – 3 Students who Received RGR Instruction and Students Who Did Not

To understand whether RGR positively impacted students' reading outcomes, researchers conducted a series of multiple regression analyses for three timepoints controlling for baseline reading scores (fall 2021), IEP-status, and gender. Overall, all students showed growth across the four time points in Grades 1-3, but Grades 1-2 students who received RGR-aligned reading instruction in Year 1 had significantly greater growth compared to those who did not receive RGR-aligned instruction.

Key Finding 3 Grade 1 students whose teacher used RGR-aligned reading instruction had significantly higher NWEA MAP® Reading Growth scores compared to students who did not receive RGR-aligned instruction (Figure 3). Specifically, Grade 1 students who received RGR-aligned reading instruction in the 2021-22 school year, had significantly higher reading scores in spring 2022 (Hedge's $g = .29$, 9 percentile points), fall 2022 ($g = .32$, 11 percentile points), winter 2023 ($g = .30$, 10 percentile points)⁴.

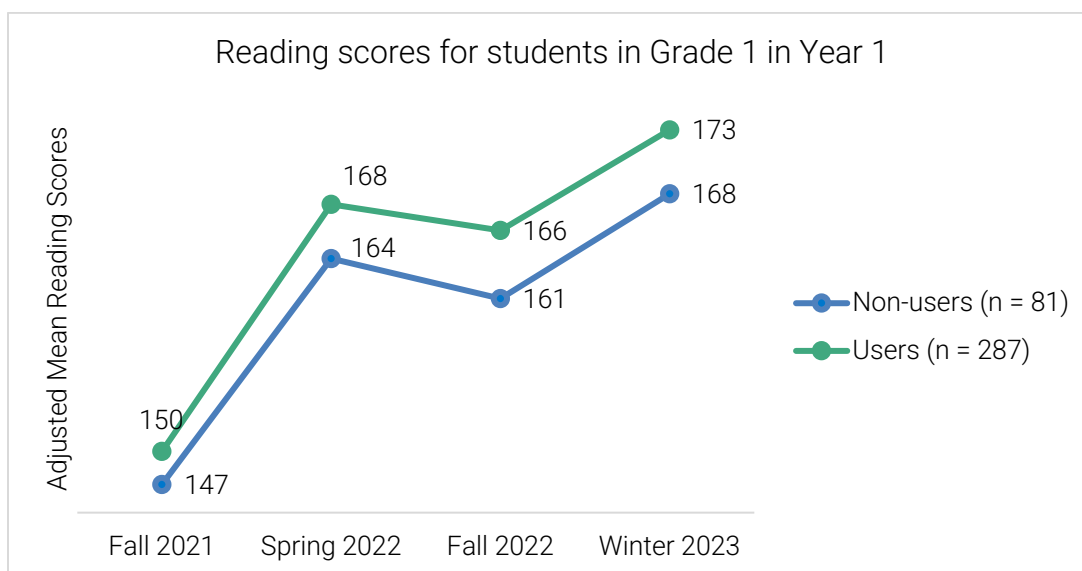


Figure 3. There was a significantly greater increase in adjusted NWEA MAP® Reading scores from fall 2021 to winter 2023 for students who received RGR instruction in Grade 1 in the 2021-22 school year compared to those who did not.

⁴ See Appendix C for Cohen's d effect size estimates. When interpreting the magnitude of the effect sizes using these estimates (i.e., Cohen's d and Hedge's g), it is important to keep in mind that a broad reading achievement measure was used, and the study was conducted in the field (vs. a lab setting); these conditions are associated with smaller effect sizes. Thus, the standards for interpreting field experiment effect sizes should be applied, whereby an effect size larger than .10 - .15 should be considered large and substantive (Hill et al., 2008; Kraft, 2020).

Key Finding 4. Grade 2 students whose teacher used RGR-aligned reading instruction had significantly higher NWEA MAP® Reading scores compared to students who did not receive RGR-aligned instruction (Figure 4). Specifically, Grade 2 students who received RGR-aligned reading instruction in the 2021-22 school year, had positive trending scores in spring 2022 (Hedge’s $g = .19$, 6 percentile points), and significantly higher scores in fall 2022 ($g = .37$, 13 percentile points). The reading scores for these students did not decline over the summer in contrast to the typical pattern (i.e., “summer slide”, Quinn & Polikoff, 2017).

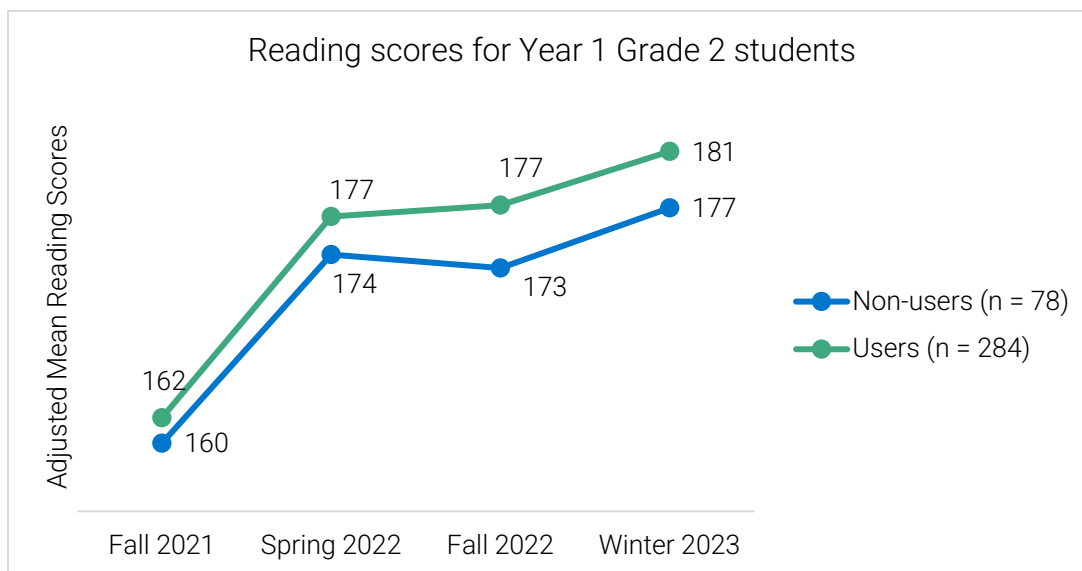


Figure 4. There was a significantly greater increase in adjusted NWEA MAP® Reading scores from fall 2021 to fall 2023 for students who received RGR instruction in Grade 2 in the 2021-22 school year compared to those who did not.

Researchers also investigated whether there was a positive impact of receiving RGR-aligned instruction for Grade 3 students and did not find any significant results (see Appendix C).

Conclusions and Recommendations

In sum, the results of this study suggest that there is a positive effect of Really Great Reading-aligned literacy instruction on Grade 1-2 students' reading outcomes. The data indicate that students who received RGR instruction had substantial and significant positive effects on reading outcomes that lasted into the subsequent school year (i.e., Grades 2-3) compared to students who do not receive RGR instruction. It is also notable that this study took place in a context that serves students traditionally underrepresented in education research (i.e., predominantly Black students) following the COVID-19 pandemic and suggests that using evidence-based strategies such as RGR may be an important step forward to address learning loss and disparities among students (Dorn et al., 2020; Horsford et al., 2021). This is particularly true as RGR was found to limit the learning loss typically seen over the summer months (i.e., "summer slide," Quinn & Polikoff, 2017).

Given the positive outcome findings of the impact analysis among the sample, this study provides results to satisfy ESSA evidence requirements for Level II (*Moderate Evidence*). Specifically, this study met the following criteria:

- ✓ Quasi-experimental design
- ✓ Proper design and implementation
- ✓ Statistical controls through covariates
- ✓ At least one statistically significant, positive finding

Researchers recommend the following next steps for the RGR team:

- seek to gather more detailed data about the different ways educators use the program (e.g., types and frequency of RGR learning activities completed by students) to inform specific implementation recommendations;
- consider recruiting a comparison district for Kindergarten-Grade 4 students to better understand how elementary school students who received RGR instruction compared to elementary school students who received reading instruction from other programs; and,
- work to partner with a learning education agency that uses a literacy assessment more closely aligned to what RGR teaches (e.g., DIBELS) as the broader reading assessment used in present study (i.e., MAP® Reading Growth) measures skills beyond what RGR teaches (e.g., literary concepts) and therefore may have diminished the evidence of effectiveness.

Acknowledgements

The authors would like to extend their deepest thanks to Avery Wall who supported the preparation of this report in numerous ways.

References

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). COVID-19 and learning loss—disparities grow and students need help. *McKinsey & Company, December, 8*, 6-7.

Fortson, K., Gleason, P., Kopa, E., & Verbitsky-Savitz, N. (2015). Horseshoes, hand grenades, and treatment effects? Reassessing whether nonexperimental estimators are biased. *Economics of Education Review, 44*, 100-113.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*, 172-177.

Horsford, S. D., Cabral, L., Touloukian, C., Parks, S., Smith, P. A., McGhee, C., Qadir, F., Lester, D., & Jacobs, J. (2021). Black education in the wake of COVID-19 and systemic racism: Toward a theory of change and action. *Black Education Research Collective*. Teachers College, Columbia University.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*, 241-253.

Quinn, D., & Polikoff, M. (2017). *Summer learning loss: What is it, and what can we do about it*. Brookings Institution.

What Works Clearinghouse. (2022). What Works Clearinghouse procedures and standards handbook, version 5.0. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at <https://ies.ed.gov/ncee/wwc/Handbooks>

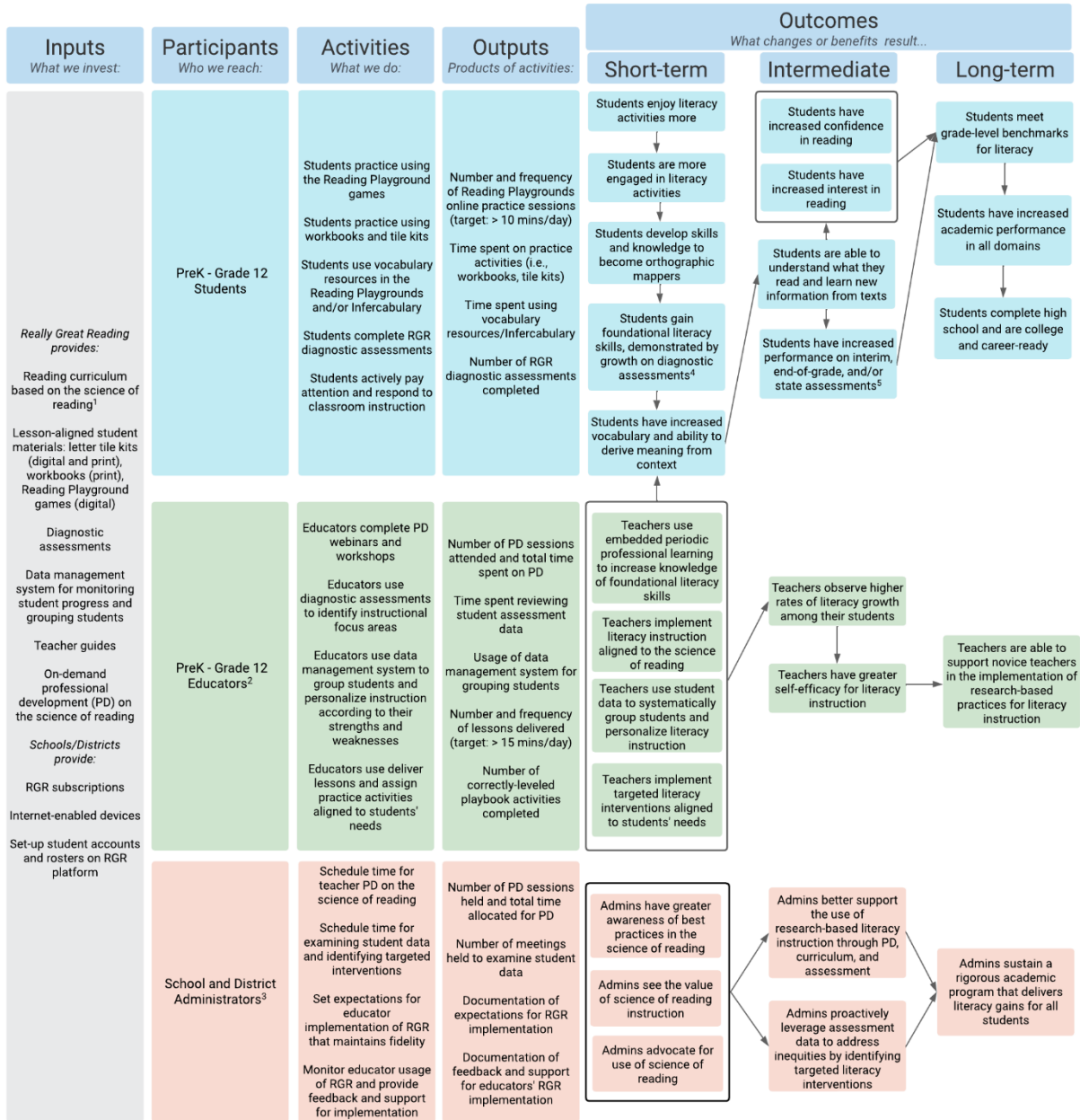
Woods, A. D., Morgan, P. L., Wang, Y., Farkas, G., & Hillemeier, M. M. (2023). Effects of having an IEP on the reading achievement of students with learning disabilities and speech or language impairments. *Learning Disability Quarterly, 07319487231154235*.

Appendix A. Really Great Reading Logic Model

LOGIC MODEL



Problem Statement: Sixty-five percent of students are not able to read proficiently by 4th grade, which is due in part to the enduring research-practice gap. Really Great Reading (RGR) provides teachers with the tools to implement research-based science of reading instruction, which helps students develop word-level literacy using phonics, phonemic awareness, orthographic mapping, and deciphering word meaning.



¹ The science of reading is a set of research-based practices that support the development of reading by helping students to relate written text to spoken language by focusing on malleable factors that underpin reading ability, including phonics, phonemic awareness, and decoding (Peterscher et al., 2020).

² Educators may include primary classroom teachers, interventionists, Title I coordinators, paraprofessionals, and special educators.

³ School and District Administrators may include literacy/ELA instructional coaches, curriculum specialists, special education directors, district-level PD directors, principals.

⁴ Foundational literacy skills for students in Pre-K - 2nd grade include phonemic awareness, alphabetic principals, and oral reading fluency, skills for students in 3rd grade or higher include decoding, word reading, and reading fluency.

⁵ Examples of interim, end-of-grade, state assessments include DIBELS, iReady, NWEA, and MAP.



Appendix B. Additional Information on Study Design and Methods

Propensity Score Weighting

To help make the Grade 1-3 student groups (i.e., students who received RGR-aligned instruction and students who did not) as comparable as possible, propensity score weights were calculated for each student. To calculate propensity scores, researchers conducted binary logistic regression with student group as the dependent variable and grade, gender, IEP status, and NWEA MAP® Reading scores from fall 2021 (baseline) as the covariates⁵. The probability was saved as a new variable. Weights were calculated by dividing one by the probability (one/probability). Students without a weight were dropped from the final analytic sample. All analyses that included students who did not receive RGR-aligned instruction included these weights.

Baseline Equivalence

Researchers conducted baseline equivalence analyses to determine whether there were baseline differences in characteristics between students who received RGR-aligned instruction and students who did not during the 2021–22 and 2022-23 school years (Year 1 and Year 2). Specifically, researchers used chi-square analyses on student-level demographics and linear regressions for NWEA MAP® Reading Growth scores for all the timepoints compared in the study.

Baseline Equivalence for Grade 1 Students (in Year 1)

As noted in Table B1, there were no statistically significant differences between Grade 1 (in Year 1) student groups in terms of gender and IEP status.

Table B1. Baseline Equivalence Analysis of Grade 1 Student-Level Demographics by User Group in Year 1

Characteristics	Users (n = 287)		Non-users (n = 81)		Chi-squared	p-Value
	Percent	N	Percent	N		
<i>Gender</i>						
Male	49	140	44	36	1.60	.281
Female	51	147	56	45		

⁵ English language learner status and race/ethnicity were not included in the models because there was no variability in these variables among the sample.

Characteristics	Users (n = 287)		Non-users (n = 81)		Chi-squared	p-Value
	Percent	N	Percent	N		
<i>Individualized Educational Program (IEP)</i>						
Yes	7	21	5	4	0.57	.452
No	93	266	95	77		

As presented in Table B2, results from the weighted regression analysis reveal that there was a statistically significant difference between Grade 1 student groups in fall 2021 (Year 1 baseline), with a Hedge’s *g* effect size indicating that researchers should include NWEA MAP® Reading Growth scores in the final models as a control variable since it was significantly associated with the outcome. Additionally, results from the weighted regression analysis reveal that there was a statistically significant difference between these student groups in the subsequent school year (i.e., fall 2022; Year 2 baseline), and the Hedge’s *g* effect size was too large to control for in the final models (i.e., > .25), so comparative analysis for Year 2 were not conducted for these students.

Table B2. Baseline Equivalence Analysis of Grade 1 Students by User Group

Outcome Variable	Coefficient	Standard Error	t-value	p-value	Effect Size
NWEA MAP® Reading scores fall 2021	2.95	1.48	1.99	.047*	0.21
NWEA MAP® Reading scores fall 2022	7.65	3.32	2.30	.022*	0.48

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Baseline Equivalence for Grade 2 Students (in Year 1)

As noted in Table B3, there were no statistically significant differences between Grade 2 (in Year 1) student groups in terms of gender and IEP status.

Table B3. Baseline Equivalence Analysis of Grade 2 Student-Level Demographics by User Group in Year 1

Characteristics	Users (n = 284)		Non-users (n = 78)		Chi-squared	p-Value
	Percent	N	Percent	N		
<i>Gender</i>						
Male	54	152	51	40	0.12	.726
Female	46	132	49	38		
<i>Individualized Educational Program (IEP)</i>						
Yes	14	41	14	11	0.01	.941
No	86	243	86	67		

As presented in Table B4, results from the weighted regression analysis reveal that there was not a statistically significant difference between Grade 2 student groups in fall 2021 (Year 1 baseline), with a Hedge’s *g* effect size indicating that researchers should include NWEA MAP® Reading Growth scores in the final models. Additionally, results from the weighted regression analysis reveal that there was a statistically significant difference between these student groups in the subsequent school year (i.e., fall 2022; Year 2 baseline), and the Hedge’s *g* effect size was too large to control for in the final models (i.e., > .25), so comparative analysis for Year 2 were not conducted for these students.

Table B4. Baseline Equivalence Analysis of Grade 2 Students by User Group

Outcome Variable	Coefficient	Standard Error	t-value	p-value	Effect Size
NWEA MAP® Reading scores fall 2021	2.21	2.02	1.09	.275	0.12
NWEA MAP® Reading scores fall 2022	7.69	2.17	3.54	<.001***	0.40

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Baseline Equivalence for Grade 3 Students (in Year 1)

As noted in Table B5, there were no statistically significant differences between Grade 3 (in Year 1) student groups in terms of gender and IEP status.

Table B5. Baseline Equivalence Analysis of Grade 2 Student-Level Demographics by User Group in Year 1

Characteristics	Users (n = 270)		Non-users (n = 72)		Chi-squared	p-Value
	Percent	N	Percent	N		
<i>Gender</i>						
Male	47	127	56	40	1.65	.199
Female	53	143	44	32		
<i>Individualized Educational Program (IEP)</i>						
Yes	14	38	18	13	0.71	.399
No	86	232	82	59		

As presented in Table B6, results from the weighted regression analysis reveal that there were not statistically significant differences between Grade 3 student groups in fall 2021 (Year 1 baseline) or fall 2022 (Year 2 baseline), with the Hedge’s *g* effect sizes indicating that researchers should include NWEA MAP® Reading Growth scores in the final models.

Table B6. Baseline Equivalence Analysis of Grade 3 Students by User Group

Outcome Variable	Coefficient	Standard Error	t-value	p-value	Effect Size
NWEA MAP® Reading scores fall 2021	1.35	2.15	0.63	.531	0.09
NWEA MAP® Reading scores fall 2022	1.21	2.21	0.55	.585	0.08

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix C. Additional Information on Grade 1-3 Outcome Findings

Relationship Between Receiving RGR Instruction in Grades 1 – 3 with Students’ Outcomes on NWEA MAP® Reading Assessment During the First Two Years of Implementation

Partial Correlation Analyses for Grade 1 Students (in Year 1)

Table C1. Spring 2022 NWEA MAP® Reading scores by receiving RGR instruction in Year 1 for Grade 1 students ($n = 368$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p-value
Year 1 RGR Instruction	.19	.005**
Fall 2022 Reading Score	.72	<.001***
IEP Status	-.15	.005**
Gender	-.10	.077

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Table C2. Fall 2022 NWEA MAP® Reading scores by receiving RGR instruction in Year 1 for Grade 1 students ($n = 374$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p-value
Year 1 RGR Instruction	.19	<.001***
Fall 2022 Reading Score	.69	<.001***
IEP Status	.01	.822
Gender	-.03	.555

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female.

Table C3. Winter 2023 NWEA MAP® Reading scores receiving RGR instruction in Year 1 for Grade 1 students ($n = 351$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 1 RGR Instruction	.18	.001**
Fall 2022 Reading Score	.65	<.001***
IEP Status	-.02	.691
Gender	-.05	.368

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female.

Table C4. Winter 2023 NWEA MAP® Reading scores by receiving RGR instruction in Year 2 for students who first had RGR instruction in Grade 1 ($n = 350$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 2 RGR Instruction	.18	.001**
Fall 2022 Reading Score	.65	<.001***
IEP Status	-.02	.691
Gender	-.05	.368

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female.

Partial Correlation Analyses for Grade 2 Students (in Year 1)

Table C5. Spring 2022 NWEA MAP® Reading scores by receiving RGR instruction in Year 1 for Grade 2 students ($n = 347$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 1 RGR Instruction	.12	.022*
Fall 2022 Reading Score	.73	<.001***
IEP Status	-.16	.003**
Gender	-.02	.741

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. . IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Table C6. Fall 2022 NWEA MAP® Reading scores by receiving RGR instruction in Year 1 for Grade 2 students ($n = 374$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 1 RGR Instruction	.19	<.001***
Fall 2022 Reading Score	.73	<.001***
IEP Status	-.18	<.001***
Gender	-.01	.937

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. . IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Table C7. Winter 2023 NWEA MAP® Reading scores by receiving RGR instruction in Year 1 for Grade 2 students ($n = 351$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 1 RGR Instruction	.08	.165
Fall 2022 Reading Score	.66	<.001***
IEP Status	-.31	<.001***
Gender	.03	.572

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female.

Table C8. Winter 2023 NWEA MAP® Reading scores by receiving RGR instruction in Year 2 for students who first used RGR in Grade 2 ($n = 350$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 2 RGR Instruction	.08	.156
Fall 2022 Reading Score	.66	<.001***
IEP Status	-.30	<.001***
Gender	.03	.590

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Partial Correlation Analyses for Grade 3 Students (in Year 1)

Table C9. Spring 2022 NWEA MAP® Reading scores by students who received RGR instruction in Year 1 for Grade 3 students ($n = 328$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 1 RGR Instruction	.05	.411
Fall 2022 Reading Score	.80	<.001***
IEP Status	-.19	<.001***
Gender	.06	.257

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. . IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Table C10. Fall 2022 NWEA MAP® Reading scores by students who received RGR instruction in Year 1 for Grade 3 students ($n = 347$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 1 RGR Instruction	.01	.840
Fall 2022 Reading Score	.79	<.001***
IEP Status	-.20	<.001***
Gender	.03	.630

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. . IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Table C11. Winter 2023 NWEA MAP® Reading scores by students who received RGR instruction in Year 1 for Grade 3 students ($n = 327$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 1 RGR Instruction	-.11	.052 ⁺
Fall 2022 Reading Score	.78	<.001 ^{***}
IEP Status	-.29	<.001 ^{***}
Gender	.02	.717

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. . IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Table C12. Winter 2023 NWEA MAP® Reading scores by students who received RGR instruction in Year 2 for students who first received RGR instruction in Grade 3 ($n = 313$; covariates: fall 2022 NWEA MAP® Reading score (baseline), gender, and IEP status)

Variable	Partial Correlation Coefficient	p -value
Year 2 RGR Instruction	-.11	.045 [*]
Fall 2022 Reading Score	.77	<.001 ^{***}
IEP Status	-.28	<.001 ^{***}
Gender	.03	.558

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. IEP Status was coded as 0 = No, 1 = Yes, Gender was coded as 0 = Male, 1 = Female. . IEP Status was negatively associated with reading scores, which was expected based on prior literature (Woods et al., 2023).

Differences in Literacy Outcomes Among Grade 1 – 3 Students who received Really Great Reading Instruction and Students Who Did Not in Year 1

Table C13. NWEA MAP® Reading scores by RGR Usage Group (covariates: fall 2022 NWEA MAP® Reading Growth score (baseline), gender, and IEP status)

Group Comparisons	Outcome Timepoint	Outcome Grade	Coefficient	Standard Error	t-value	p> t	Hedge's g	Cohen's d
Grade 1 Usage (Year 1)	Spring 2022	1	3.98	1.02	3.89	<.001***	0.29	0.30
Grade 1 Usage (Year 1)	Fall 2022	2	5.01	1.69	2.97	.003**	0.32	0.33
Grade 1 Usage (Year 1)	Winter 2022	2	4.70	1.62	2.90	.004**	0.30	0.30
Grade 2 Usage (Year 1)	Spring 2022	2	2.80	1.45	1.94	.053+	0.19	0.18
Grade 2 Usage (Year 1)	Fall 2022	3	4.62	1.22	3.79	<.001***	0.37	0.36
Grade 2 Usage (Year 1)	Winter 2022	3	2.02	1.64	1.23	.219	0.24	0.24
Grade 3 Usage (Year 1)	Spring 2022	3	1.27	1.23	1.04	.301	0.08	0.08
Grade 3 Usage (Year 1)	Fall 2022	4	0.15	1.07	0.14	.886	0.01	0.01
Grade 3 Usage (Year 1)	Winter 2022	4	-2.14	1.15	-1.85	.065	-0.13	-0.13

+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001