



Raz-Plus ESSA Tier 2 Efficacy Study

Raz-Plus Usage in First Grade Associated
With Accelerated Growth

Technical Research Report

Megan Conrad, Ph.D. and Parnika Bhatia, Ph.D. | ExploreK12 | July 2026



Table of Contents

Executive Summary	3
Introduction	4
Study Design and Methods.....	5
Analytic Approach	6
Results	7
Discussion	10
References.....	11
APPENDIX A: Analytic Approach.....	12

Executive Summary

Problem Statement: Developing foundational literacy remains a persistent challenge, with only 31% of fourth-grade students meeting proficiency standards in 2024. Early reading deficits often predict long-term academic struggles. Teachers report difficulty finding adequate supplemental materials to manage the wide skill disparities — ranging from decoding to independent reading — found in early elementary classrooms.

The Solution and Study Design: This study evaluated the efficacy of the Raz-Plus® K-5 supplemental literacy solution for accelerating student reading growth. The study used a quasi-experimental design (QED) to examine outcomes for 4,602 first-grade students over a 16-week period (fall 2025 to winter 2025). Reading growth in classrooms with high, consistent implementation of Raz-Plus (n=1,870) was compared against a baseline-equivalent comparison group with minimal usage (n=2,732).

Key Impact Findings:

- **Accelerated Growth:** Raz-Plus treatment classrooms achieved an adjusted mean growth of 8.75 RIT points, compared to 8.07 RIT points in the comparison group, representing a 1.6-week instructional advantage.
- **Outpacing National Norms:** Implementation of Raz-Plus helped students outpace typical national growth expectations by 24.3%.
- **Support for At-Risk Students:** Students entering the year in the lowest 25th percentile showed a 2.6-week instructional advantage, confirming the platform's efficacy in supporting the most vulnerable learners.

Conclusion: For district leaders and educators, these findings validate Raz-Plus as an effective tool for bridging literacy gaps and supporting educational equity. By providing flexible, research-backed instructional resources, the solution allows teachers to move beyond "one-size-fits-all" instruction, delivering the individualized differentiation necessary to ensure students across all proficiency levels achieve accelerated reading growth.

Introduction

Developing foundational literacy and reading comprehension remains a persistent challenge for elementary educators. Over the past 30 years, reading proficiency in the United States has remained low and largely unchanged; in 2024, only 31% of fourth-grade students met proficiency standards, representing only a marginal increase from 1992 (NCES, 2024). The problem is even more acute for teachers of non-white and economically disadvantaged students, as well as children with disabilities, who are two to three times more likely to fall below grade-level standards (NCES, 2017).

Research has shown that reading deficits in early elementary school often predict persistent challenges, highlighting the crucial need for early screening and targeted support. A recent longitudinal study of over 400,000 students found that only 1 in 50 students scoring in the bottom 20% in the spring of first grade reached reading proficiency by the end of third grade (Kuhfeld, Lewis, & Morton, 2026). One cause of this pattern is differences in motivation and engagement in reading. Without support, early deficits can lead to frustration, disengagement, and avoidance of reading, causing these students to lag progressively further behind their peers. Research has found that students with lower reading skills in early elementary school are more likely to exhibit lower motivation and self-concept as readers than their peers with typical reading skills. This results in different levels of reading experience and exposure, which further contribute to differences in achievement: **striving readers are less motivated to read, so they read less, and as a result, they demonstrate smaller levels of growth than their peers, leading to further declining levels of reading motivation and self-efficacy** (Vaknin-Nusbaum, Nevo, Brande, Gambrell, 2018).

To mitigate these risks and prevent early gaps from becoming permanent deficits, teachers must be equipped with engaging, accessible, High-Quality Instructional Materials (HQIM) rooted in science-based reading instruction. However, a standardized core curriculum alone is rarely sufficient to manage the large skill disparities typical of early elementary classrooms, where proficiency can range from mastering the alphabet to reading chapter books. A recent survey found that fewer than 25% of early elementary teachers are highly satisfied with their available literacy curricula (The Harris Poll, 2025). In order to support all students, teachers require flexible, high-quality supplemental resources that they can use to deliver targeted, individualized differentiation.

Raz-Plus addresses this gap as a comprehensive K-5 supplemental literacy solution aligned to Science of Reading (SoR) principles of literacy instruction. Raz-Plus resources integrate into existing frameworks to both complement broad core instruction and supplement the curriculum by filling specific foundational gaps within a digital platform that kids actually enjoy. It equips educators with:

- **Decodable texts and targeted vocabulary tools** to reinforce early decoding skills and meaning-making.
- **Engaging books and explicit lessons** tied to research-backed strategies that deepen comprehension and fluency.
- **A motivating independent reading portal** that provides personalized reading practice appropriate to each student's Lexile® Measure.
- **Built-in professional development** and ongoing webinars to provide the structural support teachers need to execute differentiated instruction effectively.

The current study provides a rigorous evaluation of Raz-Plus's real-world impact on students' reading growth and achievement in the early elementary grades. Guided by Every Student Succeeds Act (ESSA) Tier 2 (Moderate) evidence standards, this quasi-experimental design (QED) study compared first-grade literacy growth over a 16-week period for 1,870 students in classrooms with intentional Raz-Plus usage,

compared with a baseline-equivalent comparison group of 2,732 students with no Raz-Plus usage. Because early intervention and support are most vital for those most at risk of falling behind, this analysis places an additional focus on the platform's efficacy for striving learners. Specifically, the study explores whether the use of Raz-Plus accelerates first-grade reading proficiency compared to similar peers with minimal or no Raz-Plus usage, and investigates the magnitude of these benefits for the most vulnerable learners.

Study Design and Methods

Study Design

This study used a quasi-experimental design (QED) to examine the relationship between Raz-Plus use and student literacy growth. Treatment and control classrooms were selected from a full data set based on the availability of beginning-of-year (fall 2025) and middle-of-year (winter 2025) reading proficiency scores and classroom-level Raz-Plus implementation data.

Usage and Implementation

Raz-Plus provides a range of differentiated instructional resources, including decodable texts, vocabulary development tools, and systematically designed lessons. Typical usage includes accessing a digital library of grade-appropriate texts that feature read-aloud supports for early readers, completing associated comprehension quizzes, and engaging in self-paced practice activities that reinforce foundational literacy skills and build independent reading volume.

Treatment classrooms were selected based on evidence of high and consistent implementation of Raz-Plus. To ensure a focus on active, teacher-driven engagement rather than optional or inconsistent access, we restricted the treatment group to classrooms where 90% or more of the students demonstrated frequent usage throughout the study period. The comparison group comprised classrooms where usage was minimal across the majority (90%+) of the students. Implementation within the treatment group was teacher-directed and guided by local instructional needs.

Assessing Reading Skills and Growth

The outcome measure of student literacy achievement and growth in the current study is the NWEA MAP® Growth™ Reading assessment. MAP Growth assessments are adaptive, interim assessments used to track student achievement and growth before end-of-year summative tests, to inform instructional planning, and to evaluate the efficacy of interventions. The Reading assessment for emerging readers (grades K-2) is designed to evaluate early literacy skills, including phonological awareness, print concepts, and word recognition. Tests are delivered online and typically administered at three time points in the school year: fall (around week 4), winter (around week 20), and spring (around week 32).

Performance is presented as **RIT scores** (Rasch Unit), an equal-interval scale independent of grade level, and percentile ranking based on comparison to nationwide performance for that grade and subject area. Growth is calculated as RIT score difference between tests. Since exact test dates are set by individual schools and districts, and significant variation was observed, we calculated **fall-to-winter growth** using a standardized 16-week growth score for each student:

$$\text{Standardized Growth} = \left(\frac{\text{Winter RIT} - \text{Fall RIT}}{\text{Instructional Weeks}} \right) \times 16 \text{ weeks}$$

Study Sample

The analytic sample consisted of 4,602 first-grade students, with 1,870 students in the Raz-Plus treatment group and 2,732 students in the comparison group. Within the treatment group, students averaged 55.7 days of usage (SD = 50.4) and completed an average of 160 activities (SD = 134) from fall to winter. Each week, these students typically completed an average of 10 activities, including listening to books, reading, comprehension quizzes, and foundational skills activities. Conversely, usage within the comparison group was confirmed to be low, averaging just 1.26 days of usage (SD = 3.33) and 0.7 completed activities (SD = 3.17) from fall to winter. Across both groups, students' average baseline percentile was around 56%, confirming that the sample represented the average norm-referenced performance of their national peer group. Student-level performance on the beginning-of-the-year reading assessment was used to establish baseline equivalence between the Raz-Plus and comparison groups. Because the baseline differences for both the fall RIT score ($g = 0.0715$) and the fall percentile ($g = 0.0797$) fell within the 0.05 to 0.25 standard deviation window, the sample met WWC baseline equivalence standards under the condition that the baseline fall reading scores were statistically adjusted for in the primary analysis. Consequently, the primary Hierarchical Linear Model (HLM) satisfied this requirement by incorporating the student-level fall RIT score as a grand-mean-centered covariate to isolate the true treatment effect.

Table 2. Raz-Plus Usage and Baseline Performance by Treatment Group

	Treatment Students (n = 1870)	Comparison Students (n = 2732)
	M (SD)	M (SD)
Days Usage of Raz-Plus	55.7(50.4)	1.26(3.33)
Number of Activities Completed	160(134)	0.7(3.17)
Baseline (Fall) RIT Score	157(11.9))	156(12.1)
Baseline Percentile	56.3 (27.9)	55.4 (28.6)

Analytic Approach

Primary Analysis

To examine if Raz-Plus implementation significantly improved student reading outcomes, we conducted a three-level Hierarchical Linear Model (HLM) that accounted for the nesting of students within classrooms and schools. Initial variance testing justified this approach, revealing that 6.8% of the variance in student growth was due to classroom- and school-level clustering. The analytic sample consisted of 4,602 students across 267 unique classrooms nested within 123 schools. Student-level fall-to-winter RIT growth was the outcome of interest. Because the testing window differed across usage groups, the treatment effect was standardized post hoc to a 16-week window. This corresponds to NWEA's default fall-to-winter testing interval and serves as the basis for national growth norms.

Subgroup Analysis

To determine if the efficacy of Raz-Plus implementation varied across student populations, we conducted subgroup analyses based on initial academic risk. When stratified by fall NWEA-MAP percentiles, the sample was distributed across three distinct groups. The core instructional tier, Tier 1 Instruction (25th–90th Percentile), comprised the largest segment of the population with 3,151 students ($n = 1,843$ in the comparison implementation group; $n = 1,308$ in the Raz-Plus usage group). The intervention group, containing Lower-Performing Students (Below the 25th Percentile), consisted of 917 students ($n = 573$ No usage; $n = 344$ Raz-Plus Usage). Finally, the advanced tier, representing Gifted and Advanced Students (Above the 90th Percentile), included 534 students ($n = 316$ No Usage; $n = 218$ Raz-Plus Usage). Subgroup differences across the implementation groups were evaluated using fixed-effects analysis of variance (ANCOVA), mapped to raw gains, and converted to standardized 16-week growth equivalents and relative weeks-of-learning advantages.

Results

Raz-Plus Impact on Growth (Primary Analysis)

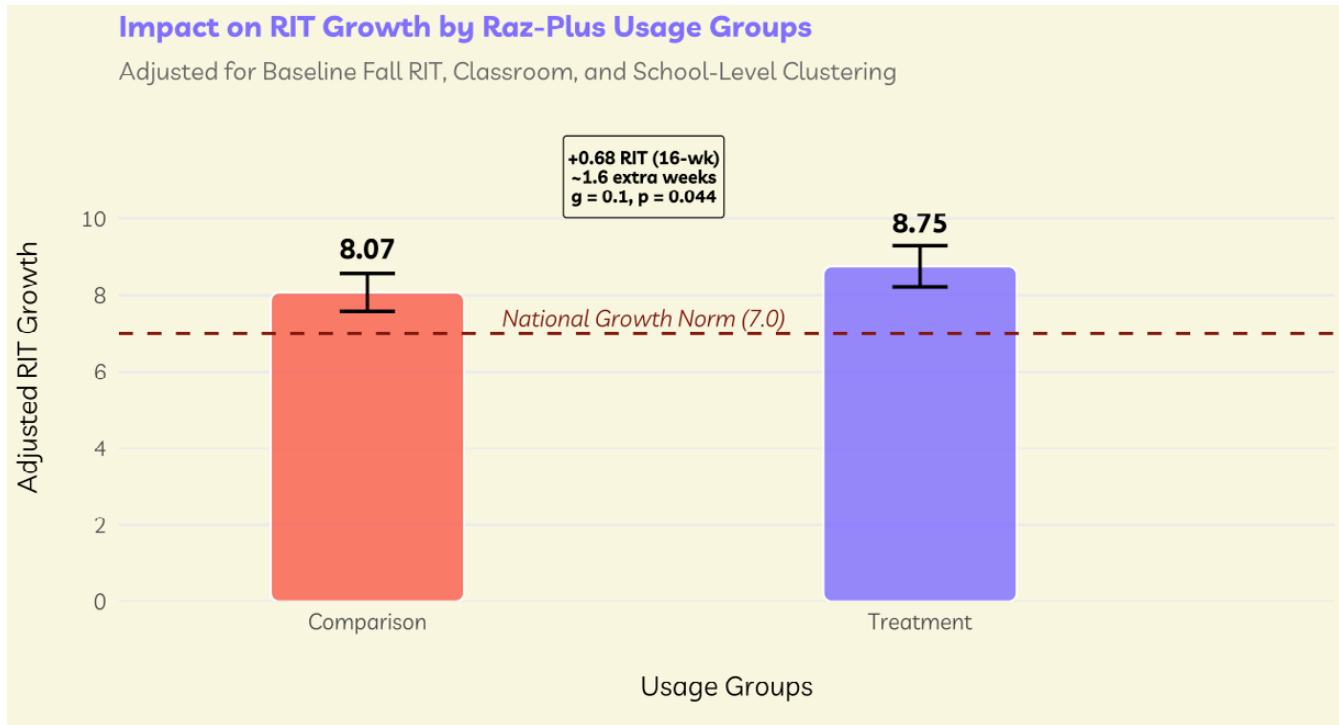
After accounting for school- and classroom-level clustering and controlling for baseline fall RIT scores, Raz-Plus usage had a statistically significant impact on reading growth. Students in the treatment classrooms achieved an adjusted mean of 8.75 RIT points of growth over a standardized 16-week window, compared to 8.07 RIT points in the comparison group — a difference of 0.68 RIT points in favor of the treatment group ($\beta = 0.68$, $SE = 0.336$, $t = 2.03$, $p = 0.044$, Hedges' $g = 0.097$). ***These results demonstrate that consistent engagement with the Raz-Plus solution reliably accelerates reading growth, even when accounting for the powerful role that nested school and classroom environments play in student outcomes.***

The +0.68 RIT growth advantage seen in the treatment group over the comparison group represents an instructional gain comparable to 1.6 extra weeks of reading instruction, based on a standardized 16-week testing window. Furthermore, given that the national NWEA MAP norm for first-grade progress during this fall-to-winter interval is 7.0 RIT points, treatment classrooms outpaced typical national growth expectations by 24.3%, adding acceleration to the baseline trajectory.

Table 3. Adjusted 16-Week Impact Results by Treatment Group

Treatment		Comparison		Results		
Adjusted Mean	Standard Deviation (SD)	Adjusted Mean	Standard Deviation (SD)	Difference	Standard Error (SE)	Effect Size (Hedges' g)
8.75	7.07	8.07	6.97	0.68	0.336	0.1

Figure 1: Adjusted 16-Week Impact Results by Treatment Group



Subgroup Analysis

To determine if the efficacy of Raz-Plus implementation varied across different student populations, we conducted subgroup analyses based on initial academic risk using an Analysis of Covariance (ANCOVA) that controlled for baseline fall RIT scores. The positive impacts of high-fidelity implementation were noted across Tier 1, 2, and 3 intervention contexts.

For students entering the academic year with the **highest initial risk (below 25th percentile)**, Raz-Plus Usage yielded an adjusted mean growth of 11.1 RIT points, compared to 9.97 RIT points in the comparison group. This net adjusted difference of 1.12 RIT points approached significance thresholds ($F = 3.54$, $p = 0.06$). On a standardized 16-week window, this translates to a substantial **2.6-week instructional learning advantage**.

Students within the 25th-90th percentile tier also benefit from Raz-Plus usage. Raz-Plus usage achieved an adjusted mean growth of 8.63 RIT points, compared to 7.79 RIT points in the low-usage comparison classrooms. This net adjusted advantage of 0.845 RIT points was statistically significant ($F = 12.90$, $p < .001$). This represents an **instructional gain comparable to 1.9 extra weeks of reading instruction over the comparison group**.

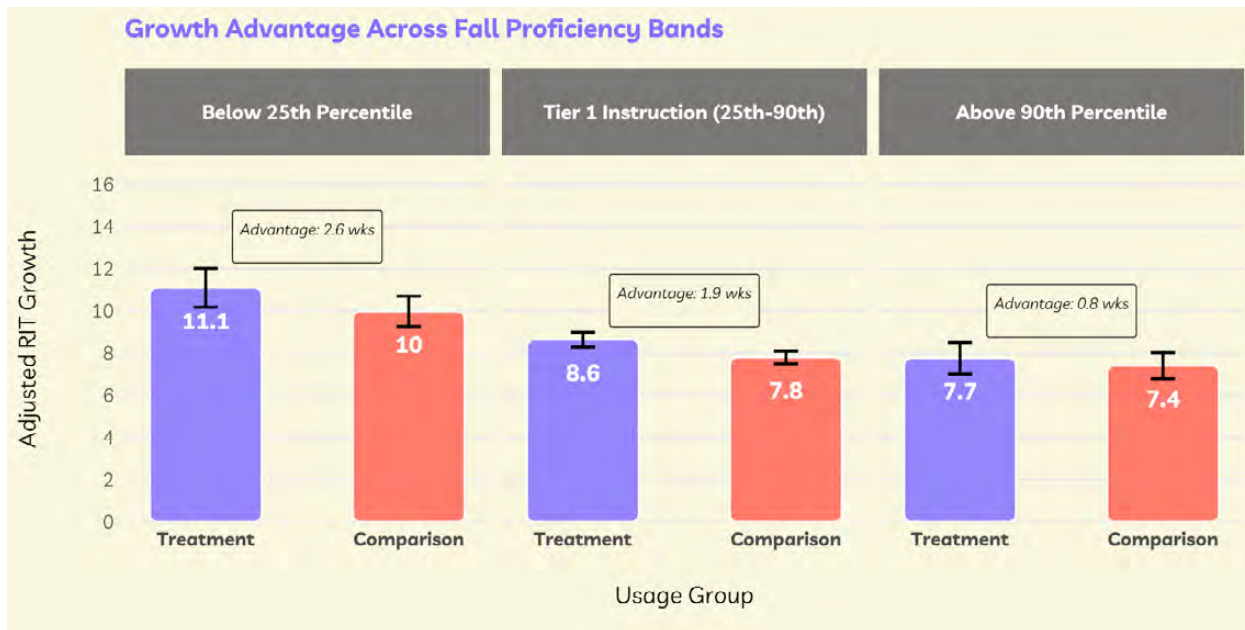
For highly proficient students, the Raz-Plus usage group achieved an adjusted mean growth of 7.75 RIT points, compared to 7.4 RIT points in the comparison group. This adjusted difference of 0.35 RIT points did not reach the threshold for statistical significance ($F = 0.511$, $p = 0.475$).

Combined, these findings indicate that the Raz-Plus solution provided critical, accelerated growth for both core tier and at-risk student populations.

Table 4. Subgroup Analysis (16 week): Impact Results by Treatment Group

	Treatment		Comparison		Results		
Student Baseline Performance	N Students	M (SD)	N Students	M (SD)	Mean Difference (SE)	F-stat	p-value
Below 25th Percentile	344	11.1 (14.2)	573	9.97 (11)	1.12(0.6)	3.54	0.06
25th-90th Percentile	1308	8.63 (10)	1843	7.79 (8.52)	0.84 (0.24)	12.9	<0.001
Above 90th Percentile	218	7.75 (8.77)	316	7.40 (7.28)	0.35 (0.49)	0.51	0.475

Figure 2: Subgroup Analysis for Baseline Performance: Impact Results by Treatment Group (16 week adjusted)



Discussion

This study evaluated the efficacy of the Raz-Plus supplemental literacy solution on first-grade reading growth over a 16-week period, specifically examining its impact on accelerating proficiency for diverse learner populations. The primary analysis revealed that consistent engagement with Raz-Plus led to statistically significant reading growth, with treatment classrooms achieving 8.75 RIT points compared to 8.07 RIT points in the comparison group. **This performance translates to a 1.6-week instructional advantage, outperforming national NWEA MAP growth norms by 24.3%.**

The subgroup analysis underscores the solution's role in promoting educational equity. While Raz-Plus provided measurable benefits for students in the 25th–90th percentile (1.9-week instructional advantage, $p < .001$), the most substantial gains were observed among the most vulnerable learners. **Students entering the year in the lowest 25th percentile of performance realized a 2.6-week instructional advantage over their peers,** demonstrating the solution's capacity to support intensive intervention needs. Although positive trends were observed for advanced learners (above the 90th percentile), these results did not reach statistical significance.

These findings offer critical implications for elementary practitioners. As many educators report dissatisfaction with existing literacy curricula and struggle to manage wide skill disparities — ranging from decoding to independent reading — Raz-Plus can fill gaps in core curricula and increase independent reading by providing access to appropriate and engaging texts with built-in support tools. By providing decodable texts, systematic lessons, and engaging student tools, the platform allows teachers to move beyond a “one-size-fits-all” approach, offering the targeted, individualized differentiation necessary to prevent early literacy deficits from becoming permanent challenges.

References

- The Harris Poll (2025). *Nemours Children's Reading BrightStart! Early Literacy Survey: A National Survey by Nemours Children's Health*. {Unpublished Report] <https://www.nemours.org/content/dam/nemours/nemours-org/en/documents/NCRBS-harris-poll-executive-summary.pdf>
- Kuhfeld, M., Lewis, K., Morton, E. (2026, June). Behind at the starting line: What Kindergarten test scores tell us about reaching proficiency in 3rd grade (Research Brief). NWEA Research. Retrieved from https://www.nwea.org/uploads/Behind-the-starting-line-what-kindergartent-test-scores-tell-us-about-reaching-proficiency-in-third-grade_NWEA_researchBrief.pdf
- National Center for Education Statistics. (2017). *The Nation's Report Card: NAEP Reading 2017 Assessment*. U.S. Department of Education, Institute of Education Sciences. <https://www.nationsreportcard.gov/>
- National Center for Education Statistics. (2024). *The Nation's Report Card: NAEP Reading 2024 Assessment*. U.S. Department of Education, Institute of Education Sciences. <https://www.nationsreportcard.gov/>
- Vaknin-Nusbaum, V., Nevo, E., Brande, S., & Gambrell, L. (2018). Developmental aspects of reading motivation and reading achievement among second grade low achievers and typical readers. *Journal of Research in Reading*, 41(3), 438–454. <https://doi.org/10.1111/1467-9817.12117>

APPENDIX A: Analytic Approach

Baseline Equivalence

To ensure the internal validity of the study, baseline equivalence between the treatment and comparison groups was evaluated on fall RIT scores and fall national percentiles. Under the What Works Clearinghouse (WWC) Standards (Version 5.0), the threshold for establishing baseline equivalence without statistical adjustment is a standardized difference. If the baseline difference falls between 0.05 and 0.25 standard deviations, the groups are considered “Equivalent with Statistical Adjustment,” requiring the primary outcome models to control for the baseline measure as a covariate. Baseline differences exceeding 0.25 fail to meet equivalence standards.

Because the baseline differences for both the fall RIT score ($g = 0.0715$) and the fall percentile ($g = 0.0797$) fell within the 0.05 to 0.25 standard deviation window, the sample met WWC baseline equivalence standards under the condition that the fall RIT was statistically adjusted for in the primary analysis. Consequently, the primary Hierarchical Linear Model (HLM) satisfied the requirement by incorporating the student-level fall RIT score as a mean-centered covariate to isolate the treatment effect.

Table 1. Baseline Equivalence Results, Fall RIT Score and Fall Percentile

Measure	Treatment Mean (SD)	Comparison Mean (SD)	Mean Diff	Pooled SD	Hedges' g	WWC Status
Fall RIT Score	157.0 (11.9)	156.0 (12.1)	0.862	12.1	0.0715	Equivalent with statistical adjustment
Fall Percentile	56.3 (27.9)	54.0 (28.6)	2.260	28.3	0.0797	Equivalent with statistical adjustment

Table 2
Hierarchical Linear Model Predicting First Grade Students' RIT Growth (Students Nested in Classrooms Nested in Schools)

	Null Model	Conditional Model
Fixed Effects — Estimate (SE)		
Intercept	9.31*** (0.22)	27.18*** (1.56)
Fall RIT score	—	−0.12*** (0.01)
Usage Tier: Raz-Plus Usage (vs. Comparison)	—	0.75* (0.37)
Random Effects — Variance		
School (Level 3)	3.43	3.86
Classroom Within School (Level 2)	0.69	1.00
Student / Residual (Level 1)	56.40	54.40
Intraclass Correlations (Null Model)		
School ICC	0.057	—
Classroom ICC	0.011	—
Total Clustering	0.068	—
Model Information		
REML Criterion	31,786.7	—
Students (Level 1)	4,602	4,602
Classrooms (Level 2)	267	267
Schools (Level 3)	123	123

Note. Random effects are reported as variances. ICCs are derived from the null model. The contrast (Raz-Plus Usage vs. No usage) is statistically significant, $t(244) = 2.03$, $p = .044$, 95% CI [0.02, 1.48].

* $p < .05$. *** $p < .001$.