Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 24-1001

Effectiveness of Structured Teacher Adaptations to an Online Content Literacy Intervention for Third Graders: A Randomized Controlled Trial During COVID-19

Jackie E. Relyea
North Carolina State University

Joshua B. Gilbert
Harvard University

Mary Burkhauser
Harvard University

Ethan Scherer
Harvard University

Douglas Mosher
Harvard University

Zhongyu Wei
Harvard University

Johanna Tvedt
Harvard University

James S. Kim
Harvard University

Scaling up evidence-based educational interventions to improve student outcomes presents challenges, especially when adapting to new contexts while maintaining fidelity. Structured teacher adaptations that integrate the strengths of experimental science (high fidelity) and improvement science (high adaptation) offer a viable solution to bridge the research-practice divide. This preregistered study was designed to examine the effectiveness of structured teacher adaptations to a Tier 1 content literacy intervention on Grade 3 students' (N = 1,914) engagement in asynchronous digital app and print-based reading activities, synchronous student-teacher interactions during the Zoom-delivered synchronous lessons, and student literacy outcomes. Using a multisite, cluster randomized trial design, a total of 95 teachers and their students in 26 K-5 elementary schools were randomly assigned to either a core treatment that emphasized fidelity of implementation or a structured teacher adaptations condition (i.e., Adaptive Treatment). In the structured teacher adaptations condition, teachers participated in Team-Based Learning activities that tightly coupled knowledge acquisition and application focused on improving student engagement during the COVID-19 pandemic when instruction was fully remote. Students in the adaptations condition outperformed students in the core treatment condition on measures of science reading comprehension (ES = .07) and science background knowledge (ES = .09). Implementation analyses indicated that student engagement with digital app activities and synchronous lessons was also higher in the Adaptive Treatment. Findings suggest that structured adaptations represent an asset-based approach that empowers teachers with both research knowledge and team-based practical knowledge to enhance student engagement and literacy outcomes.

VERSION: January 2025

Suggested citation: Relyea, Jackie E., Joshua B. Gilbert, Mary A. Burkhauser, Ethan Scherer, Douglas M. Mosher, Zhongyu Wei, Johanna N. Tvedt, and James S. Kim. (2025). Effectiveness of Structured Teacher Adaptations to an Online Content Literacy Intervention for Third Graders: A Randomized Controlled Trial During COVID-19. (EdWorkingPaper: 24 -1001). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/ztcp-v322

Abstract

Scaling up evidence-based educational interventions to improve student outcomes presents challenges, especially when adapting to new contexts while maintaining fidelity. Structured teacher adaptations that integrate the strengths of experimental science (high fidelity) and improvement science (high adaptation) offer a viable solution to bridge the research-practice divide. This preregistered study was designed to examine the effectiveness of structured teacher adaptations to a Tier 1 content literacy intervention on Grade 3 students' (N = 1,914) engagement in asynchronous digital app and print-based reading activities, synchronous student-teacher interactions during the Zoom-delivered synchronous lessons, and student literacy outcomes. Using a multisite, cluster randomized trial design, a total of 95 teachers and their students in 26 K-5 elementary schools were randomly assigned to either a core treatment that emphasized fidelity of implementation or a structured teacher adaptations condition (i.e., Adaptive Treatment). In the structured teacher adaptations condition, teachers participated in Team-Based Learning activities that tightly coupled knowledge acquisition and application focused on improving student engagement during the COVID-19 pandemic when instruction was fully remote. Students in the adaptations condition outperformed students in the core treatment condition on measures of science reading comprehension (ES = .07) and science background knowledge (ES = .09). Implementation analyses indicated that student engagement with digital app activities and synchronous lessons was also higher in the Adaptive Treatment. Findings suggest that structured adaptations represent an asset-based approach that empowers teachers with both research knowledge and team-based practical knowledge to enhance student engagement and literacy outcomes.

Keywords: structured adaptations, team-based learning, content literacy instruction, hybrid instruction, synchronous learning, asynchronous learning

Structured Teacher Adaptations to an Online Content Literacy Intervention: Asset-Based Approaches to Bridging the Science of Reading Research and Practice

Scaling evidence-based interventions to diverse educational settings remains one of the most persistent and complex challenges in literacy education research (Dede, 2006; McDonald et al., 2006; Stein et al., 2008). While decades of rigorous studies have demonstrated the efficacy of interventions in controlled conditions, their implementation in real-world classrooms often produces inconsistent outcomes (Author, 2019). This discrepancy, commonly referred to as the research-practice gap, arises from a fundamental misalignment between the controlled conditions under which evidence-based interventions are designed and the variability of the contexts in which they are implemented (K.E. Joyce & Cartwright, 2020). This challenge raises the question of how researchers and educators can achieve a delicate balance: preserving the core instructional principles of interventions with fidelity while adapting to the diverse and dynamic demands of real-world classrooms.

Historically, fidelity has been prioritized as the primary mechanism of effective scaling, maintaining the causal validity of interventions through strict adherence to protocols (Dane & Schneider, 1998; Dusenbery et al., 2003; Kaderavek & Justice, 2010). However, fidelity-focused approaches assume uniform, idealized teaching conditions and reduce teachers' roles to passive implementers of prepackaged interventions, limiting their opportunities to exercise professional judgment or tailor practices to local contexts (Authors, 2017b; Cobb et al., 2003). In contrast, adaptation offers a flexible approach that enables teachers to modify interventions to their specific classroom needs (Datnow & Castellano, 2000; Parsons et al., 2018). Research in improvement science also emphasizes the importance of adaptation for fostering teacher ownership and engagement, both of which are critical for sustainability (Lewis, 2015). Yet, allowing teachers to freely adapt an intervention risks deviations from its theoretical foundations, potentially undermining the effectiveness of evidence-based practices. This tension between fidelity and adaptation has long been framed as a trade-off, but emerging integrative frameworks suggest that the two priorities can coexist (Authors, 2017a).

Structured adaptations reconcile the competing priorities of fidelity and adaptation by providing teachers with explicit guidelines for context-sensitive modifications that preserve core instructional principles. Structured adaptations are guided by researchers and instructional leaders who establish clear parameters and offer support, enabling teachers to design modifications that maintain implementation fidelity, foster student and teacher engagement, and improve student outcomes (Authors, 2024; Neugebauer et al., 2023; Sanetti & Kratochwill, 2009). These structured adaptations empower teachers to incorporate their professional expertise and cultural insights into evidence-based practices, maintaining both the rigor and relevance of the interventions (Bryk et al., 2015; Durlak & DuPre, 2008). The COVID-19 pandemic further illuminated the urgent need for such an approach. As educators navigated disparities in resources, home learning environments, and student engagement, the importance of adaptable models became evident. Teachers' adaptations to sustain meaningful instruction underscore the critical role of structured adaptations that balance adherence to evidence-based curricula with responsiveness to contextual demands. Asset-based pedagogy reframes teacher adaptations as opportunities to leverage the cultural and linguistic resources of students and teachers, positioning them central to effective instruction rather than deviations from prescribed norms (Gabriel & López, 2024; Hattan & Kendeou, 2024).

Building on this framework, this study examines the impact of structured teacher adaptations to a content literacy intervention for third-grade students during the 2020-2021 school year. To accommodate pandemic-related constraints, the original randomized controlled trial (RCT) design was modified to eliminate the traditional control group, such that all participating students received the core treatment. Replicating the procedures and content from prior studies (e.g., Authors, 2017a, 2017b), 95 classrooms were randomly assigned to one of two conditions: in the *Core Treatment* condition, teachers implemented the content literacy intervention, called [BLINDED FOR REVIEW], while in the *Adaptive Treatment* condition, teachers implemented the Core Treatment components with support to make structured adaptations tailored to their classroom contexts. Conceptually replicating prior research (Authors, 2017a) in a fully digital context, this study explores how structured adaptations influenced student engagement and achievement outcomes, offering critical insights for scaling interventions in diverse educational settings.

Balancing Fidelity and Adaptation in Interventions

Efforts to scale evidence-based educational interventions have long grappled with a paradox: achieving both fidelity to core principles and adaptability to diverse educational contexts. These two dimensions of program implementation place specific demands on teachers, often highlighting tensions between the theoretical rigor of research-based practices and the practical realities of classrooms. Fidelity, as a cornerstone of the experimental science paradigm, emphasizes teachers' strict adherence to the core components of a program to ensure its theoretical and empirical integrity (Dane & Schneider, 1998). This approach establishes consistency across implementations, allowing researchers to attribute observed outcomes directly to the intervention while ensuring replicable results that form the basis for scaling (Domitrovich et al., 2010). Fidelity-focused implementation models prioritize control over variability, reflecting a tightly managed instructional framework in which teachers' roles are prescribed, and deviations are often seen as compromising program integrity (Rowan, 1990; Sherin & Drake, 2009).

While fidelity offers clear advantages in controlled settings, emphasizing uniformity often poses challenges for the practical scalability of interventions. Classrooms are inherently diverse and often vary in resources, cultural norms, linguistic diversity, and teacher expertise. Programs designed for one context may not seamlessly translate to others, particularly when teachers are expected to implement them without accounting for the unique needs of their students or the constraints of their environment (Coburn, 2003). These challenges are amplified in under-resourced or culturally diverse schools, where rigid adherence to intervention protocols can inadvertently limit relevance and efficacy (Goldenberg & Gallimore, 1991; Stanovich, 2003). While scientifically robust, fidelity-focused models risk narrowing opportunities for teacher agency and professional judgment, ultimately reducing the applicability of interventions in real-world contexts.

Adaptation, on the other hand, provides a compelling counterbalance, emphasizing flexibility and responsiveness to the dynamic nature of educational settings. Improvement science frameworks highlight the importance of tailoring interventions to local contexts while preserving their foundational principles (Bryk et al., 2015; Gutiérrez & Penuel, 2014). This approach views variability not as a threat to program integrity but as an opportunity to enhance its practical relevance (Coburn, 2003). Teachers play a crucial role in this process, drawing on their contextual knowledge and professional expertise to modify implementation strategies in ways that align with the cultural, linguistic, and experiential realities of their classrooms (McDonald et al., 2006). For instance, adaptation may involve adjusting pacing to accommodate diverse learner needs, integrating culturally relevant materials, or leveraging students' linguistic assets to promote engagement and comprehension (Goldenberg & Gallimore, 1991).

The perceived dichotomy between fidelity and adaptation has often framed these dimensions as competing priorities: fidelity ensures program consistency, while adaptation promotes contextual relevance. This framing creates a paradox—fidelity's emphasis on control may appear to restrict flexibility, while adaptation's focus on variability might seem to undermine standardization. However, emerging evidence highlights the potential for fidelity and adaptation to function as complementary forces. Fidelity preserves the theoretical and empirical foundations of interventions, keeping core principles intact, while teachers' adaptations enhance the applicability of interventions by aligning them with the realities of diverse classrooms. Interventions may achieve consistency and flexibility for equitable and effective outcomes across varied educational contexts by delineating "non-negotiable" components while empowering teachers to make context-sensitive adaptations that maintain instructional integrity. Recent research supports this integrated implementation approach, showing that "structured" adaptations can not only maintain fidelity but also improve teacher engagement and student outcomes by fostering a balance between rigor and responsiveness (e.g., Authors, 2017a, 2017b; M. Vaughn et al., 2020). When viewed as synergistic rather than competing, fidelity and adaptation create pathways for scalable, sustainable, and contextually meaningful interventions that address the dual demands of scientific rigor and practical relevance.

The Promise of Structured Teacher Adaptations

Structured teacher adaptations provide a practical framework for balancing fidelity and adaptation, potentially addressing the complex challenge of scaling evidence-based interventions across diverse educational settings. Rather than relying on modifications without clear guidance or ad hoc changes, structured adaptations establish explicit parameters that define essential components and identify areas where flexibility can address specific contextual needs (Authors, 2023; Maniates, 2017). Structured adaptations center on collaboration between teachers and researchers, leveraging teachers' professional expertise and contextual knowledge to adapt implementation in ways that preserve the intervention's core principles while enhancing its relevance and scalability (Bryk, 2015; Lemons et al., 2014). By doing so, structured adaptation can facilitate the integration of research-based practices into varied contexts without compromising their foundational instructional goals.

Figure 1 illustrates this balance, positioning the concept of structuring teacher adaptations in the high-fidelity, high-adaptation quadrant (Authors, 2023). This illustration integrates the strengths of experimental science, which emphasizes adherence to core principles, and the strengths of improvement science, which prioritizes responsiveness to local contexts. This dual emphasis bridges the research-practice divide, positioning structured adaptations as a mechanism for achieving scientific rigor and contextual relevance.

A critical feature of structured adaptations is their collaborative development, involving both researchers and teachers in a joint process to design and refine the framework. This partnership respects the theoretical foundations of the intervention and accounts for the practical demands of varied educational settings (Brownell et al., 2006). Collaboration involves initial training, ongoing support, and structured feedback mechanisms, creating professional learning environments that promote inquiry, peer learning, and collective problem-solving (Elmore, 1996; Frank et al., 2011). Embedded professional development models, such as team-based learning (TBL; Michaelsen & Sweet, 2008) and peer coaching (B. Joyce & Showers, 1980), further enhance teachers' ability to align instructional practices with the intervention's core principles while addressing classroom-specific challenges (Authors, 2017a; Dobb et al., 2017). These processes foster teacher agency and ownership, potentially making interventions more sustainable and impactful (Pugach & Johnson, 2002).

The necessity for an integrative implementation approach became particularly evident during the COVID-19 pandemic when educators navigated unprecedented disruptions. Remote learning environments magnified the need for flexible and context-sensitive approaches to sustain instruction and student engagement. Traditional fidelity-focused models often fall short in accommodating the realities of students and teachers (Authors, 2017b). Structured adaptations, however, provided a pathway to uphold instructional goals by accommodating changes in delivery methods and supporting students through online platforms. This flexibility allowed interventions to remain scientifically grounded and contextually relevant to the needs of students and teachers. While empirical research on structured adaptations is still evolving, existing studies demonstrate the effectiveness of this approach, highlighting that it can simultaneously maintain fidelity and improve student outcomes (e.g., Lemons et al., 2014; Neuman et al., 2021). For example, in the previous study (Authors, 2017a), students in the structured adaptations condition achieved 0.12 standard deviations (SD) higher on standardized reading comprehension assessments compared to those in the core condition.

The current study builds on this evidence by comparing a fidelity-focused core condition of the content literacy intervention to a structured adaptations condition implemented during the pandemic. Structured adaptations positioned teachers as co-creators of knowledge in tailoring evidence-based practices to align with the contextual demands of their classrooms. By fostering collaboration and supporting modifications, structured adaptations aimed to enhance both instructional relevance and scalability while maintaining instructional coherence (Dede, 2006).

Team-Based Learning in Support of Structured Teacher Adaptations

One effective approach to building teachers' capacity for identifying structured adaptations is TBL. It is an evidence-based professional development framework that emphasizes collaboration, shared accountability, and adaptive problem-solving among educators to enhance instructional practices and innovation (Michaelsen & Sweet, 2008). Unlike traditional, top-down professional development approaches, TBL facilitates structured collaboration and peer-assisted learning strategies, enabling teachers to share expertise, co-construct knowledge, and design instructional practices tailored to their unique classroom contexts (Lemons et al., 2014). This participatory model aligns with the principles of asset-based pedagogy (López, 2024), valuing teachers' professional agency by positioning them as active contributors to the educational ecosystem. It empowers educators to integrate their cultural and contextual knowledge into instructional design, fostering a more responsive and inclusive approach to teaching (Authors 2023; Martinez et al., 2024).

Empirical evidence supports TBL as an asset-based approach to bridging the science of reading research and practice. TBL has demonstrated strong efficacy across diverse professional domains, including STEM, medicine, humanities, and the social sciences (e.g., Leupen, 2020; Michaelsen & Sweet, 2003; Roosien et al., 2023). As a pedagogy, by moving away from deficit-oriented knowledge transmission models, TBL focuses on active, team-based knowledge construction. It is built around four essential elements: (a) forming groups, (b) fostering accountability within teams, (c) providing feedback, and (d) designing assignments that promote individual learning and team development. Two critical components of TBL are the individual readiness assurance test (iRAT) and the team readiness assurance test (tRAT), which encourages participants to develop a foundational understanding of key concepts before applying them to collaborative problem-solving. These elements enable meaningful integration of knowledge acquisition and application, creating professional learning environments that emphasize inquiry, peer learning, and collective problem-solving.

The relevance of TBL was especially salient during the global COVID-19 pandemic, which underscored the importance of fostering belongingness, competence, and autonomy among educators facing unprecedented challenges. Self-determination theory posits that the broader sociocultural context can facilitate or undermine motivation and well-being (Ryan & Deci, 2000). More recently, literacy scholars and educational psychologists (Gabriel & López, 2024; Hattan & Kendeou, 2024) have argued that self-determination theory can potentially facilitate asset-based models of reading through its emphasis on changing the broader professional learning contexts of teachers and the classroom contexts where teachers interact with their students. Importantly, Ryan and Deci (2020) suggested that self-determination theory "shares with many constructivist and post-modern approaches to education a concern with cultural internalizations and impositions, and a recognition of layered forms of hegemony. It stands as an example of theory that can be both empirically grounded and critical, and thus merits consideration alongside other critical educational theories" (p. 9).

In many ways, TBL represents a novel approach for disrupting power asymmetries (a) between researchers who produce generalizable knowledge and practitioners who value local, contextualized knowledge, and (b) between administrators who have power to make policy decisions and practitioners who have limited voice and agency. In the current discourse around the science of reading, practitioners are often viewed as passive recipients of knowledge who must adhere faithfully to evidence-based practices and programs, which are selected and purchased by school administrators. Through the implementation of TBL activities, our aim was to translate self-determination theory into a model for professional learning. Ultimately, our aim was to create professional learning contexts that supported teachers' knowledge of research-based practices, provided teachers with autonomy to make structured adaptations informed by research and local contexts, and fostered teachers' connections during a global pandemic. The pivot to online TBL was also informed by emerging research on the flexibility of this learning models. Research also highlights the feasibility of implementing TBL in fully online contexts (Clark et al., 2018), demonstrating its adaptability as a solution for fostering collaborative learning and dialogue during periods of disruption (Roosien et al., 2023).

In the current intervention, TBL served as a collaborative platform for teachers to co-develop structured adaptations with researchers and peers, grounding their decisions in the intervention's theoretical foundations. Teachers participated in iRAT and tRAT activities, designed and facilitated by researchers, to deepen their understanding of the intervention's purpose and components. For example, iRAT questions included: *What problem was the intervention designed to help teachers address? How can the app and trifolds help students?* and *Which of the following words are related to the word "diagnosis"?* During the researcher-led tRAT sessions, teachers clarified key concepts and discussed how the app and trifolds enhance student learning by introducing essential vocabulary and providing repeated exposures to important words. Moreover, these sessions also supported collaborative planning on adapting the intervention to meet specific classroom needs. For instance, teachers co-developed strategies with their peers to foster student engagement with the app and trifolds, including modeling app use, creating scaffolded activities like scavenger hunts and breakout room discussions, and introducing methods to motivate participation. They also discussed barriers such as varying levels of home support by planning structured classroom time for the app and trifold activities and maintaining regular communication with parents. Through these sessions, teachers collaboratively refined adaptation plans, set measurable goals, and developed strategies to monitor progress. This process enabled teachers to translate evidence-based practices into practical, context-specific strategies that aligned with the intervention's instructional objectives while respecting the unique needs of their classrooms.

Why Structured Adaptations Matter to Content Literacy Instruction

A growing body of research highlights the effectiveness of content-integrated literacy instruction in improving students' language, literacy, and engagement outcomes (e.g., Authors, 2021, 2024a, 2024b, 2024c; Connor et al., 2017; Guthrie et al., 2004; Hwang et al., 2022). Content literacy instruction emphasizes schema building, a process through which students organize and integrate new information by connecting it to prior knowledge, facilitating deeper understanding and knowledge transfer across contexts (Anderson & Pearson, 1984; Kintsch & van Dijk, 1978). This process is particularly vital when students engage with complex informational texts, which often present unique challenges due to their abstract concepts, dense language, and complex structures, requiring cognitive integration (Martin & Duke, 2010). These cognitive demands could hinder comprehension without targeted instructional support.

Structured teacher adaptations provide a practical approach for addressing these challenges, enabling teachers to tailor instruction in ways that maintain academic rigor while ensuring accessibility and relevance for students. For instance, constructing schemas around the abstract scientific concept of *systems*, such as the human body, requires students to develop mental frameworks that illustrate the interdependence of concepts like *organs* and *functions*. To scaffold this schema-building process, teachers can make modifications to the core curriculum by implementing an extended lesson that offers opportunities for students to practice vocabulary, syntax, and sentence structure in an engaging, low-stakes format (e.g., “mad libs” activity). Additionally, incorporating student presentations about *systems* encourages students to synthesize information and apply their knowledge in new contexts, thus deepening comprehension and facilitating knowledge transfer (Gick & Holyoak, 1983). These extended lesson activities can help students establish connections between ideas, refine their conceptual understanding, and deepen their grasp of abstract scientific concepts.

There is emerging theoretical and empirical evidence that promoting teacher engagement can fuel and fire student engagement. Self-determination theory, with its emphasis on autonomy, competence, and relatedness, has provided a foundation for instructional practices that support student engagement. Such teacher-direct practices include fostering mastery goals, promoting situational interest, and highlighting useful applications of knowledge (Guthrie et al., 2013; Patall, 2018; Reeve et al., 2022, Schraw & Lehman, 2001). Science instruction provides an ideal context to enact these practices because science is not part of high-stakes accountability policies governing teachers or schools in most U.S. schools. Therefore, science instruction offers more opportunities for teachers to experiment with engagement-enhancing practices. We hypothesized that building autonomy, competence, and relatedness among teachers who participated in TBL activities would spill over into classrooms in ways that supported their students’ engagement with both asynchronous and synchronous learning activities. For example, TBL afforded teachers opportunities to adapt Zoom-delivered synchronous lesson by extending them and providing more opportunities for students to discuss concepts related to the human body system, using games designed to foster student interests, build conceptual knowledge, and foster dialogic activities. TBL also provided teachers with structured adaptations designed to enhance student engagement with asynchronous learning activities using the app and trifolds. These tools were designed to promote language, literacy, and content knowledge development by engaging students in vocabulary practice, engagement with science texts, and schema building. However, their asynchronous nature requires intentional scaffolding as the lack of immediate guidance and feedback often affects students' motivation and ability to navigate complex informational content independently (Lucas et al., 2020). Teacher adaptations can increase student engagement by modeling app usage, monitoring student progress with the app, and providing opportunities for students to share their learning from the app during class discussions. Additionally, celebrating students' progress fosters a sense of accomplishment, enhancing motivation and sustaining active participation. These strategies transform the app into an interactive, student-centered resource, making it more accessible while fostering meaningful engagement.

The Current Study

Research indicates that targeted engagement strategies enhance student motivation, build confidence in navigating complex material, and foster deeper connections to content (Darling-Hammond & Cook-Harvey, 2018; Ryan & Deci, 2000). The present study examined the effectiveness of structured teacher adaptations in a content literacy intervention delivered remotely during the COVID-19 school closures. Two treatment conditions were implemented: (a) a core treatment condition (Core Treatment), which replicated implementation procedures from prior experiments, and (b) a core treatment plus structured teacher adaptations condition (Adaptive Treatment), where teachers were given opportunities to modify and extend program activities to support student engagement in both asynchronous and synchronous learning contexts. In the Adaptive Treatment condition, teachers were encouraged to apply their professional judgment and insights to instructional delivery while adhering to the intervention's core principles. They also participated in TBL sessions, designed to foster collaboration and the co-construction of knowledge to increase the relevance and effectiveness of the intervention. The core components of the intervention in both conditions included asynchronous reading activities using a digital app and print-based text reading, as well as Zoom-delivered synchronous lessons led by classroom teachers.

This study makes a unique contribution to the literature by investigating the causal impact of structured teacher adaptations during the pandemic on third graders' engagement and learning outcomes. We addressed three research questions:

- **RQ1:** Compared to Core Treatment, what are the effects of Adaptive Treatment on student engagement in asynchronous digital app and print-based reading activities, as well as synchronous curriculum activities?
- **RQ2:** Compared to Core Treatment, what are the effects of Adaptive Treatment on students' science vocabulary knowledge depth, science background knowledge, science content reading comprehension, and domain-general reading comprehension outcomes? Are the effects of Adaptive Treatment on student outcomes consistent across school sites?
- **RQ3:** Compared to Core Treatment, what is the effect of Adaptive Treatment on teachers' fidelity to the intervention?

Methods

Study Design and Participants

This study employed a cluster (school) randomized controlled trial (RCT) involving 2,247 third-grade students and their 95 classroom teachers from 26 elementary schools in an urban district in the southeastern United States. Classroom teachers, stratified by schools, were randomly assigned to either the Adaptive or Core Treatment condition. The consort diagram in

Figure 2 illustrates the flow of the randomization process and attrition rates for each condition. Attrition was 13.4% in the Adaptive Treatment condition (n = 1,088 to 942) and 16.1% in the Core Treatment condition (n = 1,159 to 972), resulting in a final analytic sample of 1,914 students. A test for differential attrition indicated no statistically significant differences between conditions (p = .12).

Table 1 presents student characteristics by treatment condition. The two groups were generally balanced across most characteristics, with a few differences: the Adaptive Treatment group included a higher proportion of Hispanic students (38% vs. 29%, $p < .01$), a slightly greater representation of medium SES students (40% vs. 36%, $p < .01$), and fewer high SES students (21% vs. 28%, $p < .001$) compared to Core Treatment. These minimal baseline differences support the internal validity of the study design and a causal interpretation of the results. This RCT is preregistered, and the replication data and code are accessible at the following URL: [BLINDED].

Intervention Treatment Conditions

Core and Adaptive Treatment

Table 2 provides an overview of how the core components, asynchronous learning activities using the digital app and print-based trifolds and synchronous online learning via Zoom, were operationalized in both Core and Adaptive Treatment conditions, along with acceptable adaptations implemented by Adaptive Treatment teachers. Both Core and Adaptive Treatment teachers implemented the core components of the intervention. In January 2021, all participating teachers attended a 60-minute synchronous online training session led by the research team on Zoom. The session focused on the intervention's theory of change, empirical evidence, curriculum framework, and detailed lesson plans. The research team emphasized the importance of schema-building in supporting reading comprehension in a content-rich curriculum, highlighting how learners integrate new information with prior knowledge to construct deeper understanding and improve comprehension of complex texts. Teachers were provided with practical strategies to facilitate schema construction, enabling students to engage more effectively with complex texts, strengthen their conceptual understanding, and transfer their learning to new literacy and content-based contexts (Kintsch, 1993; McNamara et al., 1996).

To maintain an interactive and engaging format, each session was limited to 25 participants, with multiple sessions offered to accommodate teachers' availability. In addition to the synchronous training, teachers received electronic lesson materials, a pacing calendar, and instructional videos to familiarize themselves with students' app and trifolds activities. These resources for the at-home activities were also made available to students' families in both English and Spanish versions. The research team provided teachers with ongoing support through frequent email communication and synchronous office hours to address questions and troubleshooting issues throughout the implementation of the intervention.

Digital App. The research team developed the digital app to incorporate gamified learning experiences to support students' language and literacy development at home. The app featured 15 interactive activities per science book, targeting both code- and meaning-based skills, with a read-aloud option for students to follow along. These activities (see Table A1 in Appendix A for detailed descriptions and examples) included tasks such as identifying sight words and syllables, understanding word parts like prefixes and suffixes, spelling with morphemes, constructing sentences, and enhancing reading comprehension by focusing on main ideas and textual content. The activities were also designed to develop metalinguistic awareness by engaging students with humor, idioms, and metaphors, enhancing their ability to analyze and interpret language nuances and deepening their understanding of its abstract and figurative aspects. The app supported personalized learning by tailoring activity difficulty to align with students' initial reading proficiency levels, providing an engaging and adaptive platform for developing literacy skills.

Print-Based Books and Trifolds. In addition, all students received four print-based science books, each accompanied by a paper trifold. After reading each book, students were encouraged to work independently at home on the trifold, which featured activities to deepen their understanding of a key concept word (e.g., *skeletal, muscular*). Specifically, students completed tasks such as filling in blanks in sentences and identifying semantic associations (e.g., *What word is related to the word skeletal?*) using the answer bank. The trifold also included three questions on vocabulary and reading comprehension, prompting students to reflect on and apply what they had learned from the book.

Synchronous Lessons. From February to March, teachers conducted 15 synchronous Zoom lessons (30–35 minutes each) over five weeks in the human body systems unit. These lessons focused on building students' conceptual understanding of the muscular, skeletal, and nervous systems, stem cells, and strengthening morphological knowledge, vocabulary development, and reading comprehension. The lessons integrated structured activities to promote critical thinking and collaboration, encouraging students to engage meaningfully with the content.

Each lesson began with activities aimed at setting learning goals, sparking curiosity, and activating prior knowledge. Teachers encouraged students to observe, make predictions, and ask questions about human body systems by incorporating the knowledge that students had acquired from the app and trifolds activities into the class discussions. Teachers also acknowledged students' efforts in preparatory tasks (e.g., app and trifolds activities), reinforcing their progress and motivation. A key focus of the lessons was vocabulary development, in which teachers guided students in analyzing target words and their morphological structures (e.g., prefixes like micro- or bio-). Students also explored connections between new vocabulary and broader scientific concepts using visual tools like concept maps.

In small breakout groups, students applied their vocabulary knowledge and conceptual understanding by collaboratively annotating images and texts, synthesizing information, and explaining related scientific concepts. For example, students were encouraged to apply their understanding of the prefix *micro*- when examining an image of a microscope, collaboratively creating descriptions. During collaborative research activities, students explored real-world applications of the concepts they were learning. For example, teachers introduced examples about astronauts in space, explaining how muscles weaken in microgravity and why exercise is essential to keep them strong during long missions. With their peers, students investigated questions like "*Why are muscles important, and how can we maintain them?*" Students first worked independently to make predictions and conduct research using science books and trifolds, then collaborated in groups to discuss their findings, synthesize key takeaways, and present their insights to the class.

Adaptive Treatment

Adaptive Treatment teachers additionally incorporated structured teacher adaptations to the core components to enhance student engagement and learning outcomes. Table 3 specifies the type and nature of content-based and procedural adaptations completed by teachers in the Adaptive Treatment group. These teachers along with their school's literacy facilitator completed four online modules developed by the researchers. This two-hour asynchronous course was designed to prepare them to make intentional and effective adaptations, with the primary goal of increasing student engagement with the app and trifolds activities. The course also sought to deepen teachers' understanding of the science of reading by emphasizing explicit instruction in word analysis and guiding them in applying morphological awareness and vocabulary learning in the core curriculum.

To further support the Adaptive Treatment group, the research team conducted a 60-minute synchronous TBL meeting with teachers and literacy facilitators from each Adaptive Treatment school. The session began with an exploration of potential barriers to engagement, such as students' lack of motivation to complete the activities independently, and challenges in integrating resources into existing schedules. The researchers facilitated a collaborative dialogue, encouraging teachers to share their insights and expertise in developing shared action plans to address these challenges. To guide this process, the researchers provided a list of adaptive strategies to increase student engagement with the app and trifolds. These strategies included modeling app and trifold usage for students, providing opportunities for them to share their learning during the synchronous Zoom lessons, setting explicit expectations for app use, involving families to encourage participation, incentivizing engagement, and leveraging app data to monitor and track usage. Teachers and literacy facilitators, working collaboratively, selected the strategies from the list and engaged in problem-solving discussions to tailor the chosen adaptations to their school contexts.

Moreover, Adaptive Treatment teachers delivered an additional 35-minute extension lesson on Day 16, which was not included in the Core Treatment condition. This lesson (see Appendix B) was designed to enhance students' schema for the scientific concept of *systems* by expanding their understanding of the human body lessons and encouraging abstract thinking through targeted language and vocabulary development. During the TBL meeting, the research team shared a list of suggested strategies to support the delivery and adaptations of this lesson. These included interactive activities such as "mad libs" to introduce the concept of *systems*, modeling scientifically accurate presentations, and incorporating multimedia elements to enhance engagement. Teachers and literacy facilitators selected and adapted the strategies from the list to align with their instructional goals, making the lesson accessible and engaging for their students.

Measures

Student Outcome Measures

Student Engagement. Student engagement in asynchronous learning activities with the digital app and print-based reading was assessed. For the digital app, engagement was measured across five key metrics using backend data: (a) *aver access app library*, which indicated whether students accessed the digital library containing science books and related activities; (b) *the number of science books completed*, reflecting how many books and their corresponding activities students finished; (c) *the number of target words accessed*, capturing the number of multisyllabic or lesson-relevant words students encountered and the frequency of their exposure during specific app activities; (d) *total time spent on app activities*, measuring the cumulative time students engaged with the app; and (e) *overall app activity accuracy*, ranging from 0 to 100%, serving as an indicator of students' mastery of the app content based on their performance in answering activity questions correctly.

Engagement with the app activities closely aligned with the synchronous curriculum lessons was measured using seven metrics: (a) *ever access curriculum lessons*, indicating whether students accessed any curriculum-aligned lessons, (b) *proportion of all curriculum lessons completed*, (c) *proportion of interactive read-aloud lessons completed*, (d) *proportion of word-sleuthing lessons completed*, (e) *accuracy on curriculum activities*, reflecting students' performance on curriculum-related activities, (f) *accuracy on end-of-unit quizzes*, indicating mastery of unit content, and (g) *total time spent on curriculum lessons.*

Additionally, students' motivational engagement with the app was assessed through an in-app self-reported survey focusing on three dimensions: (a) *enjoyment of app activities*, which was rated on a 4-point scale ranging from 1 (*I didn't like it*) to 4 (*I liked it a lot*), capturing students' enjoyment of their experience with the app; (b) self-competence beliefs, assessed on a 3-point scale (1 = *ok reader*, 2 = *good reader*, 3 = *great reader*), reflecting students' perceptions of their reading abilities; and (c) *perceived task difficulty*, measured on a 3-point scale (1 = *too easy*, 2 = *just right*, 3 = *too hard*), indicating how challenging students found the app activities. Finally, engagement with print-based trifolds activities was evaluated by (a) *whether any trifolds were returned* and (b) *the number of trifolds returned*, indicating task participation and completion.

Science Vocabulary Knowledge Depth. We utilized a researcher-developed semantic association task (Appendix C) to assess students' science vocabulary knowledge depth at the conclusion of the intervention. This task measured students' understanding of science academic words explicitly taught or incidentally encountered during the lesson unit and their ability to identify and connect semantically related words (Stahl & Fairbanks, 1986). Each item presented a target word (e.g., *carnivore*) paired with four options (e.g., *fruit, care, meat,* and *prey*), asking students to select two words semantically related to the target. Distractors were designed to include words that were either phonetically similar to the target word in their initial or final sounds or semantically contrasting. All questions and options were read aloud to students. Each item was scored as 1 for correct selections only and 0 otherwise. The internal consistency (Cronbach's α) was .90.

Science Background Knowledge. Students' background knowledge was assessed across three science topics: monkeys, birds, and skyscrapers. For each topic, students listened to a passage and three corresponding question items read aloud to them (see Appendix D). In the monkey topic, for instance, students answered questions such as identifying which muscles can be controlled or which muscles in a monkey never rest. Students responded individually, and their answers were scored dichotomously (1 = *correct*, 0 = *incorrect*). The internal consistency for the nine-item measure across all topics was .55.

Science Content Reading Comprehension. After the science background knowledge assessment was administered, science content reading comprehension was evaluated post-intervention using a 29-item multiple-choice instrument (Appendix E) designed to measure students' understanding of main ideas and concepts in science content passages. The assessment passages were categorized along a transfer continuum—near, mid, and far—based on their relevance to the instructional content. Passage content varied in prior exposure, ranging from extensive familiarity with target vocabulary (near transfer) to no prior exposure (far transfer), as well as in contextual similarity to the instructional focus on human body systems. The near-transfer passage involved scientists studying monkey heart attack recovery, and the mid-transfer passage described adaptations in migratory birds' skeletal and muscular systems. The far-transfer passage, unrelated to previously taught vocabulary or contexts, explored the anatomy of a skyscraper as a nonliving system.

Each passage included multiple-choice questions designed to assess main idea identification, word or phrase meaning, scientific concept understanding, and knowledge integration, with responses scored dichotomously. The internal consistency for the full assessment was .86, with subscale reliabilities for the near-, mid-, and far-transfer passages at .72, .63, and .68, respectively.

Domain-General Reading Comprehension. We measured students' domain-general reading comprehension using two state-standardized assessments: Measure of Academic Progress (MAP) and End-of-Grade (EOG) assessments.

MAP Reading Assessment. The MAP Reading Assessment (NWEA, 2019) was used to evaluate Grade 3 students' domain-general reading comprehension. This computerized adaptive test adjusts the difficulty of each question based on a student's responses to prior items (Thum & Kuhfeld, 2020). The assessment consists of 43 unique items for each student and typically takes 25–40 minutes to complete. It provides a composite score derived from four strands: narrative and informational text comprehension, vocabulary use and functions, foundational skills, and language and writing. Scores are reported on the Rasch Unit (RIT) scale, calibrated using item response theory, to reflect students' reading proficiency. Test-retest reliability coefficients for composite scores in [STATE BLINDED] ranged from .79 to .86 (NWEA, 2019).

Statewide EOG Reading Assessment. The EOG Reading Assessment for Grade 3, administered by the [STATE BLINDED] Department of Public Instruction (DPI) in the spring, was used as a measure of students' domain-general reading comprehension. The assessment aligns with the [STATE BLINDED] Standard Course of Study for English Language Arts and evaluates mastery of foundational reading skills, vocabulary, and comprehension of literary and informational texts. The assessment provides a summative evaluation of a student's mastery of grade-level standards, providing a snapshot of performance at the end of the academic year. It includes multiple-choice questions that assess literal and inferential comprehension, vocabulary in context, and text analysis. Item response theory-scaled EOG test scores were used to estimate overall reading proficiency and to model variation in performance across subgroups. Internal consistency for the assessment is approximately .90 ([STATE BLINDED] DPI, 2020).

Fidelity of Implementation

We assessed treatment fidelity by examining two key dimensions: differentiation and responsiveness (Dane & Schneider, 1998). Differentiation refers to the unique features of a program intended to facilitate meaningful change (Power et al., 2005). Differentiation was measured through a 13-item teacher survey that captured self-reported strategies for promoting student engagement in digital app-based and print-based activities. The survey items included practices such as modeling, setting completion expectations, dedicating instructional time, communicating with families, and following up with students who did not complete activities. Teachers in both treatment conditions rated each item on a 5-point Likert scale (1 = *never* to 5 = *every day*), with higher scores indicating a greater frequency of implementation. The internal consistency was .83.

Responsiveness, defined as the degree of participant engagement in a program (Sánchez et al., 2007), was evaluated based on the quality of student-teacher interactions observed in audio-recorded synchronous Zoom classes. We developed a rubric (Appendix F) to assess three dimensions of interaction: engagement, questioning, and feedback. Each dimension was rated on a 4-point Likert scale (1 = *low*, 2 = *medium low*, 3 = *medium high*, and 4 = *high*) based on evidence from the recorded lessons. Engagement was measured as the frequency and depth of back-and-forth exchanges, dialogic interactions, and follow-up questions, with higher scores indicating more sustained reciprocal dialogue. Questioning assessed the use of open-ended prompts to elicit higher-order thinking, capturing the extent to which teachers moved beyond closed-ended questions to encourage analysis and reasoning. Feedback focused on cognitive engagement through clarifications, connections to prior learning, and positive reinforcement, with higher scores reflecting more detailed and constructive responses.

We rated 153 synchronous lesson sessions in Lessons 3, 4, 7, and 8 of the 10-day science unit. These lessons were selected due to their interactive and varied instructional activities, providing opportunities to observe variations in engagement, questioning, and feedback. To ensure consistent rubric application, two researchers, blinded to the treatment condition, jointly rated each lesson from four classroom sessions and resolved scoring discrepancies to align their interpretations of the criteria. After this calibration process, one researcher independently rated the remaining sessions, and the second researcher independently rated a random 20% of these sessions to assess reliability. The overall agreement was .91 (Cohen's κ = .79).

Data Analysis

To address RQs1 and 2, we employed a three-level hierarchical linear model (HLM) to account for the nested structure of the data in our multi-site, cluster-randomized study design. The model includes students (level 1), teachers (level 2), and schools (level 3). The reduced-form model is expressed as follows:

$$Y_{ijk} = \alpha_k + \beta_1 Adaptive_{jk} + \sum_{p=2}^{11} \beta_p COV_{p_{ijk}} + \zeta_{jk} + \epsilon_{ijk} \qquad \text{(Equation 1)}$$

$$\zeta_{jk} \sim N(0, \sigma^2_\zeta)$$

$$\epsilon_{ijk} \sim N(0, \sigma^2_\epsilon),$$

where Y_{ijk} denotes the outcome for student i in teacher j's classroom in school k, and α_k represents the fixed effect of school k. The term β_1 captures the adjusted intention-to-treat (ITT) effect of the Adaptive Treatment, and $\beta_p COV_{p_{ijk}}$ represents a set of student-level covariates (i.e., baseline MAP reading scores, gender, race/ethnicity, English learner status, home language, individualized education plan status, and neighborhood poverty). The variable $Adaptive_{jk}$ is a binary indicator of whether the teacher was assigned to the Adaptive Treatment condition. The model also includes a teacher-level random effect, ζ_{jk}, and student-level residual, ϵ_{ijk}, both of which are assumed to follow normal distributions.

For RQ2, we further examined heterogeneous treatment effects on student learning outcomes across school sites using a model that incorporated random intercepts for schools and a random slope for the treatment condition at Level 3 (Equation 1). This specification allowed for the estimation of both the average treatment effect across schools and the variability (standard deviations) in treatment effects across school sites. To address missing data on student demographics, we implemented multiple imputations. All analyses were conducted using Stata 17.0 (StataCorp, 2021).

For the analysis of teachers' fidelity of implementation for RQ3, we used Ordinary Least Squares (OLS) regression models with school fixed effects and teacher-level covariates, including prior intervention experience, years of teaching, National Board Certification status, and state reading course completion. The model is specified as:

$$Y_{jk} = \alpha_k + \beta_1 Adaptive_{jk} + \sum_{p=2}^{5} \beta_p COV_{jk} + \zeta_{jk} \quad \text{(Equation 2)}$$

$$\zeta_{jk} \sim N(0, \sigma_{\zeta}^2),$$

where Y_{jk} denotes the outcome for teacher j in school k, and α_k represents the fixed effect of school k. The term β_1 captures the adjusted ITT effect of the Adaptive Treatment, while $\beta_k COV_{jk}$

includes teacher-level covariates. The teacher-level residual ζ_{jk} is assumed to be normally distributed.

Results

RQ1: The Effect of Structured Adaptations on Student Engagement in Asynchronous and Synchronous Learning Activities

Table 4 presents descriptive statistics and treatment effects comparing student engagement outcomes between the Adaptive and Core Treatment conditions. Students in the Adaptive Treatment condition demonstrated significantly higher engagement in digital app activities compared to those in the Core Treatment condition. Specifically, the Adaptive Treatment group was more likely to access the app library ($\beta = 0.05$, $p < .05$), complete more science books on the app library ($\beta = 0.29$, $p < .001$), and access more target words ($\beta = 0.24$, $p < .05$). They also spent slightly more time on app activities ($\beta = 0.17$, $p < .10$), though this difference was marginal. No significant difference was observed in overall app activity accuracy ($\beta = -0.02$, $p > .05$) between the two groups. Students in the Adaptive Treatment condition also reported significantly higher enjoyment of lesson activities ($\beta = 0.10$, $p < .05$) and stronger self-competence beliefs ($\beta = 0.11$, $p < .05$) compared to their counterparts. Perceived task difficulty, however, did not differ significantly between the two groups ($\beta = -0.01$, $p > .05$). For print-based books/trifolds activities, students in the Adaptive Treatment condition returned a slightly higher number of trifolds, but this difference was not statistically significant ($\beta = 0.17$, $p > .05$). Similarly, the likelihood of returning any trifolds was comparable across groups ($\beta = 0.01$, $p > .05$). There were also no significant differences between the groups on synchronous lesson activities (*ps* > .05).

RQ2: The Effect of Structured Adaptations on Student Learning Outcomes and Heterogeneous Treatment Effects

Table 5 shows descriptive statistics and the effects of Adaptive Treatment on student learning outcomes. No statistically significant difference was found between the Adaptive and Core Treatment groups for science vocabulary knowledge depth (effect size [ES] = 0.02, $p > .05$). However, students in the Adaptive Treatment group scored slightly higher on science background knowledge, with a small but statistically significant effect (ES = .09, $p < .05$). For science content reading comprehension**,** students in the Adaptive Treatment condition performed better overall across all passages compared to those in the Core Treatment condition (ES = .07, $p < .05$). Analysis by passage type revealed that the Adaptive Treatment group outperformed the Core Treatment group on the near-transfer passage (ES = .11, $p < .05$), while differences for mid-transfer (ES = .05, $p > .05$) and far-transfer passages (ES = .03, $p > .05$) were not statistically significant. In terms of domain-general reading comprehension, no significant differences were observed between the two groups for either MAP scores (ES = .004, $p > .05$) or EOG reading scores (ES = -0.02, $p > .05$).

The analysis of heterogeneous treatment effects using a random intercept and random slope model revealed minimal variability in treatment effects across school sites. Point estimates of SD for treatment effects were near 0 for all student learning outcomes, indicating that the effects of Adaptive Treatment were consistent across school sites.

RQ3: Teachers' Fidelity of Implementation

Differentiation in teacher-reported strategies for promoting asynchronous reading activities varied between the treatment conditions (see Table 6). For digital app activities, teachers in the Adaptive Treatment group were significantly more likely to follow up with students who did not complete activities compared to those in the Core Treatment group ($p < .01$). Other app-related strategies, such as encouraging students to complete activities and modeling the use of the app library, showed positive trends but were not statistically significant ($ps > .05$).

For print-based reading activities, significant differences emerged in favor of the Adaptive Treatment condition. Teachers in this group were more likely to model the use of trifolds, set expectations for the number of trifolds, and communicate with families about trifolds ($ps < .05$). Other strategies, such as following up with students who did not complete trifolds, did not differ significantly between the groups ($ps > .05$).

Responsiveness, as reflected in observed student-teacher interactions during the synchronous lessons, was consistently higher in the Adaptive Treatment condition across all three assessed dimensions (see Table 7). Specifically, teachers in the Adaptive Treatment group demonstrated significantly greater engagement with students, provided more details and frequent use of feedback, and employed higher levels of questioning to elicit deeper reasoning and critical thinking compared to the Core Treatment group ($ps < .01$).

Discussion

This experimental study was designed to assess the effectiveness of structured teacher adaptations to a third-grade content literacy intervention that integrated asynchronous print and digital reading activities with synchronous lessons. Using a within-school cluster randomized trial design, 95 third-grade teachers and their students were randomly assigned to one of two conditions: the Core Treatment condition, which replicated intervention components validated in previous research, and the Adaptive Treatment condition, which incorporated structured adaptations with TBL activities building on the Core Treatment. The Adaptive Treatment condition aimed to empower teachers to apply both general research knowledge and their professional expertise to enhance implementation during the COVID-19 pandemic.

Our findings provide a comprehensive and detailed account of implementation processes and an outcome-wide analysis (VanderWeele, 2017) of student learning using multiple measures of engagement and learning. Results indicate that third graders whose teachers were in the Adaptive Treatment condition demonstrated higher engagement levels during both asynchronous digital activities and synchronous lessons compared to those in the Core Treatment condition. Furthermore, this increased student engagement was also reflected in better outcomes for students' science background knowledge and science reading comprehension. Notably, the null effects on science vocabulary knowledge depth and domain-general reading suggest no evidence of adverse effects. Although no single factor led to these gains, the most critical difference between the two conditions was the integration of TBL activities in the Adaptive Treatment that enabled teachers to improve student engagement with asynchronous learning and synchronous lessons in ways that fostered the development of background knowledge and comprehension of science tests. Below, we highlight key findings for each research question and discuss broader implications.

Effectiveness of Structured Teacher Adaptations on Student Engagement

Our first aim was to examine whether teachers in the Adaptive Treatment condition could enhance student engagement with asynchronous learning activities and synchronous lessons. In many ways, there is clear evidence of a cascading series of effects linking upstream TBL activities for teachers and downstream student engagement processes. In operationalizing the adaptive treatment, we emphasized "active ingredients" embedded in the TBL activities, which afforded teachers opportunities to collaborate with the research team and their peers, including literacy coaches. As noted in Table 3, TBL discussions enabled teachers to customize communication with families, utilize learning analytics from digital app data to promote equitable access to digital resources (Bernacki, 2025), and deliver an extension lesson to facilitate students' mastery of core science concepts (e.g., *system*) introduced during the synchronous Zoom lessons. These practices, moreover, emphasized active learning and time on task, which are two critical components of effective and flexible teaching identified by an American Psychological Association's task force on pivot teaching during the pandemic (Cavazos et al., 2024). While no differences were observed in engagement with print-based activities and reported greater motivation and task difficulty alignment compared to those taught by Core Treatment teachers. These findings suggest that the structured adaptations originating from TBL activities significantly increased student participation in digital, game-based activities.

Furthermore, these findings underscore the importance of an asset-based approach in driving student engagement (Gabriel & López, 2024). Strategies co-developed during TBL sessions provided teachers with collaborative scaffolding methods and problem-solving approaches that were both responsive to students' challenges and grounded in their strengths. This process allowed teachers to draw on students' prior knowledge, curiosity, and interests in topics in the unit, as well as their ability to connect science content to real-world experiences. By situating instructional strategies within contexts relevant to students' daily lives, teachers validated students' contributions, fostered active participation, and cultivated a sense of competence and belonging (Gray et al., 2018; López, 2017).

Student engagement in both asynchronous and synchronous learning contexts was further supported by the deliberate integration of knowledge from asynchronous activities into synchronous class discussions. Teachers in the Adaptive Treatment group used app participation data to co-design follow-ups that celebrated students' progress, addressed areas needing targeted support, and reinforced connections between vocabulary learned in app-based activities and broader scientific concepts discussed during the synchronous lessons. This deliberate integration likely helped students see the relevance of their learning across contexts and modalities, fostering a cohesive and engaging educational experience.

Additionally, the design of the app may have contributed to promoting student engagement by integrating gamified learning experiences with features that supported multiple aspects of language and literacy development. Its read-aloud functionality for science books, coupled with activities targeting both constrained and unconstrained literacy skills (Paris, 2005; Snow & Matthews, 2016), offered structured opportunities for students to practice and expand their literacy competencies. Activities designed to cultivate metalinguistic awareness, such as engaging with humor, idioms, and metaphors, encouraged students to analyze and reflect on language at a deeper level (Spector, 1996). These features likely enhanced exposure to target vocabulary and cognitive engagement with language while reinforcing the importance of leveraging students' analytical thinking and prior experiences as assets in language and literacy development (Zipke, 2008).

The findings underscore the synergy between TBL-driven adaptive strategies and the multifaceted digital app design in fostering student engagement during the challenging pandemic period. Adaptive Treatment teachers used collaborative problem-solving to address barriers, implement responsive scaffolding, and support students' engagement in navigating self-directed tasks. Simultaneously, the app's gamified elements and activities targeting both foundational and higher-order literacy skills encouraged independent learning and deeper cognitive engagement, especially in contexts demanding autonomy and self-regulation (Authors, 2021; Bond, 2020;

Zimmerman, 2002). Together, these approaches illustrate the potential of integrative adaptive teaching practices with digital tools to enhance participation and promote active learning in online educational contexts (Archambault et al., 2022).

Effectiveness of Structured Teacher Adaptations on Learning Outcomes

The second research aim was to examine whether students in the Adaptive Treatment condition enjoyed larger gains in science vocabulary knowledge depth, science background knowledge, and domain-specific (i.e., science) and domain-general reading comprehension. Beyond engagement outcomes, Adaptive Treatment also significantly impacted learning outcomes, particularly in science background knowledge (ES = .09) and science content reading comprehension (ES = .07). Notably, however, the positive impact on domain specific science reading comprehension was largely driven by the near-transfer passive (ES = .11) rather than the mid-transfer (ES = .05) and far-transfer (ES = .03) passages. These findings suggest that the Adaptive Treatment supported students' comprehension of a near-transfer passage about how scientists help monkeys recover from heart attacks, which included all the directly taught vocabulary from the teacher-directed synchronous lessons. More broadly, the findings underscore the need to move beyond the simple view of reading to a more complete view of reading that models heterogeneity at the reader and passage levels (Francis et al., 2018). This perspective calls on researchers to use statistical models that more precisely assess for whom and on which tasks online and in-person literacy interventions are most effective.

In addition, the effects on student reading outcomes suggest that even small adjustments to teacher professional learning can have a modest and positive impact on students' understanding of disciplinary concepts in science reading passages. A key TBL adaptation was the inclusion of an extension lesson on the concept of systems, specifically designed for the

Adaptive Treatment group. In this lesson, teachers connected the scientific concept of systems to students' prior knowledge, facilitating schema-building and integrating new information with existing schemas. Bu deepening students' conceptual understanding and encouraging them to apply their learning to novel contexts, this approach supported text comprehension and retention of specialized content (Kintsch, 1993; McNamara et al., 1996). These targeted scaffolding practices demonstrate how adaptive instruction can address the cognitive demands of disciplinary learning while fostering deeper connections to the learning material.

It is noteworthy that the Adaptive Treatment condition not only supported teachers in adhering to the core instructional framework but also enhanced the quality of student-teacher interactions during the synchronous lessons, as shown in Table 7. While both groups incorporated evidence-based practices that supported schema building and language and literacy skills, the Adaptive Treatment condition extended these practices by introducing additional scaffolding and differentiation tailored to the cognitive and linguistic demands of science texts. These findings demonstrate how structured, teacher-driven modifications in a shared instructional framework can optimize instructional experience and effectiveness, offering opportunities for comprehension and knowledge integration without compromising fidelity to the intervention's core principles (Authors, 2017b; Bryk et al., 2015; Goldenberg & Gallimore, 1991).

The broader significance of these findings is particularly evident when considering the study context. With approximately 40% of students from low-SES backgrounds and implementation occurring during the COVID-19 pandemic, the Adaptive Treatment condition demonstrated its potential to mitigate systemic inequities in education. Pandemic-related disruptions disproportionately affected under-resourced communities, exacerbating educational disparities (Authors, 2022; Kuhfeld et al., 2022). Despite these challenges, students' learning gains suggest that the Adaptive Treatment condition provided additional support that amplified the impact of content literacy instruction, leading to stronger domain-specific learning outcomes. While both treatment conditions maintained learning continuity, the Adaptive Treatment teachers leveraged the collaborative TBL framework to align evidence-based practices with students' needs. This participatory approach enabled teachers to foster deeper reading engagement and comprehension and support equitable and impactful learning outcomes even under adverse conditions (Authors, 2023).

Nevertheless, the lack of additional gains in science vocabulary knowledge suggests that both treatment conditions benefited equally from the core components of the intervention, leaving limited room for further enhancement in this outcome. The null findings for domain-general reading comprehension also diverge from prior work (Authors, 2017a), which reported a modest advantage (0.12 SD) for students in the adaptive condition on reading comprehension. Unlike Authors, 2017a's study, conducted in an in-person learning environment, this study was implemented in a remote context, where the challenges of navigating non-traditional instructional settings may have constrained students' learning potential. This inconsistency may also reflect the inherent difficulty of transferring content-specific gains to broader literacy domains (Authors, 2024a). Domain-general reading comprehension develops incrementally over extended periods and requires sustained practice (Paris, 2005; Snow & Matthews, 2016), which was likely disrupted by pandemic-related constraints such as reduced instructional hours and inconsistent reading instruction. Moreover, statewide standardized assessments used to measure these skills may have lacked the sensitivity to detect the nuanced, intervention-specific effects, potentially obscuring more subtle impacts.

The analysis of the heterogeneous treatment effect revealed minimal variability in outcomes across school sites, highlighting the consistency in the impact of the Adaptive Treatment condition regardless of geographic or demographic differences. This consistency is particularly noteworthy given the disparities in resources, instructional conditions, and student demographics that often characterize educational systems (Hanushek, 1997; Raudenbush et al., 1998). By balancing flexibility with fidelity, structured teacher adaptations enabled schools, irrespective of their starting conditions, to align evidence-based practices with their unique contexts while preserving the integrity of the instructional core principles (Bryk et al., 2015; Goldenberg & Gallimore, 1991). These findings align with equity-driven pedagogical frameworks, demonstrating how well-designed adaptations can mitigate contextual disparities and promote comparable learning opportunities for all students. The results emphasize the potential of structured teacher adaptations to serve as a scalable approach to bridging the research-practice divide and advancing educational equity across diverse instructional settings (Coburn, 2003).

Structured Adaptations and Fidelity of Implementation

Our third aim was to investigate whether teachers in the Adaptive Treatment condition could maintain levels of fidelity comparable to those in the Core Treatment condition, particularly in terms of differentiation and responsiveness components of fidelity (Dane & Schneider, 1998). Teachers in the Adaptive Treatment condition exhibited higher fidelity across these dimensions, providing insights into balancing adherence to core instructional principles with context-sensitive adaptations. Differentiation, operationalized through teacher-reported strategies for asynchronous activities, highlights how fidelity can extend beyond rigid adherence to protocols by incorporating teacher-driven scaffolding practices tailored to diverse student needs (Power et al., 2005). Adaptive Treatment teachers' emphasis on personalized follow-ups for app activity completion, modeling trifold usage, setting expectations, and engaging in family communication (see Table 6) exemplified fidelity to the intervention's core engagement goals while simultaneously addressing the unique demands of asynchronous learning and varied student needs. These findings support the notion that fidelity and adaptation can coexist, fostering flexibility without compromising alignment with the core principles (Authors, 2023).

It is noteworthy that the greater responsiveness observed in the Adaptive Treatment condition highlights how fidelity can encompass the quality of instructional interactions. Teachers in this condition demonstrated higher levels of open-ended questioning, meaningful feedback, and dialogic engagement, fostering reciprocal exchanges and active student participation during the synchronous lessons. These practices demonstrate how structured adaptations enabled teachers to maintain the pedagogical integrity of the intervention while elevating the quality of student-teacher interactions. These findings align with conceptualizations of fidelity that emphasize both adherence to prescribed practices and the quality of their implementation (Dane & Schneider, 1998; O'Donnell, 2008). By incorporating responsiveness, structured adaptations bridged the theoretical rigor of fidelity with the practical demands of adaptation, underscoring the critical role of teacher agency in achieving scalable and sustainable outcomes (Authors, 2017b). In doing so, structured teacher adaptations not only safeguarded the integrity of the intervention's core principles but also empowered educators to exercise professional judgment to respond to contextual demands, thus enhancing its relevance and impact.

These findings contribute to the broader discourse on scaling evidence-based interventions by demonstrating the importance of integrating fidelity and adaptation (Dede, 2006; Hill & Erickson, 2019). Rather than constraining innovation, fidelity, when conceptualized inclusively, can guide meaningful modifications that enhance both implementation quality and educational equity (Bryk et al., 2015; Durlak & DuPre, 2008). This integrated perspective advances our understanding of how structured adaptation frameworks can bridge the research-practice gap, ensuring that interventions remain scientifically rigorous while responsive to the diverse and evolving needs of classrooms (Goldenberg & Gallimore, 1991; K.E. Joyce & Cartwright, 2020).

Limitations and Future Directions

Several limitations merit scrutiny in future research. First, the absence of a traditional treatment-versus-control (business-as-usual) comparison group limits the ability to assess the unique contributions of structured adaptations relative to regular instructional practices. While the study design ensured equitable access to instructional opportunities during school closures, it precluded establishing a baseline comparison to determine the specific contributions of adaptive treatments. Second, the mechanisms underlying the observed effects were unexplored in this study. Examining whether factors such as enhanced teacher responsiveness or increased student engagement mediate the relationship between adaptive treatments and literacy outcomes could offer a more nuanced understanding of how these interventions produce their effects and help identify critical leverage points for maximizing their effectiveness. Finally, there is growing attention to using learning analytics to study collaborative learning (Yan et al., 2025). Future research should aim to capture the depth and variability of teachers' collaborations within the TBL activities. Furthermore, incorporating artificial intelligence-based engagement metrics, such as real-time interaction patterns, prosodic features, or behavioral indicators, would offer a richer and more dynamic understanding of teacher engagement as they co-construct adaptations with researchers and their peers, and then implement and fine-tune them in the classroom.

Conclusion

There is emerging evidence that structured teacher adaptations can enhance program effectiveness (e.g., Authors, 2017a; Lemons et al., 2014; Neuman et al., 2011), particularly amidst diverse challenges. When supported by researchers and aligned with core principles, these adaptations maintain fidelity while increasing student engagement and learning outcomes by addressing unique student and local context needs. For policymakers and educational leaders, our findings advocate for policies that support adaptive frameworks. Such policies should integrate core intervention components with flexible, localized adaptation strategies, ensuring that evidence-based programs are robust yet responsive to contextual demands. Finally, this study suggests that asset-based learning environments build autonomy, competence, and belongness supports (Ryan & Deci, 2000, 2020), which are both causes and consequences of more effective literacy instruction and learning. In essence, structuring teacher adaptations may represent an evidence-based model for supporting fidelity and flexibility and humanizing learning spaces for teachers and their students.

Table 1

Characteristics of Student Participants by Treatment Condition and Balance Checks (N = 1,914)

Characteristics	Overall (N = 1,914)		Adaptive treatment (n = 942)		Core treatment (n = 972)		Balance checks[b]	
	M[a]	*SD*	M[a]	*SD*	M[a]	*SD*	Difference	z-score
Male	.49	.50	.50	.50	.49	.50	0.01	0.58
Black	.37	.48	.35	.48	.38	.49	-0.01	-0.47
Asian	.09	.29	.07	.26	.11	.31	-0.03†	-1.66
Hispanic	.34	.47	.38	.49	.29	.45	0.06**	3.14
White	.17	.38	.17	.37	.18	.39	0.01	0.26
Other	.03	.17	.02	.15	.04	.19	-0.01†	-1.76
AIG	.12	.33	.13	.34	.12	.33	0.01	0.79
English learners	.25	.43	.27	.44	.22	.42	0.04†	1.67
IEP	.08	.28	.08	.28	.09	.28	0.00	-0.07
Low SES	.39	.49	.41	.49	.36	.48	0.00	0.02
Medium SES	.37	.48	.37	.48	.36	.48	0.05**	2.76
High SES	.25	.43	.22	.41	.28	.45	-0.05***	-4.12
Baseline MAP reading	189.56	18.06	189.06	17.85	190.05	18.26	0.85	-1.04

Note. AIG = Academically or Intellectually Gifted (AIG) program. IEP = Individual Education Plan. SES = Socio-economic Status (at neighborhood level). MAP = Measure of Academic Progress.

[a]Mean (M) values are proportion estimates for categorical variables or raw means for continuous variables.

[b]Adaptive-Core treatment group differences are regression coefficients from multilevel models including the treatment indicator, school fixed effects, and random effects for teacher. p-values below 0.01 are significant when applying the Benjamini-Hochberg (BH) correction. Results for baseline MAP reading are based on multiple imputation for missing data.

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < .001$.

Table 2

Operationalization of the Core Components in Core and Adaptive Treatment and Acceptable Adaptations

Core components	Operationalization in Core Treatment	Operationalization in Adaptive Treatment	Acceptable adaptations
Asynchronous activities: • digital app • print-based trifolds	• Teachers introduce the app and trifolds to students and parents through video and letter. • Teachers access the app and trifolds and guide students in navigating features. • Teachers review app data reports to monitor student participation.	• Core Treatment PLUS • Teachers and literacy coaches engage in Team-Based Learning (TBL) sessions, meeting synchronously with the research team to set goals, envision student outcomes, and address challenges collaboratively. • Teachers co-develop strategies with peers to foster student engagement[a] Teachers receive app participation data to monitor progress and develop adaptive strategies during TBL sessions.	• Customize the strategies based on teachers' knowledge of individual student needs • Optimize communication and incentives to enhance student and family engagement • Utilize app data for targeted support, identifying students who need additional assistance
Synchronous online lessons	• Teachers receive basic training on the intervention curriculum and lesson materials. • Teachers deliver 15-day scripted lessons.	• Core Treatment PLUS • Teachers deliver a Day 16 extension lesson to deepen students' knowledge of the word *system*. • Teachers and literacy coaches at the same school meet synchronously with the research team to strategize how to teach the extension lesson to their students.	• Foster enjoyment and appreciation of language • Make changes to expand students' understanding of the word to more abstract contexts • Determine the presentation format to include interactive elements • Adjust lesson timing to meet student needs and school contexts (e.g., extending lessons over two days)

Note. [a]The strategies include modeling how to use, encouraging students to share their learnings, setting expectations, communicating with families, incentivizing, monitoring progress, and following up with students needing extra support.

Table 3

Type and Nature of Adaptations Completed by Adaptive Treatment Teachers

Component	Content Adaptations	Procedural Adaptations
Asynchronous activities	• App/trifold engagement ○ Introduced the purpose of using the app/trifolds and connected it to classroom content learning. ○ Highlighted student learning from the app/trifolds through projects, discussions, or presentations. ○ Emphasized the relevance and benefits of the app/trifolds to students and parents.	• App/trifold engagement ○ Modeled how to use the app during virtual class sessions. ○ Provided structured in-class time for app/trifold engagement. ○ Set clear expectations by integrating app/trifold use into routines or assignments. ○ Encouraged consistent use through regular communication with parents. ○ Incentivized usage with rewards. ○ Monitored engagement and offered individual follow-ups for encouragement and support.
Synchronous online lessons	• Extended lesson activities ○ Encouraged abstract thinking by connecting the concept of *systems* to classroom content and broader contexts. ○ Reinforced understanding of *systems* through mad libs, discussions, and presentations. ○ Provided word banks and visuals to support English learners.	• Extended lesson activities ○ Set clear expectations and modeled an example presentation to demonstrate expectations. ○ Guided students in completing and refining their *system* mad libs. ○ Facilitated peer sharing and engagement through structured opportunities. ○ Encouraged the use of multimedia and creative elements in presentations.

Table 4

Descriptive Statistics and the Effects of Adaptive Treatment on Student Engagement with the Asynchronous Learning Activities

Variable	Adaptive treatment			Core treatment			Adaptive treatment effect	
	M	*SD*	*n*	M	*SD*	*n*	β	*SE*
Engagement with game-based app activities								
Ever access app library (1 = yes, 0 = no)	0.65	0.48	942	0.60	0.49	972	0.05*	0.03
Number of science books completed on app library	0.82	1.60	942	0.47	1.16	972	0.29***	0.08
Number of target words	2.12	1.79	942	1.84	1.74	972	0.24*	0.10
Total time spent on app activities[a]	1.72	1.77	942	1.54	1.73	972	0.17†	0.10
Overall app activity accuracy[b] (std)	0.02	1.03	543	0.10	1.00	516	-0.02	0.05
Engagement with app activities aligned with synchronous curriculum lessons								
Ever access curriculum lessons	0.83	0.38	942	0.85	0.36	972	-0.002	0.02
Proportion of all curriculum lessons completed	0.57	0.40	942	0.59	0.39	972	0.01	0.02
Proportion of interactive read aloud lessons completed	0.59	0.39	942	0.62	0.39	972	0.01	0.02
Proportion of word sleuthing lessons completed	0.56	0.39	942	0.59	0.38	972	0.01	0.02
Accuracy on curriculum activities (std)	-0.01	1.02	774	0.06	1.00	820	0.001	0.04
Accuracy on end-of-unit quizzes (std)	0.00	1.01	573	0.02	1.01	603	0.03	0.05
Total time spent on curriculum lessons[b]	4.15	2.10	942	4.27	2.01	972	0.01	0.14
Perceived motivation and task challenges with app activities								
Enjoyment of app activities	2.92	0.69	932	2.86	0.71	966	0.10*	0.05
Reader self-competence beliefs	2.19	0.55	932	2.16	0.55	966	0.11*	0.05
How difficult was the task	2.01	0.39	932	2.01	0.38	966	-0.01	0.05
Print-based trifolds activities								
Return any trifolds (1 = yes, 0 = no)	0.11	0.31	942	0.09	0.29	972	0.01	0.01
Number of trifolds returned	0.97	3.22	942	0.78	2.85	972	0.17	0.14

Note. M = mean, SD = standard deviation. Point estimates for the Adaptive Treatment effect derived from multilevel models including the Adaptive Treatment indicator, school fixed effects, teacher random effects, student demographics, and baseline Measure of Academic Progress (MAP) reading scores and therefore differ from the raw difference in means.

[a]Log-transformed minutes. We add 1 to the total time spent on curriculum before taking the log.

[b]Sample sizes for app activity accuracy are reduced because we only have accuracy data for students who logged onto the app.

†$p < 0.10$, *$p < 0.05$, ***$p < 0.001$.

Table 5

Descriptive Statistics and the Effects of Adaptive Treatment on Student Learning Outcomes

Variable	Adaptive treatment			Core treatment			Adaptive treatment effect		
	M	*SD*	*n*	M	*SD*	*n*	β	*SE*	*SD* (across school sites)
Science vocabulary knowledge depth	20.02	7.92	942	20.51	7.70	972	0.02	0.03	0.01
Science background knowledge	4.56	2.06	942	4.53	2.05	972	0.09*	0.04	0.00
Reading comprehension									
Science content reading comprehension									
All passages	13.00	6.53	942	13.17	6.50	972	0.07*	0.03	0.00
Near-transfer passage	4.96	2.68	942	4.91	2.61	972	0.11**	0.04	0.08
Mid-transfer passage	4.29	2.35	913	4.36	2.35	950	0.05	0.04	0.00
Far-transfer passage	4.14	2.28	882	4.25	2.34	917	0.03	0.04	0.00
Domain-general reading comprehension									
MAP	191.87	18.76	902	193.01	18.19	915	0.004	0.03	0.00
EOG reading	433.74	9.55	869	434.45	9.94	899	-0.02	0.03	0.01

Note. M = mean, SD = standard deviation. MAP = Measure of Academic Progress. EOG = End-of-grade. The SD of treatment effects across school sites is derived from a random slopes multilevel model that allows the Adaptive Treatment effect to vary across schools.
**p* < 0.05, ***p* < 0.01.

Table 6

*Fidelity of Implementation: Differentiation in Teacher Strategies for Promoting Asynchronous Activities in Adaptive and Core Treatment Conditions (*N *= 95)*

Variable	Adaptive treatment (n = 48)		Core treatment (n = 47)		Adaptive treatment effect	
	M	*SD*	M	*SD*	β	*SE*
Digital app activities						
Modeled use of the app library	4.11	1.18	3.91	1.19	0.27	0.20
Encouraged students to complete app activities	4.32	1.04	3.94	1.22	0.37†	0.20
Set expectations books or minutes	3.45	1.41	3.47	1.23	0.08	0.23
Set aside instructional time to work on the app	3.70	1.44	3.72	1.14	0.11	0.22
Communicated with families about app	3.59	1.39	3.55	1.23	0.08	0.22
Provided opportunities to share experience with the app	3.26	1.42	3.09	1.21	0.19	0.24
Followed-up with students not completing activities	3.74	1.22	3.32	1.14	0.55**	0.20
Print-based books/trifolds activities						
Modeled use of the trifolds	2.35	1.45	1.83	1.24	0.48*	0.21
Set expectations for the number of trifolds	2.40	1.40	1.81	1.23	0.55*	0.22
Set aside instructional time to work on the trifolds	2.21	1.50	1.89	1.36	0.30	0.22
Communicated with families about trifolds	2.49	1.38	1.91	1.30	0.53*	0.22
Provided opportunities to share experience with trifolds	2.17	1.39	1.91	1.24	0.31	0.22
Followed-up with students not completing activities	2.06	1.46	1.83	1.24	0.29	0.21

Note. †$p < 0.10$, *$p < 0.05$, **$p < 0.01$.

Table 7

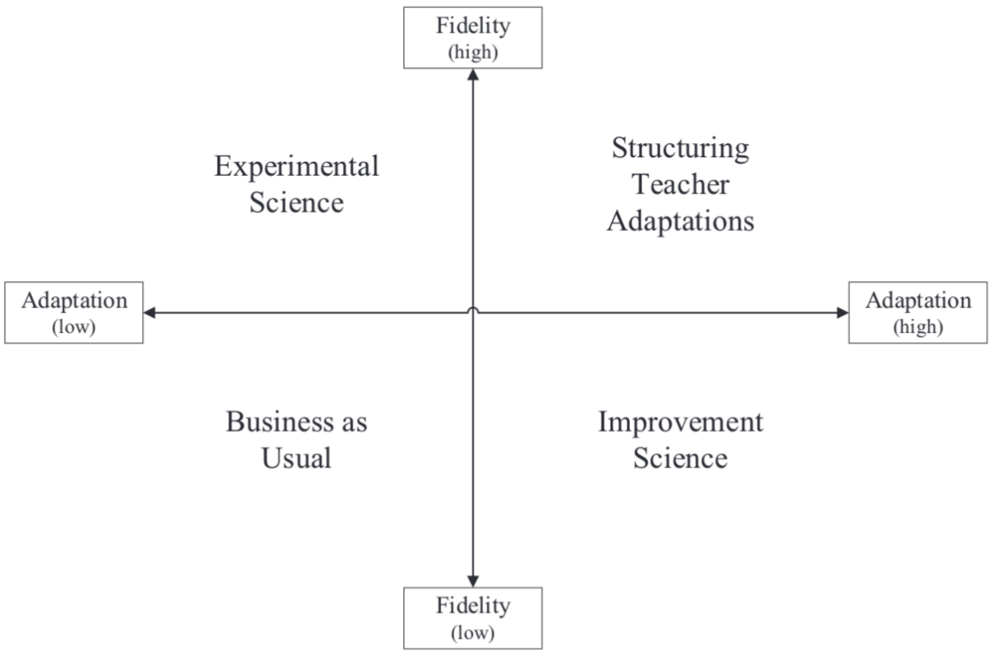
Fidelity of Implementation: Responsiveness Reflected in Student-Teacher Interaction in Adaptive and Core Treatment Conditions

Student-teacher interaction dimensions	Adaptive treatment		Core treatment		Adaptive treatment effect	
	M	*SD*	M	*SD*	β	*SE*
Engagement	2.95	0.67	2.25	0.60	0.77**	0.22
Feedback	3.46	0.45	2.71	0.68	0.80***	0.17
Questioning	3.07	0.74	2.26	0.40	1.00***	0.20

**$p < 0.01$, ***$p < 0.001$.

Figure 1

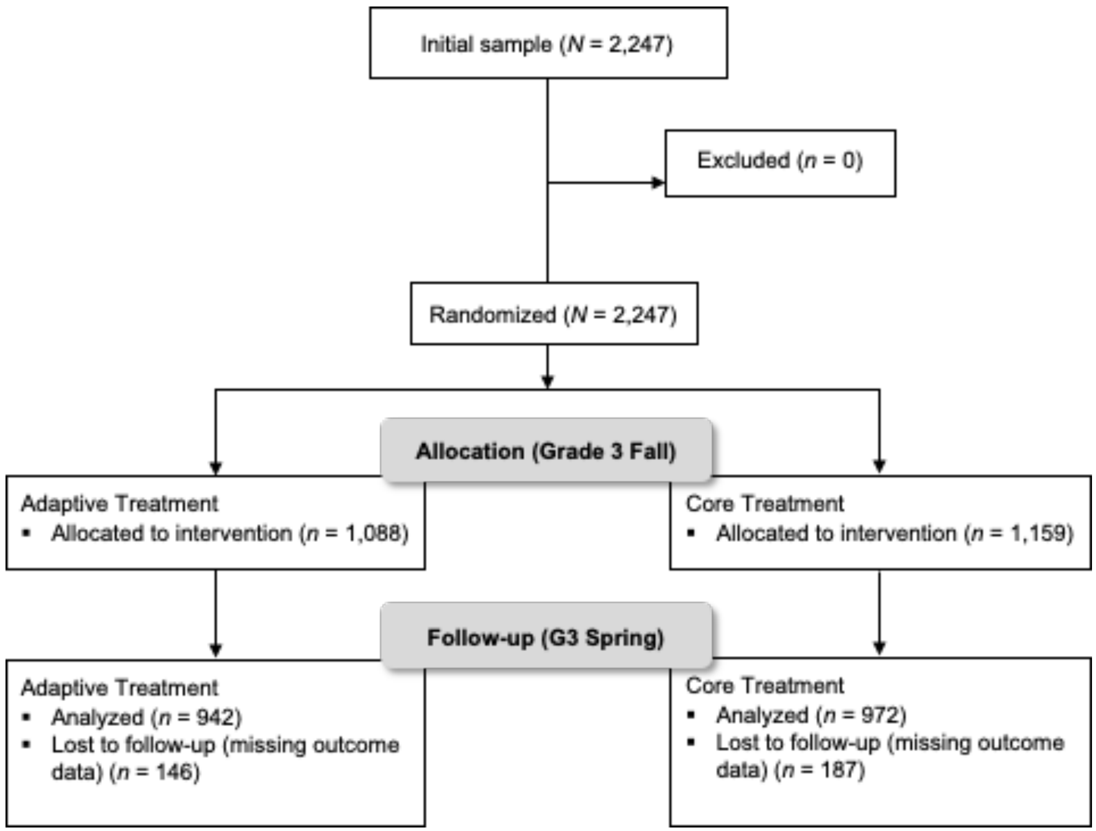
A Quadrant Framework for Teacher Adaptations and Fidelity in Interventions



Note. Source from [BLINDED FOR PEER REVIEW]

Figure 2

Consort Diagram for the Randomization Process



References

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 255–291). New York: Longman, Inc.

Archambault, L., Leary, H., & Rice, K. (2022). Pillars of online pedagogy: A framework for teaching in online learning environments. *Educational Psychologist, 57*(3), 178–191. https://doi.org/10.1080/00461520.2022.2051513

Bernacki, M. L. (2025). Leveraging learning theory and analytics to produce grounded, innovative, data-driven, equitable improvements to teaching and learning. *Journal of Educational Psychology, 117*(1), 1–11. https://doi.org/10.1037/edu0000933

Bond, M. (2020). Facilitating student engagement through the flipped learning approach in K-12: A systematic review. *Computers & Education, 151,* 103819. https://doi.org/10.1016/j.compedu.2020.103819

Brownell, M. T., Adams, A., Sindelar, P., Waldron, N., & Vanhover, S. (2006). Learning from collaboration: The role of teacher qualities. *Exceptional Children*, *72*(2), 169-185. https://doi.org/10.1177/001440290607200203

Bryk, A. S. (2015). 2014 AERA distinguished lecture: Accelerating how we learn to improve. *Educational Researcher, 44*(9), 467-477. https://doi.org/10.3102/0013189X15621543

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better.* Cambridge, MA: Harvard Education Press.

Cavazos, J. T., Hakala, C. M., Schiff, W. B., White, J. A., & Baskin, H. M. (2024). Flexible teaching during a pandemic and beyond: A reflection on lessons learned from the society for the teaching of psychology's pivot teaching committee. *Scholarship of Teaching and Learning in Psychology, 10*(4), 626–641. https://doi.org/10.1037/stl0000342

Clark, M., Merrick, L., Styron, J., Dorius, C., Bender, H., Johnson, J., Chapman, J., Gillette, M., Dorneich, M., O'Dwyer, B., Grogan, J., Brown, T., Leonard, B., Rongerude, J., & Winter, L. (2018). *Off to on: Best practices for online team-based learning.* Center for Excellence in Learning and Teaching Publications.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9-13. https://doi.org/10.3102/0013189X032001009

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12. https://doi.org/10.3102/0013189X032006003

Connor, C. M., Dombek, J., Crowe, E. C., Spencer, M., Tighe, E. L., Coffinger, S., Zargar, E., Wood, T., & Petscher, Y. (2017). Acquiring science and social studies knowledge in kindergarten through fourth grade: Conceptualization, design, implementation, and efficacy testing of content-area literacy instruction (CALI). *Journal of Educational Psychology, 109*(3), 301–320. https://doi.org/10.1037/edu0000128

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23-45. http://dx.doi.org/10.1016/S0272-7358(97)00043-3

Darling-Hammond, L., & Cook-Harvey, C. M. (2018). *Educating the whole child: Improving school climate to support student success.* Palo Alto, CA: Learning Policy Institute. https://doi.org/10.54300/145.655.

Datnow, A., & Castellano, M. (2000). Teachers' responses to success for all: How beliefs, experiences, and adaptations shape implementation. *American Educational Research Journal*, *37*(3), 775-799. https://doi.org/10.3102/00028312037003775

Dede, C. (2006). Scaling up: Evolving innovations beyond ideal settings to challenging contexts of practice. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 551-566). Cambridge University Press.

Dobbs, C. L., Ippolito, J., & Charner-Laird, M. (2017). Scaling up professional learning: Technical expectations and adaptive challenges. *Professional Development in Education*, *43*(5), 729-748. https://doi.org/10.1080/19415257.2016.1238834

Domitrovich, C.E., Gest, S.D., Jones, D., Gill, S., & DeRousie, R.M.S. (2010). Implementation quality: Lessons learned in the context of the Head Start REDI trial. *Early Childhood Research Quarterly, 25*(3), 284–298. https://doi.org/10.1016/j.ecresq.2010.04.001

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3–4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237–256. https://doi.org/10.1093/her/18.2.237

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*(1), 1–26. https://doi.org/10.17763/haer.66.1.g73266758j348t33

Francis, D. J., Kulesz, P. A., & Benoit, J. S. (2018). Extending the simple view of reading to account for variation within readers and across texts: The Complete View of Reading (CVRi). *Remedial and Special Education, 39*(5), 274–288. https://doi.org/10.1177/0741932518772904

Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N., & Porter, S. (2011). Focus, fiddle, and friends: Experiences that transform knowledge for the implementation of innovations. *Sociology of Education, 84*(2), 137–156. https://doi.org/10.1177/0038040711401812

Gabriel, R., & López, F. (2024). The role of asset-based pedagogy in an interactive view of reading. *Educational Psychologist, 59*(4), 233–249. https://doi.org/10.1080/00461520.2024.2394031

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6

Goldenberg, C., & Gallimore, R. (1991). Local knowledge, research knowledge, and educational change: A case study of early Spanish reading improvement. *Educational Researcher, 20*(8), 2–14. https://doi.org/10.3102/0013189X020008002

Gray, D. L., Hope, E. C., & Matthews, J. S. (2018). Black and belonging at school: A case for interpersonal, instructional, and institutional opportunity structures. *Educational Psychologist, 53*(2), 97-113. https://doi.org/10.1080/00461520.2017.1421466

Guthrie, J. T., Klauda, S. L., & Ho, A. N. (2013). Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly, 48*(1), 9–26. https://doi.org/10.1002/rrq.035

Guthrie, J. T., McRae, A., & Klauda, S. L. (2007). Contributions of concept-oriented reading instruction to knowledge about interventions for motivations in reading. *Educational Psychologist, 42*(4), 237–250. https://doi.org/10.1080/00461520701621087

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*(2), 141–164. https://doi.org/10.2307/1164207

Hattan, C., & Kendeou, P. (2024). Expanding the science of reading: Contributions from educational psychology. *Educational Psychologist, 59*(4), 217–232. https://doi.org/10.1080/00461520.2024.2418048

Hill, H. C., & Erickson, A. (2019). Using implementation fidelity to aid in interpreting program impacts: A brief review. *Educational Researcher, 48*(9), 590-598. https://doi.org/10.3102/0013189X19891436

Hwang, H., Cabell, S. Q., & Joyner, R. E. (2022). Effects of integrated literacy and content-area instruction on vocabulary and comprehension in the elementary years: A meta-analysis. *Scientific Studies of Reading, 26*(3), 223–249. https://doi.org/10.1080/10888438.2021.1954005

Joyce, B., & Showers, B. (1980). Improving in-service training: The messages of research. *Educational Leadership, 37*(5), 379-385.

Joyce, K. E., & Cartwright, N. (2020). Bridging the gap between research and practice: Predicting what will work locally. *American Educational Research Journal, 57*(3), 1045-1082. https://doi.org/10.3102/0002831219866687

Kaderavek, J. N., & Justice, L. M. (2010). Fidelity: an essential component of evidence-based practice in speech-language pathology. *American Journal of Speech-language Pathology*, *19*(4), 369–379. https://doi.org/10.1044/1058-0360(2010/09-0097)

Kintsch, W. (1993). Information accretion and reduction in text processing: Inferences. *Discourse Processes, 16*(1–2), 193–202. https://doi.org/10.1080/01638539309544837

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Kuhfeld, M., Soland, J., & Lewis, K. (2022). Test score patterns across three COVID-19-impacted school years. *Educational Researcher, 51*(7), 500-506. https://doi.org/10.3102/0013189X221109178

Lemons, C. J., Fuchs, D., Gilbert, J. K., & Fuchs, L. S. (2014). Evidence-based practices in a changing world reconsidering the counterfactual in education research. *Educational Researcher, 43*, 242–252. https://doi.org/10.3102/0013189X14539189

Leupen, S. (2020). Team-based learning in STEM and the health sciences. In J. J. Mintzes & E. M. Walter (Eds.), *Active learning in college science: The case for evidence-based practice* (pp. 219–232). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-33600-4_15

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54-61. https://doi.org/10.3102/0013189X15570388

López, F. (2024). Asset-based pedagogy. In P. A. Schutz & K. R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 433–457). Routledge.

Lucas, M., Nelson, J., & Sims, D. (2020). *Schools' responses to Covid-19: Pupil engagement in remote learning.* National Foundation for Educational Research.

Maniates, H. (2017). Teacher adaptations to a core reading program: Increasing access to curriculum for elementary students in urban classrooms. *Literacy Research and Instruction*, *56*(1), 68–84. https://doi.org/10.1080/19388071.2016.1210706

Martin, N.M., & Duke, N.K. (2010). Interventions to enhance informational text comprehension. In A. McGill-Franzen & R.L. Allington (Eds.), *Handbook of reading disability research* (pp. 345–361). London, UK: Routledge.

Martinez, L. R., Fishstrom, S., Vaughn, S., Capin, P., Carlson, C. D., Andress, T. T., & Francis, D. J. (2024). Supporting knowledge and language acquisition of secondary emergent bilinguals through social studies instruction. *Reading Research Quarterly, 59*(3), 349-370. https://doi.org/10.1002/rrq.541

McDonald, S., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher, 35*(3), 15–24. https://doi.org/10.3102/0013189X035003015

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1–43. https://doi.org/10.1207/s1532690xci1401_1

Michaelsen, L. K., & Sweet, M. (2011). Team-based learning. *New Directions for Teaching and Learning*, 2011(128), 41-51. https://doi.org/10.1002/tl.467

Neugebauer, S. R., Sandilos, L., DiPerna, J., Hunter, L., Hart, S. C., & Ellis, E. (2023). Exploring teacher implementation of a universal social-emotional learning program under routine conditions. *The Elementary School Journal*, *124*(1), 157-192. https://doi.org/10.1086/725675

Neuman, S. B., Samudra, P., & Danielson, K. (2021). Effectiveness of scaling up a vocabulary intervention for low-income children, pre-K through first grade. *The Elementary School Journal*, *121*(3), 385-409. https://doi.org/10.1086/712492

NWEA. (2019). *MAP® growth™ technical report*. https://www.nwea.org/uploads/2021/11/MAP-Growth-Technical-Report-2019_NWEA.pdf

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research, 78*(1), 33–84. https://doi.org/10.3102/0034654307313793

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184–202. https://doi.org/10.1598/RRQ.40.2.3

Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research, 88*(2), 205–242. https://doi.org/10.3102/0034654317743198

Patall, E. A., Steingut, R. R., Vasquez, A. C., Trimble, S. S., Pituch, K. A., & Freeman, J. L. (2018). Daily autonomy supporting or thwarting and students' motivation and engagement in the high school science classroom. *Journal of Educational Psychology, 110*(2), 269–288. https://doi.org/10.1037/edu0000214

Power, T. J., Blom-Hoffman, J., Clarke, A. T., Riley-Tillman, T. C., Kelleher, C., & Manz, P. H. (2005). Reconceptualizing intervention integrity: A partnership-based framework for linking research with practice. *Psychology in the Schools, 42*, 495–507. http://dx.doi.org/10.1002/pits.20087

Pugach, M. C., & Johnson, L. J. (2002). *Collaborative practitioners, collaborative schools* (2nd ed.). Denver: Love Publishing.

Raudenbush, S. W., Fotiu, R. P., & Cheong, Y. F. (1998). Inequality of access to educational resources: A national report card for eighth-grade math. *Educational Evaluation and Policy Analysis, 20*(4), 253-267. https://doi.org/10.3102/01623737020004253

Reeve, J., Ryan, R. M., Cheon, S. H., Matos, L., & Kaplan, H. (2022). *Supporting Students' Motivation: Strategies for Success.* Taylor & Francis.

Roossien, L., Boerboom, T., Spaai, G., van Klaveren, L. M., Dolmans, D., & de Vos, R. (2023). Opening the black box of team-based learning (TBL): A study of verbal interactions in online application sessions. *Medical Teacher, 46*(6), 832-841. https://doi.org/10.1080/0142159X.2023.2285249

Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. *Review of Research in Education, 16*, 353–389. https://doi.org/10.3102/0091732X016001353

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sánchez, V., Steckler, A., Nitirat, P., Hallfors, D., Cho, H., & Brodish, P. (2007). Fidelity of implementation in a treatment effectiveness trial of Reconnecting Youth. *Health Education Research, 22*, 95–107. http://dx.doi.org/10.1093/her/cyl052

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, 38, 445–459.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review, 13*(1), 23–52. https://doi.org/10.1023/A:1009004801455

Sherin, M. G., & Drake, C. (2009). Curriculum strategy framework: Investigating patterns in teachers' use of a reform-based elementary mathematics curriculum. *Journal of Curriculum Studies, 41*(4), 467–500. https://doi.org/10.1080/00220270802696115

Snow, C. E., & Matthews, T. J. (2016). Reading and language in the early grades. *The Future of Children*, *26*(2), 57-74. https://dx.doi.org/10.1353/foc.2016.0012.

Spector, C. C. (1996). Children's comprehension of idioms in the context of humor. *Language, Speech, and Hearing Services in Schools, 27*(4), 307–313. https://doi.org/10.1044/0161-1461.2704.307

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*(1), 72–110. https://doi.org/10.3102/00346543056001072

StataCorp. (2021). *Stata Statistical Software: Release 17.* College Station, TX: StataCorp LLC.

Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., Yen, L., Fuchs, L. S., & Compton, D. L. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis, 30*(4), 368–388. https://doi.org/10.3102/0162373708322738

Thum Y. M., Kuhfeld M. (2020). *NWEA 2020 MAP growth norms for student and school achievement status and growth*. NWEA.

VanderWeele, T. J. (2017). Outcome-wide epidemiology. *Epidemiology, 28*(3), 399–402. https://doi.org/10.1097/EDE.0000000000000641

Vaughn, M., Parsons, S. A., & Massey, D. (2020). Aligning the science of reading with adaptive teaching. *Reading Research Quarterly, 55*, S299-S306. https://doi.org/10.1002/rrq.351

Yan, L., Martinez-Maldonado, R., Swiecki, Z., Zhao, L., Li, X., & Gašević, D. (2025). Dissecting the temporal dynamics of embodied collaborative learning using multimodal learning analytics. *Journal of Educational Psychology, 117*(1), 106–133. https://doi.org/10.1037/edu0000905

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Zipke, M. (2008). Teaching metalinguistic awareness and reading comprehension with riddles. *The Reading Teacher, 62*(2), 128–137. https://doi.org/10.1598/RT.62.2.4