Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 24-1001

Asset-Based Implementation of Structured Adaptations in an Online Third-Grade Content Literacy Intervention

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 presents challenges, particularly in adapting to new contexts while maintaining fidelity. Structured adaptations that integrate the strengths of experimental science (high fidelity) and improvement science (high adaptation) represent a novel design framework for supporting the equitable implementation of research-based practices and programs. This preregistered study examined the effectiveness of structured 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, the quality of synchronous student-teacher interactions during the Zoom-delivered lessons, and student learning outcomes during the COVID-19 school closures. Using a cluster randomized trial design, 95 teachers and their students in 26 elementary schools were randomly assigned to either a high-fidelity core treatment or a structured adaptation condition. In the latter, teachers participated in Team-Based Learning activities that tightly coupled knowledge acquisition and application, with a focus on improving student engagement. Students in the adaptations condition outperformed students in the core treatment condition on science reading comprehension (ES = .07) and science background knowledge (ES = .09). Implementation analyses of the synchronous lessons indicate that the structured adaptations also improved norms of social interaction between students and teachers, resulting in stronger engagement, better feedback, and dialogic questioning. These findings suggest the importance of intentionally building in opportunities for teachers to adapt instruction within a clear framework. When teachers have support to work together and combine researchbased strategies with classroom knowledge, they can more effectively engage students and improve learning outcomes.

VERSION: August 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). Asset-Based Implementation of Structured Adaptations in an Online Third-Grade Content Literacy Intervention. (EdWorkingPaper: 24-1001). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/ztcp-v322

Asset-Based Implementation of Structured Adaptations in an Online Third-Grade Content Literacy Intervention

Jackie E. Relyea (Corresponding Author)
Affiliation: College of Education, North Carolina State University
E-mail: jrelyea@ncsu.edu
ORCID: https://orcid.org/0000-0002-7560-7136

Joshua B. Gilbert
Affiliation: Harvard Graduate School of Education
E-mail: joshua_gilbert@g.harvard.edu
ORCID: https://orcid.org/0000-0003-3496-2710

Mary Burkhauser
Affiliation: Harvard Graduate School of Education
E-mail: mary_burkhauser@gse.harvard.edu
ORCID: https://orcid.org/0000-0001-9142-7269

Ethan Scherer
Affiliation: Harvard Graduate School of Education
E-mail: mary_burkhauser@gse.harvard.edu
ORCID: https://orcid.org/0000-0002-6548-7724

Douglas Mosher
Affiliation: Harvard Graduate School of Education
E-mail: douglas_mosher@gse.harvard.edu
ORCID: https://orcid.org/0000-0001-5722-9218

Zhongyu Wei
Affiliation: Harvard Graduate School of Education
E-mail: zhongyu_wei@g.harvard.edu
ORCID: https://orcid.org/0009-0006-2704-0581

Johanna Tvedt
Affiliation: Harvard Graduate School of Education
E-mail: johanna_tvedt@gse.harvard.edu
ORCID: https://orcid.org/0009-0005-7003-8842

James S. Kim
Affiliation: Harvard Graduate School of Education
E-mail: james_kim@gse.harvard.edu
ORCID: https://orcid.org/0000-0002-6415-5496

Abstract

Scaling up evidence-based educational interventions presents challenges, particularly in adapting to new contexts while maintaining fidelity. Structured adaptations that integrate the strengths of experimental science (high fidelity) and improvement science (high adaptation) represent a novel design framework for supporting the equitable implementation of research-based practices and programs. This preregistered study examined the effectiveness of structured 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, the quality of synchronous student-teacher interactions during the Zoom-delivered lessons, and student learning outcomes during the COVID-19 school closures. Using a cluster randomized trial design, 95 teachers and their students in 26 elementary schools were randomly assigned to either a high-fidelity core treatment or a structured adaptation condition. In the latter, teachers participated in Team-Based Learning activities that tightly coupled knowledge acquisition and application, with a focus on improving student engagement. Students in the adaptations condition outperformed students in the core treatment condition on science reading comprehension (ES = .07) and science background knowledge (ES = .09). Implementation analyses of the synchronous lessons indicate that the structured adaptations also improved norms of social interaction between students and teachers, resulting in stronger engagement, better feedback, and dialogic questioning. These findings suggest the importance of intentionally building in opportunities for teachers to adapt instruction within a clear framework. When teachers have support to work together and combine research-based strategies with classroom knowledge, they can more effectively engage students and improve learning outcomes.

Asset-Based Implementation of Structured Adaptations in an Online Third-Grade Content Literacy Intervention

Efforts to bring evidence-based literacy interventions into everyday classrooms have long faced a set of practical and conceptual challenges (Dede, 2006; McDonald et al., 2006; Stein et al., 2008). Although many of these programs yield strong results in controlled studies, their effects tend to be less predictable when implemented across diverse school settings (J. Kim, 2019). The gap between research and practice is not merely a matter of implementation fidelity but reflects real tensions between idealized program conditions and the variability of classroom life (Joyce & Cartwright, 2020). Thus, scaling is not only a technical challenge but also a matter of educational equity, particularly when instructional models assume standard conditions and narrowly define fidelity in ways that constrain teacher responsiveness to students' needs (O'Donnell, 2008). Addressing this challenge requires implementation designs that preserve the conceptual integrity of research-based instruction while also enabling instructional decisions that reflect local priorities and classroom-level variability.

Educational equity has been conceptualized in varied ways across research and policy, reflecting different assumptions about fairness, opportunity, and learning (Levinson et al., 2022). In this study, we define equity as the intentional design of instructional opportunities that address structural and instructional barriers, enabling teachers to make principled adaptations responsive to students' cultural, linguistic, and contextual assets (Gay, 2018; Ladson-Billings, 2014). This framing recognizes that equitable outcomes depend on when teachers can exercise professional judgment in a coherent instructional model. By treating student variability as an asset rather than a constraint, this conception of equity informed our hypothesis that teachers empowered to adapt instruction systematically would be better positioned to foster students' engagement and learning.

While equity-oriented instruction calls for contextual responsiveness, fidelity has, historically, been prioritized as the primary mechanism for scaling, with effectiveness tied to strict adherence to standardized protocols (Dane & Schneider, 1998; Dusenbury et al., 2003; Kaderavek & Justice, 2010). Fidelity-driven approaches often rest on the assumption that if teachers closely follow a program's prescribed steps, outcomes will follow. But this logic rarely holds up in classrooms with varied student needs, shifting priorities, and instructional constraints (Cobb et al., 2003; Quinn & J. Kim, 2017). In contrast, adaptation offers a more responsive stance, allowing teachers to adjust interventions based on the realities of their settings (Datnow & Castellano, 2000; Parsons et al., 2018). Research in improvement science highlights the role of adaptation in fostering teacher agency and ownership, both of which are essential for sustainability (Lewis, 2015). Still, when adaptations are made without clear guidance or boundaries, they risk fragmenting the intervention and undermining its core logic. Building on prior research on teacher adaptations to curricula (e.g., Brown, 2009; J. Kim et al., 2017; Vaughn et al., 2020), this study examines how implementation models can structure teacher agency in ways that maintain program coherence while allowing for responsiveness to local context.

One approach to designing this balance is through structured adaptations, a design framework that supports teacher decision-making within defined instructional boundaries (J. Kim & Mosher, 2023). This framework positions teachers to exercise professional judgment in making deliberate adaptations, maintaining intervention fidelity while deepening both teacher and student engagement to strengthen learning outcomes. Structured adaptations are not intended to resolve the tension between fidelity and adaptation; instead, they reframe it. The central question becomes how teachers interpret and enact core instructional principles in ways that are responsive to their instructional context (Cohen et al., 2025; Quinn & J. Kim, 2017). While researchers may offer conceptual guidance and tools, teachers determine how instruction unfolds, adjusting pacing, scaffolds, and routines to support participation and achieve learning goals (J. Kim & Mosher, 2023). Fidelity, in this sense, is understood as alignment with the conceptual aims of the intervention rather than adherence to fixed procedures. Teachers apply their pedagogical content knowledge (Shulman, 1986) and contextual expertise to modify delivery while preserving the integrity of the intervention. These adjustments support accessibility and relevance for their students (Bryk et al., 2015; Durlak & DuPre, 2008). Structured adaptations reflect a design stance that positions teacher agency as essential to the implementation of research-based instruction across varied contexts.

Building on this framework, this study examined the impact of structured 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., J. Kim et al., 2017; Quinn & J. Kim, 2017), 95 classrooms were randomly assigned to one of two conditions. In the *Core Treatment* condition, teachers implemented the content literacy intervention, called Model of Reading Engagement (MORE), 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 J. Kim et al. (2017) 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.

Structured Adaptation as an Asset-Based Implementation Design

Fidelity and adaptation inevitably coexist in efforts to bridge education research and classroom practice (Century & Cassata, 2016; Hill & Erickson, 2019). The critical question, then, is not whether to prioritize one over the other but how implementation designs can create conditions that support principled decision-making while preserving the integrity of core instructional goals. Asset-based pedagogy speaks directly to this challenge by centering the cultural, experiential, and professional assets that both students and teachers bring to learning environments (López, 2024). It recognizes their local knowledge, professional judgment, and contextual awareness as essential to high-quality implementation (Gabriel & López, 2024). As López (2024) argues, asset-based pedagogy invites us to see variation not as a threat to consistency but as a resource to leverage for responsive teaching. Yet much of the current science of reading discourse continues to reflect an instrumental model of knowledge use, framing research as a set of prescriptions to be transferred into practice, with little attention to how local adaptation mediates implementation (Hattan & Kendeou, 2024).

The Reading First initiative, the “academic cornerstone” of the No Child Left Behind Act (2008), illustrates the limitations of this transfer model. Intended to translate the National Reading Panel (2000) meta-analytic findings into prescriptive definitions of scientifically based reading practices, Reading First prioritized fidelity as a central mechanism of reform. However, large-scale evaluations revealed that adaptation inevitably occurred, resulting in substantial variation in how practices were implemented across classrooms (Correnti & Rowan, 2007; Gamse et al., 2008). Ethnographic studies further demonstrated that adaptations were not merely deviations but often reflected teachers' efforts to reconcile externally mandated practices with their students' needs and instructional contexts (Coburn & Woulfin, 2012). Some coaches cultivated autonomy-supportive environments that enabled teachers to integrate core principles meaningfully, while others enforced compliance-driven environments that constrained teacher agency and instructional responsiveness (Gallucci, 2008). The uneven implementation of Reading First suggests that fidelity, absent attention to the relational and contextual dimensions of practice, is insufficient to achieve sustained instructional improvement. This history underscores the need for implementation designs that deliberately support structured adaptation rather than treating it as an undesirable byproduct. Structured adaptation, when grounded in asset-based principles, acknowledges the variability of instructional contexts and positions teachers as professional actors capable of interpreting and enacting core practices (Honig, 2006). Research on autonomy-supportive environments shows that contexts fostering agency, ownership, and professional discretion enhance psychological well-being, prosocial engagement, and sustained motivation (e.g., Collie & Martin, 2017; Reeve et al., 2020). In literacy intervention contexts, autonomy-supportive designs help teachers internalize the conceptual foundations of interventions, make principled instructional decisions, and develop adaptive expertise necessary for long-term reform (Parsons et al., 2018).

Ideally, developing adaptive expertise requires autonomy-supportive environments that cultivate both routine expertise—the consistent, high-quality enactment of research-based practices—and adaptive expertise, characterized by flexible, principle-driven reasoning (Mehta & Fine, 2019). As Coburn (2003) argues, external reforms depend not on superficial compliance but on achieving depth by penetrating classrooms, shifting norms of social interaction, and deepening teachers' understanding of foundational principles. Adaptive expertise requires not only the ability to execute prescribed practices but the professional judgment to discern why practices matter, under what conditions they are effective, and how they can be responsibly adapted without introducing lethal mutations (Datnow & Castellano, 2020). Cultivating this capacity demands implementation designs grounded in program-agnostic principles: transferable frameworks that support teachers in making contextually responsive decisions while maintaining coherence with core instructional aims. In this framework, structured adaptations are not an afterthought or a deviation, but an intentional, asset-based design stance that honors professional reasoning, local knowledge, and sustained engagement with research-based literacy practices.

Viewing Structured Adaptations Through the Lens of Self-Determination Theory

Drawing on self-determination theory (SDT; Deci & Ryan, 1985, 2000), we conceptualize structured adaptations as a motivational design intended to cultivate both the *skill* and *will* teachers need to co-construct usable knowledge and engage in principled, context-responsive literacy instruction (J. Kim & Mosher, 2023). *Skill* refers to the capacity to enact and adapt research-based practices effectively; *will* refers to the internal motivation to take ownership of instructional improvement over time.

Structured adaptations underscore the idea that teachers need both general research knowledge and situated local knowledge to make interventions "work reliably over diverse contexts and populations" (Bryk, 2015, p. 469). Accordingly, the concept of structured adaptations calls for broadening the discourse from a narrow focus on the science of reading to the broader science of teaching reading well, particularly for students from varied economic, racial, linguistic, and ability backgrounds (Hattan & Kendeou, 2024; Y.-S. Kim & Snow, 2024). Structured adaptations bridge the known (general research-based knowledge) and the new (contextualized teacher knowledge generated through principled adaptation) in service of equitable literacy instruction that honors the learning of *all* students.

Viewed through the lens of SDT, structured adaptations are also intended to foster the motivational conditions necessary for teachers to persist, adapt, and sustain instructional improvement over time. SDT posits that individuals have three innate psychological needs—autonomy, competence, and relatedness—that must be satisfied for internalization and sustained engagement to occur (Deci & Ryan, 2000; Niemiec & Ryan, 2009). When these needs are neglected, teachers may comply superficially with externally mandated reforms without true ownership or lasting change. Consistent with SDT, we view teachers not as passive implementers but as active participants whose professional motivation can be cultivated or undermined by the design of implementation supports (Bryk et al., 2015; Kennedy, 2016).

The present intervention was designed not simply to deliver instructional content but to structure the motivational environment in which new practices are learned and enacted. Each core component was deliberately aligned with one of SDT psychological needs: structured adaptations fostered autonomy through bounded opportunities for professional judgment; the MORE curriculum provided affordances for competence with clear pedagogical structures and supports; and Team-Based Learning (TBL) created conditions for relatedness that promoted shared purpose and mutual accountability. This design logic positioned teachers not as passive implementers but as active participants in shaping instruction within systems that respect their professional agency, cultivate instructional efficacy, and sustain collegial connections.

Autonomy: Structured Adaptations as Bounded Professional Agency

In SDT, autonomy is defined as the experience of volition and psychological ownership over one's actions, not simply the presence of choice but the alignment of action with personal values and professional identity (Deci & Ryan, 2000; Ryan & Deci, 2020). Autonomy-supportive environments foster this alignment by offering choices within defined boundaries that respect individuals' expertise and goals (Su & Reeve, 2011). In contrast, implementation models that prioritize procedural fidelity often compromise this need, positioning teachers as passive implementers and minimizing space for professional judgment (Coburn, 2003; Spillane et al., 2002). The structured adaptation framework developed in this study was designed to support autonomy by embedding guided decision-making opportunities within a coherent, bounded instructional design (J. Kim & Mosher, 2023; Maniates, 2017). These opportunities preserved the theoretical integrity of the intervention, allowing principled variation in response to local classroom conditions.

In this study, teachers exercised agentive judgment through purposeful modifications, adjusting pacing, adapting interaction structures, and introducing scaffolds intended to support student access and engagement in remote learning. These decisions were not deviations from the intervention logic but local adaptations and enactments of disciplined inquiry in practice (Shulman, 1981). The structured adaptation framework treated teacher autonomy not as a liability to manage but as a fundamental psychological need essential to internalizing instructional practices and sustaining professional engagement (Deci & Ryan, 1994; Reeve et al., 2020). Within clearly defined instructional parameters, teachers were able to exercise agency in ways that reinforced their sense of competence and sustained motivation throughout implementation.

Competence: Curriculum Design to Support Instructional Agency

Instructional change is unlikely to take root unless teachers experience themselves as effective in navigating and enacting its demands (Cohen & Ball, 2001). In SDT, this sense of competence is not defined as task completion but as the perception of effectiveness in meeting personal or professional goals in structured environments (Deci & Ryan, 2000). Competence is fostered when individuals engage with systems that provide clear expectations, optimal challenges, and aligned tools for meaningful success (Connell & Wellborn, 1991; Niemiec & Ryan, 2009).

The MORE curriculum was deliberately designed as an instructional infrastructure to cultivate this condition. It offered a coherent pedagogical architecture, specifying pedagogical goals, sequencing cognitively demanding tasks, and embedding routines to support the enactment of content-rich literacy instruction (J. Kim, Burkhauser, et al., 2021; J. Kim, Relyea, et al., 2021; Mosher et al., 2024). Curriculum components, including conceptually coherent text sets in thematic units, interactive read-alouds, structured vocabulary routines, and guided discussions, established conceptual continuity and instructional coherence, enabling teachers to enact evidence-based literacy instruction with greater clarity and confidence.

Additionally, implementation tools such as trifolds and a digital app supported synchronous learning by reducing the cognitive and logistical demands of lesson preparation without diminishing rigor. These home-based learning resources externalized key instructional elements (e.g., texts, tasks, and scaffolds), which freed teachers to focus on student interaction and moment-to moment instructional decisions during live sessions. This design aimed to make the demands of content-rich literacy instruction more manageable in remote contexts, reinforcing teachers' sense of instructional efficacy. Instructional systems that support competence need to provide structure, aligning with the professional goals that teachers value (Ahmadi et al., 2023). When teachers perceive instructional demands as feasible and professionally worthwhile, they are more likely to engage deeply, persist through challenges, adapt instruction, and integrate complex, research-informed practices into daily teaching (Kennedy, 2016; Ryan & Deci, 2020).

Relatedness: Team-Based Learning as a Social Ecology for Professional Belonging

The construct of relatedness in SDT emphasizes that individuals internalize new practices more fully when they experience meaningful connection through relationships grounded in mutual respect, shared goals, and psychological trust (Ryan & Deci, 2020). In professional learning, this need is essential to how teachers come to see themselves as contributors to, rather than consumers of, instructional knowledge. The MORE intervention was designed to cultivate this condition by embedding TBL (Michaelsen & Sweet, 2011) as the core structure for collaboration. TBL did not function as an add-on to professional development; rather, it operated as a social ecology that organized professional relationships to support shared inquiry, context-sensitive adaptation, and collective instructional decision-making (Hord, 1997; Michaelsen & Sweet, 2011).

The collaborative architecture of TBL was enacted through intentionally sequenced routines that positioned teachers as instructional reasoners working in community with peers. Teachers participated in the Readiness Assurance Test—first individually (iRAT) and then in teams (tRAT)—to unpack the conceptual foundations of the MORE intervention and clarify its design logic (Roossien et al., 2023). These activities were not used to enforce compliance but to initiate inquiry, elicit uncertainty, and support adaptations grounded in local instructional contexts (Swanson et al., 2019). This participatory model aligns with asset-based pedagogies that view teachers' local knowledge as central to instructional improvement (López, 2024). It affirms professional agency as essential to decision-making and empowers teachers to integrate their professional and contextual expertise into research-based instruction in ways responsive to varied classroom demands (J. Kim & Mosher, 2023; Martinez et al., 2024). Through TBL, relatedness became instrumental rather than incidental: teachers held each other accountable, their classroom expertise mattered, and they collectively owned instructional improvement. These dynamics became especially salient during the COVID-19 pandemic when professional isolation and instability intensified the need for collegial structures that provided clarity and connection.

The function of relatedness in this intervention extended beyond collegial support to address the epistemic authority shaping how instructional reforms are understood and enacted (Bridwell-Mitchell, 2015). TBL created space for collaboration and redefined whose knowledge counts in the design of instructional practice. In contrast to top-down models that treat fidelity as compliance and position teachers as implementers of externally developed programs, the TBL structure affirmed their role as instructional agents whose situated expertise was critical to adaptation and design (Butler & Schnellert, 2012). Structured collaboration enabled teachers to examine the conceptual foundations of the intervention, clarify its pedagogical aims, and adjust implementation to align with local conditions while maintaining coherence with core principles (Cohen et al., 2025). Adaptation was framed not as deviation but as a form of instructional reasoning and ownership grounded in professional judgment (Parsons et al., 2018). In this design, relatedness served as an interpersonal affordance and a structured condition, enabling shared responsibility, pedagogical agency, and sustained engagement in the work of research-informed instructional improvement.

Why Structured Adaptations Matter to Content Literacy Instruction

Content literacy instruction involves building conceptual schema—mental frameworks that support the integration of new information with prior knowledge to promote comprehension and transfer (Anderson & Pearson, 1984; Kintsch & van Dijk, 1978). This process becomes crucial when students engage with informational texts in science, which often include abstract systems, discipline-specific language, and unfamiliar structures that require inferential reasoning (Martin & Duke, 2010). Learning from these texts depends on instructional supports that organize conceptual demands and make disciplinary thinking accessible through explicit language-based routines.

Empirical research has demonstrated the promise of content-integrated literacy instruction to improve students' language, reading comprehension, and engagement, especially when implemented through coherent, conceptually focused units (e.g., Cabell et al., 2025; Guthrie et al., 2007; J. Kim, Relyea et al., 2021; Mosher et al., 2024; Relyea et al., 2024; see review by Hwang et al., 2022). Yet, translating these benefits into equitable outcomes across varied instructional contexts requires more than strong curriculum materials. Teachers need tools and support to implement instruction in ways that account for variability in student engagement, classroom conditions, and available instructional time (Coburn, 2003; Durlak & DuPre, 2008). Structured adaptations address this need by guiding how instructional routines can be adjusted while maintaining alignment with core learning goals.

In this study, adaptation was designed as a bounded process, guided by team-based professional learning and supported through shared curricular materials. Teachers implemented the same science curriculum using pacing guides and lesson components. Within this structure, they made adjustments to support student engagement across asynchronous and synchronous modalities. In asynchronous contexts, where student motivation was more challenging to sustain, teachers integrated the literacy app and trifold into instructional routines by modeling their use during live sessions, monitored student progress, and incorporated student reflections into class discussion. These routines made student progress visible, promoted accountability, and created a sense of accomplishment that supported continued participation (Krechevsky et al., 2013). The app and trifold functioned as interactive instructional components used to deepen conceptual understanding while keeping students engaged.

In synchronous instruction, teachers in the Adaptive Treatment condition exercised guided autonomy to expand core lessons to support schema development around abstract scientific systems, such as the interdependence of organs and functions in the human body. These adaptations included student presentations, vocabulary games, and sentence-level activities (e.g., "Mad Libs" formats) that provided repeated, varied opportunities to work with disciplinary language (Snow & Uccelli, 2009). These extensions provided students with multiple, accessible opportunities to process and express disciplinary content, supporting deeper understanding and the application of core concepts across instructional contexts (Fries et al., 2021).

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 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 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 maintaining fidelity to core intervention 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 adaptations during the COVID-19 pandemic on third graders' engagement and learning outcomes. We addressed three research questions (RQs):

- **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? To what extent 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?

Focal outcome measures were selected to reflect the underlying theory of change guiding the intervention (J. Kim, Relyea, et al., 2021). Engagement was included as an indicator of the intervention capacity to sustain participation, which was especially critical in remote learning environments. Domain-specific vocabulary knowledge was assessed as students' ability to form and use networks of semantically related science words as a mechanism for schema development. The assessment of background knowledge captured students' conceptual understanding of science content. Reading comprehension was conceptualized as a measure of students' ability to transfer their vocabulary and background knowledge to make sense of science texts. Collectively, these outcomes provide a basis for testing whether structured adaptations move the needle on the core capacities that the Adaptive Treatment condition of the intervention aims to strengthen.

Methods

Study Design and Participants

This study employed a cluster (school) 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 13.9% in the Adaptive Treatment condition (n = 1,117 to 962) and 15.7% in the Core Treatment condition (n = 1,130 to 952), resulting in a final analytic sample of 1,914 students. A chi-squared test for differential attrition indicated no statistically significant differences between conditions ($p = .21$).

Table 1 presents student characteristics by treatment condition. The conditions were generally balanced across student characteristics and baseline reading test scores, with a few exceptions. Namely, the Adaptive Treatment group included a higher proportion of Hispanic students (38% vs. 29%, $p < .01$), a slightly greater representation of medium SES students (38% vs. 35%, $p < .01$), and fewer high SES students (22% 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: https://doi.org/10.7910/DVN/JDUIKT.

Intervention Treatment Conditions

Core and Adaptive Treatment

Both Core and Adaptive Treatment teachers implemented the core components of the intervention. Table 2 provides an overview of how these 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 the acceptable adaptations implemented by Adaptive Treatment teachers.

Both Core and Adaptive Treatment teachers participated in researcher-led professional development. In January 2021, all participating teachers attended a 60-minute synchronous online training session led by the research team on Zoom. To maintain an interactive and engaging format, each session was limited to 25 participants, with multiple sessions offered to accommodate teachers' availability. 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 a deeper understanding from complex texts. Teachers were introduced to 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).

In addition to the synchronous training, teachers received electronic lesson materials, a pacing calendar, and instructional videos to familiarize themselves with the student app and trifold activities. Additionally, 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 of the ebook option for students to follow along. These activities (see Table A1 in Appendix A for detailed descriptions and examples) included identifying sight words and syllables, understanding word parts (prefixes and suffixes), spelling with morphemes, constructing sentences, and enhancing reading comprehension with a focus on main ideas and textual content. The activities were also designed to develop metalinguistic awareness through humor, idioms, and metaphors and enhance their ability to analyze and interpret language nuances and deepen 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. The app activities were also made available to students' families in both English and Spanish versions.

Print-Based Books and Trifolds. 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, all teachers conducted 15 synchronous Zoom lessons (each lasting 30-35 minutes) over a five-week period 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 their 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 the 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 adaptations to the core components to enhance student engagement and learning outcomes. Table 3 shows 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 researcher-developed online modules. 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 school. 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 worked together to refine adaptation plans, set goals, and create monitoring strategies. This collaborative work enabled teachers to translate research evidence into practical approaches tailored to their specific classroom contexts, while maintaining alignment with intervention goals.

Teachers first examined obstacles to student engagement, particularly students' reluctance to complete independent activities and scheduling constraints within the existing curriculum. The researchers facilitated discussions where teachers contributed their classroom expertise toward developing action plans. To guide this process, the research team offered specific adaptive strategies for increasing student interaction with the app and trifold materials: demonstrating app functionality, creating opportunities for students to present their learning during synchronous sessions, establishing clear usage expectations, engaging families as partners, creating incentive structures, and using app analytics for usage monitoring. Teachers and literacy facilitators then selected from these options, adapting them through collaborative problem-solving to fit their particular school environments.

Moreover, Adaptive Treatment teachers implemented an optional 35-minute extension lesson on Day 16, which was not included in the Core Treatment condition. All teachers in the Adaptive Treatment condition opted to implement it; thus, there was no variation in whether the lesson was delivered, but variation in how it was enacted, reflecting the autonomy-supportive structure of the Adaptive Treatment condition. This flexible approach allowed teachers to adapt and extend instruction based on their students' needs, background knowledge, and instructional contexts.

The extension 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 provided implementation suggestions, such as using "Mad Libs" exercise 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) *ever 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 used 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 zero 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. Students read each passage independently; neither the texts nor the questions were read aloud. 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: the 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 scale, calibrated using item response theory, to reflect students' reading proficiency. Test-retest reliability coefficients for composite scores ranged from .79 to .86 (NWEA, 2019).

Statewide EOG Reading Assessment. The EOG Reading Assessment for Grade 3, administered in the spring, was used as a measure of students' domain-general reading comprehension. The assessment aligns with the state 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 (North Carolina Department of Public Instruction, 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. 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. Consistent with prior research on classroom discourse (Fredricks et al., 2004), 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 (Skinner & Pitzer, 2012). 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 (Alexander, 2008). Feedback focused on cognitive engagement through clarifications, connections to prior learning, and positive reinforcement, with higher scores reflecting more detailed and constructive responses (Hattie & Timperley, 2007).

We rated 153 synchronous lesson sessions in Lessons 3, 4, 7, and 8 of the 15-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 inter-rater reliability. The overall agreement was .91 (Cohen's $\kappa = .79$).

Data Analysis

To address RQs 1 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:

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

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

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

where Y_{ijk} denotes the outcome for student i in teacher j's classroom in school k, and α_k represents the intercept for school k. β_1 captures the covariate-adjusted intention-to-treat (ITT) effect of 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 standard deviation of treatment effects across school sites. To address missing covariate data, we used multiple imputation. 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:

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

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

where Y_{jk} denotes the outcome for teacher j in school k, and α_k represents the intercept for school k. The term β_1 captures the adjusted ITT effect of 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, Adaptive Treatment resulted in a higher proportion of students accessing the app library ($\beta = 0.06$, $p < .05$), completed more science books (e.g., groups of 15 activities) on the app (β = 0.31, $p < .001$), and led to more frequent exposure to target words (β = 0.27, $p < .05$). It also likely contributed to slightly more time spent on app activities (β = 0.17, $p < .10$), although this effect was marginal. The intervention, however, did not produce differences in overall accuracy on app activities (β = -0.003, $p > .05$).

In addition, the Adaptive Treatment led to higher levels of student-reported engagement, with students reporting greater enjoyment of lesson activities (β = 0.10, $p < .05$) and stronger self-competence beliefs (β = 0.10, $p < .05$). It did not produce differences in students' perceived task difficulty (β = -0.01, $p > .05$). For print-based book/trifold activities, Adaptive Treatment may have supported slightly more frequent return of trifolds (β = 0.16, $p > .05$), though this difference was not statistically significant. Similarly, the probability of returning any trifolds was comparable across groups (β = 0.01, $p > .05$). The intervention also did not produce differences in engagement during 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, Adaptive Treatment led to higher scores on science background knowledge compared to the Core Treatment group, with a small but statistically significant effect (ES = 0.09, $p < .05$). Students in the Adaptive Treatment group outperform those in the comparison group on science content reading comprehension across all passages (ES = 0.07, $p < .05$), demonstrating its positive impact on students' understanding of science texts. By passage type, Adaptive Treatment resulted in significantly higher performance on the near-transfer passage (ES = 0.12, $p < .01$), while it had no significant effect on the mid-transfer (ES = 0.07, $p < .10$) and far-transfer passages (ES = 0.02, $p > .05$). Adaptive Treatment did not affect domain-general reading comprehension, with no significant differences on MAP scores (ES = -0.02, $p > .05$) or EOG reading scores (ES = -0.03, $p > .05$).

We conducted a series of sensitivity checks on the ITT analysis by adding a covariate for the number of exposures to target vocabulary, which differed between conditions due to the 35-minute extension lesson in the Adaptive Treatment group. When controlling for word exposure, the estimated ES for science background knowledge in the Adaptive Treatment group decreased only slightly from 0.09 to 0.08. For science content reading comprehension, both overall and for each of the three passage types, the ES differences were negligible, all less than 0.01. This minimal change suggests that the additional lesson alone did not fully account for the observed treatment effects.

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 (e.g., 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 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 built on the Core Treatment by adding structured adaptations with TBL activities. The Adaptive Treatment condition aimed to increase teachers' perceptions of autonomy, competence, and relatedness to increase (a) their motivation to work with the content literacy intervention and (b) the fit of the intervention to meet teacher and student needs 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 condition 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 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 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 the 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, students in the Adaptive Treatment group demonstrated greater motivation and stronger alignment between task difficulty and interest compared to students in the Core Treatment group. These findings suggest that the structured adaptations developed through TBL activities significantly increased student participation in digital, game-based activities.

Notably, the observed differences in the quality of student-teacher interactions were not simply a function of the extension lesson but reflected qualitative differences in classroom discourse. Although both groups implemented the same core synchronous lessons, analyses of classroom audio recordings (Table 7) showed that teachers in the Adaptive Treatment condition facilitated more dialogic instruction, offered more substantive feedback, and elicited greater student participation. These findings indicate that Adaptive Treatment influenced the quality, not just the quantity, of instructional engagement. Rather than being an artifact of added instructional time, the differences in the quality of student-teacher interactions reflect how the adaptation model supported bounded flexibility and professional agency, in which teachers made principled adjustments to promote students' deeper cognitive and social engagement. This interpretation aligns with views of scale that emphasize shifts in classroom interaction norms as markers of instructionally responsive and equitable implementation (Coburn, 2003). By equitable implementation, we mean instructional enactment in which teachers use both research-based guidance and local professional judgment to make structured adaptations that promote access and engagement for all learners in their instructional context (Levinson et al., 2022; OECD, 2018).

Furthermore, the findings suggest that structured adaptations, supported through TBL sessions, enabled teachers to design scaffolding routines that remained anchored in core instructional goals while also responsive to students' sense-making of science content. Rather than adhering to scripts, teachers treated shared strategies as starting points and adjusted them based on student thinking and classroom dynamics. This flexibility allowed teachers to draw on students' prior knowledge, interests, and lived experiences, connecting unit topics to familiar cultural or community contexts. When instruction was grounded in students' experiences, their contributions became more visible in classroom discourse, fostering a stronger sense of competence and belonging (Gray et al., 2018).

These patterns reflect our conceptualization of equity in practice. Adaptive Treatment teachers were not just permitted to adapt; they were given structured opportunities to do so in more systematic ways that acknowledged students' varied learning contexts and created accessible entry points into core content. For instance, they built class time for app work into their lessons, celebrated incremental progress, and drew on families' experiences with health and medicine to extend science conversations. In doing so, they enacted asset-based pedagogy, not by simplifying content, but by using knowledge of students' lives to strengthen its relevance and reach (López, 2024).

Although these adaptations in this study did not directly leverage students' full cultural or linguistic resources, they reflected an asset-based stance by centering teachers' knowledge of their students and contexts as essential for equitable access. Unlike general guidance for differentiation in many curriculum programs, the structured adaptation model in this study offered a collaborative design infrastructure, operationalized through TBL activities, that supported teachers in making context-sensitive decisions while maintaining coherence with shared goals and instructional principles (Bryk et al., 2015; Penuel et al., 2017). While this study focused on adaptations to support student engagement, the model is not limited to this domain. Student engagement was prioritized given its central importance during remote learning and its links to student motivation, learning, and equity (Fredricks et al., 2004; Skinner & Pitzer, 2012).

Student engagement across asynchronous and synchronous contexts was further supported by teachers' deliberate integration of knowledge from asynchronous activities into live class discussions. Adaptive Treatment teachers used app participation data to co-design follow-up activities that celebrated progress, addressed emerging needs, and reinforced connections between vocabulary introduced in the app and broader science concepts discussed in class. This kind of integration likely helped students recognize the continuity of their learning across settings, making instruction feel more coherent and relevant.

The app itself may also have supported engagement by combining gamified features with tools that targeted key aspects of literacy development. Its read-aloud functionality for science books, coupled with activities targeting both constrained and unconstrained literacy skills (Paris, 2005), offered structured opportunities for practice and extension. Activities designed to foster metalinguistic awareness, such as playing with idioms, humor, and metaphors, encouraged students to think more deeply about language (Spector, 1996). These features likely enhanced vocabulary exposure and language engagement while also drawing on students' prior experiences and analytical thinking as learning resources (Zipke, 2008).

These results suggest that TBL-supported adaptations and the app's design features worked in tandem to support student engagement during pandemic-era instruction. Teachers collaborated to address participation barriers, scaffold student effort, and sustain engagement during self-directed work. At the same time, the app's interactive structure and literacy activities supported autonomy, motivation, and deeper processing, which are key factors for student success in online environments (Bond, 2020; Zimmerman, 2002). These results suggest that combining adaptive teaching practices with well-designed digital tools can help sustain student participation and learning in remote or hybrid learning environments where instructional continuity is harder to maintain (Archambault et al., 2022).

Effectiveness of Structured Adaptations on Learning Outcomes

The second research aim was to examine the effects of Adaptive Treatment on student performance 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 passage (ES = .12) rather than the mid-transfer (ES = .07) and far-transfer (ES = .02) passages. These findings suggest that 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 impact on student reading outcomes suggests that even modest adjustments to professional learning and instructional structure, such as a collaboratively planned extension lesson, can strengthen students' conceptual understanding when integrated into an adaptive implementation model. The Adaptive Treatment group included the extension lesson on the concept of *systems*, collaboratively designed during TBL session to deepen content understanding. In this lesson, teachers connected the scientific concept of systems to students' prior knowledge, facilitating schema-building and integrating new information with existing schemas. By strengthening 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. The Adaptive Treatment condition achieved dual objectives: maintaining fidelity to the core instructional framework while enriching student-teacher interactions during synchronous sessions. The study findings demonstrate how structured, teacher-initiated 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 (Bryk et al., 2015; Goldenberg & Gallimore, 1991; Quinn & J. Kim, 2017).

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 (Kuhfeld et al., 2022; Relyea et al., 2022). Despite these challenges, positive effects on student learning 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 teachers in both treatment conditions maintained learning continuity, 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 (J. Kim & Mosher, 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 (J. Kim et al., 2017), which reported a modest advantage (ES = 0.12) for students in the adaptive condition on reading comprehension. Unlike J. Kim et al.'s (2017) 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 (J. Kim et al., 2024). Domain-general reading comprehension develops incrementally over extended periods and requires sustained practice (Paris, 2005), 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 (Francis et al., 2022; Gilbert et al., 2023).

The heterogeneous treatment effects analyses 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 (Raudenbush et al., 1998). By balancing flexibility with fidelity, structured 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 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 Adaptive Treatment teachers could maintain levels of fidelity comparable to Core Treatment teachers, particularly in terms of differentiation and responsiveness components of fidelity (Dane & Schneider, 1998). Adaptive Treatment teachers 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 (J. Kim & Mosher, 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 (J. Kim et al., 2017). In doing so, structured 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). When conceptualized inclusively, fidelity can guide innovation by guiding principled modifications that enhance implementation quality and educational equity (Bryk et al., 2015; Durlak & DuPre, 2008). This integrated perspective advances understanding of how structured adaptation frameworks can bridge the research-practice gap, ensuring that interventions are scientifically rigorous while responsive to the diverse and needs of classrooms (Goldenberg & Gallimore, 1991; Joyce & Cartwright, 2020). While our findings demonstrate the promise of structured adaptations supported by researcher-teacher collaboration, we do not claim that this model is scalable in its current form. Instead, the study highlights transferable design principles, such as bounded instructional choices and collaborative structures (Bryk et al., 2015; Su & Reeve, 2011), that can inform the development of scalable, context-responsive interventions.

Implications for Practice

We view structured adaptations as program-agnostic implementation design principles for enhancing asset-based pedagogies to promote equitable implementation of evidence-based literacy programs and practices. Grounded in self-determination theory (Ryan & Deci, 2000), we designed structured adaptations to cultivate autonomy-supportive professional contexts that enhance the professional well-being of teachers (Deci & Ryan, 2012). We translated SDT into an implementation design framework in which teachers were asked to work in teams with a specific focus in mind, set realistic goals to work toward that focus, challenge their understanding of a content literacy program, and felt greater belonging with their colleagues and students during the isolation triggered by school closures during COVID-19. In many ways, the empirical findings supporting Adaptive Treatment over Core Treatment replicate findings from other instantiations of this idea in very diverse contexts, including summer reading involving school and home connections (J. Kim et al., 2017), early literacy programs using peer learning (Lemons at el., 2014), and preschool vocabulary building (Neuman et al., 2021). Thus, the principles underlying Adaptive Treatment are strongly grounded in theoretical research and replicated across a wide range of literacy programs and practices.

Limitations and Future Directions

Several limitations warrant attention in future research. First, the absence of a traditional treatment-versus-control (business-as-usual) comparison group limits the ability to assess the unique contribution of structured adaptations relative to typical instructional practices. Although the study design ensured equitable access to instructional opportunities during pandemic-related school closures, it precluded establishing a baseline against which to assess the distinct effects of Adaptive Treatment.

Second, the mechanisms underlying the observed effects were not explored in this study. Future research should investigate whether improvements in teacher responsiveness or student engagement function as mediators linking structured adaptations to student learning outcomes. Identifying such mechanisms may clarify how specific features of adaptive instruction produce their effects and help understand critical leverage points for maximizing their effectiveness.

Third, the study did not include teacher-reported perspectives on autonomy, competence, or relatedness. While implementation data reflect evidence of teacher agency, future work should incorporate teacher interviews or survey measures to understand better how teachers perceived autonomy-supportive features of the intervention and how these perceptions shaped their instructional decisions.

Finally, there is growing interest in applying learning analytics to examine collaborative learning processes (Yan et al., 2025). Future studies should aim to capture the depth and variability of teachers' engagement in TBL activities and explore how such collaboration influences implementation. Integrating artificial intelligence-based engagement metrics, such as real-time interaction patterns, prosodic signals, or behavioral indicators, could further enhance understanding of teacher engagement as they co-construct adaptations with researchers and their peers.

Conclusion

There is emerging evidence that structured adaptations can enhance program effectiveness (e.g., J. Kim et al., 2017; 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 belongingness 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* = 962)		Core treatment (*n* = 952)		Balance checks[b]	
	M[a]	*SD*	M[a]	*SD*	M[a]	*SD*	Difference	z-statistic
Black	0.37	0.48	0.36	0.48	0.38	0.48	0.00	-0.22
Other	0.03	0.17	0.02	0.16	0.04	0.19	-0.01	-1.42
Asian	0.09	0.29	0.07	0.26	0.11	0.31	-0.04	-1.77†
Hispanic	0.34	0.47	0.38	0.49	0.29	0.46	0.06	3.15**
White	0.17	0.38	0.16	0.37	0.19	0.39	-0.01	-0.50
AIG	0.12	0.33	0.13	0.33	0.12	0.33	0.01	0.48
Male	0.49	0.50	0.50	0.50	0.48	0.50	0.02	0.76
English learner	0.25	0.43	0.27	0.44	0.22	0.42	0.04	1.58
Home Language: English	0.37	0.48	0.36	0.48	0.37	0.48	-0.02	-0.87
IEP	0.08	0.28	0.09	0.28	0.08	0.28	0.00	0.10
Low SES	0.39	0.49	0.40	0.49	0.37	0.48	0.00	-0.22
Med SES	0.37	0.48	0.38	0.49	0.35	0.48	0.05	2.85**
High SES	0.25	0.43	0.22	0.41	0.28	0.45	-0.04	-3.91***
Baseline MAP Reading	189.56	18.06	189.01	17.85	190.12	18.26	-1.08	-1.26

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 proportions 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 at a 5% false discovery rate. 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 o Introduced the purpose of using the app/trifolds and connected it to classroom content learning. o Highlighted student learning from the app/trifolds through projects, discussions, or presentations. o Emphasized the relevance and benefits of the app/trifolds to students and parents.	• App/trifold engagement o Modeled how to use the app during virtual class sessions. o Provided structured in-class time for app/trifold engagement. o Set clear expectations by integrating app/trifold use into routines or assignments. o Encouraged consistent use through regular communication with parents. o Incentivized usage with rewards. o Monitored engagement and offered individual follow-ups for encouragement and support.
Synchronous online lessons	• Extended lesson activities o Encouraged abstract thinking by connecting the concept of *systems* to classroom content and broader contexts. o Reinforced understanding of *systems* through mad libs, discussions, and presentations. o Provided word banks and visuals to support English learners.	• Extended lesson activities o Set clear expectations and modeled an example presentation to demonstrate expectations. o Guided students in completing and refining their *system* mad libs. o Facilitated peer sharing and engagement through structured opportunities. o 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			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	962	0.60	0.49	952	0.06*	0.03
Number of science books completed on app library	0.81	1.60	962	0.47	1.15	952	0.31***	0.08
Number of target words	2.11	1.79	962	1.85	1.75	952	0.27*	0.10
Total time spent on app activities[a]	1.71	1.77	962	1.54	1.73	952	0.17†	0.10
Overall app activity accuracy[b] (std)	0.03	1.03	554	0.09	1.00	505	-0.003	0.05
Engagement with app activities aligned with synchronous curriculum lessons								
Ever access curriculum lessons	0.83	0.38	962	0.85	0.36	952	-0.002	0.02
Proportion of all curriculum lessons completed	0.57	0.40	962	0.60	0.39	952	0.00	0.02
Proportion of interactive read-aloud lessons completed	0.62	0.40	962	0.65	0.39	952	0.00	0.02
Proportion of word sleuthing lessons completed	0.59	0.39	962	0.62	0.39	952	0.001	0.02
Accuracy on curriculum activities (std)	-0.02	1.02	792	0.07	1.00	820	-0.004	0.04
Accuracy on end-of-unit quizzes (std)	-0.01	1.00	585	0.03	1.01	591	0.03	0.05
Total time spent on curriculum lessons[b]	4.16	2.10	962	4.26	2.01	952	0.01	0.14
Perceived motivation and task challenges with app activities								
Enjoyment of app activities	2.92	0.69	952	2.86	0.71	946	0.10*	0.05
Reader self-competence beliefs	2.19	0.55	952	2.16	0.55	946	0.10*	0.05
How difficult was the task	2.01	0.39	952	2.01	0.38	946	-0.01	0.05
Print-based trifolds activities								
Return any trifolds (1 = yes, 0 = no)	0.10	0.30	962	0.09	0.29	952	0.01	0.01
Number of trifolds returned	0.95	3.19	962	0.79	2.88	952	0.16	0.14

Note. 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. *P*-values below 0.002 are significant when applying the Benjamini-Hochberg (BH) correction at a 5% false discovery rate.

[a]Log-transformed minutes. We add 1 to the total time spent on the 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[a]	20.00	7.92	962	20.54	7.69	952	0.02	0.03	0.00
Science background knowledge[a]	4.56	2.05	962	4.53	2.05	952	0.09*	0.04	0.00
Reading comprehension									
Science content reading comprehension[a]									
All passages	13.01	6.52	962	13.16	6.51	952	0.07*	0.03	0.00
Near-transfer passage	4.96	2.67	962	4.90	2.61	952	0.12**	0.04	0.00
Mid-transfer passage	4.31	2.35	933	4.35	2.35	930	0.07	0.04	0.03
Far-transfer passage	4.14	2.28	901	4.25	2.34	898	0.02	0.04	0.00
Domain-general reading comprehension									
MAP	191.70	18.82	918	193.20	18.10	899	-0.02	0.03	0.00
EOG reading	433.66	9.52	888	434.55	9.97	880	-0.03	0.03	0.04

Note. Point estimates for the Adaptive Treatment effect were derived from multilevel models that included the Adaptive Treatment indicator, school fixed effects, teacher random effects, student demographics, and baseline Measure of Academic Progress (MAP) reading scores; therefore, these estimates differ from the raw difference in means. The standard deviation (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*-values below 0.006 are significant when applying the Benjamini-Hochberg (BH) correction at a 5% false discovery rate. MAP = Measure of Academic Progress. EOG = End-of-grade.
[a]For all researcher-developed assessments, we used Expected a Posteriori latent trait scores derived from a two-parameter logistic IRT model as our outcome variables, standardized to mean 0 and standard deviation 1 in our sample.
*$p < 0.05$, **$p < 0.01$.

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 the 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. †$p < 0.10$, *$p < 0.05$, **$p < 0.01$. *P*-values below 0.003 are significant when applying the Benjamini-Hochberg (BH) correction at a 5% false discovery rate.

Table 7

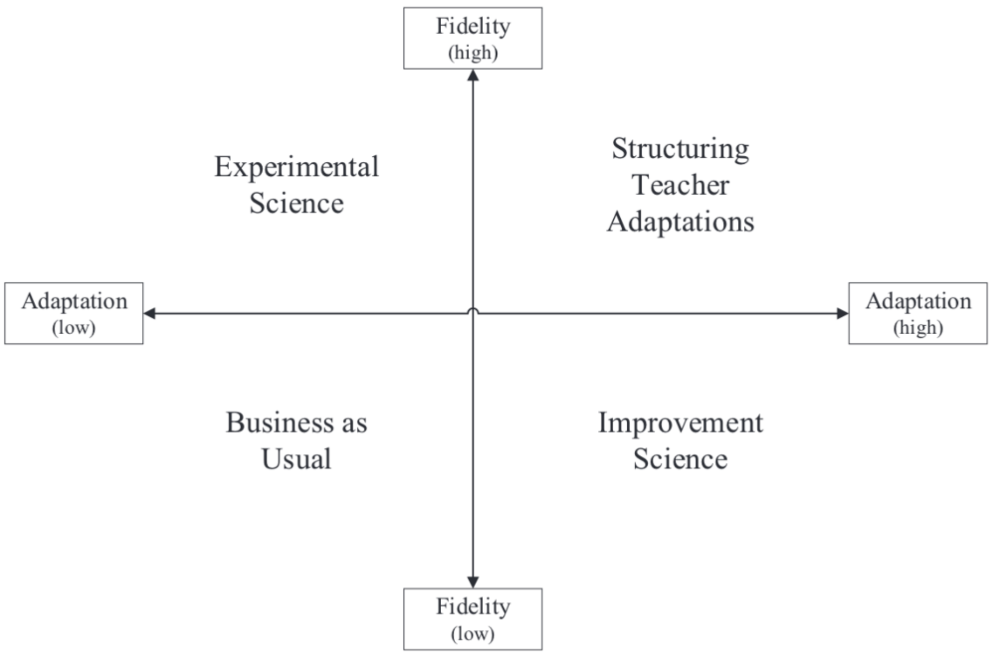
Fidelity of Implementation: Responsiveness Reflected in Student-Teacher Interaction in Adaptive and Core Treatment Conditions

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

Note. $^{**}p < 0.01$, $^{***}p < 0.001$. *P*-values below 0.02 are significant when applying the Benjamini-Hochberg (BH) correction at a 5% false discovery rate.

Figure 1

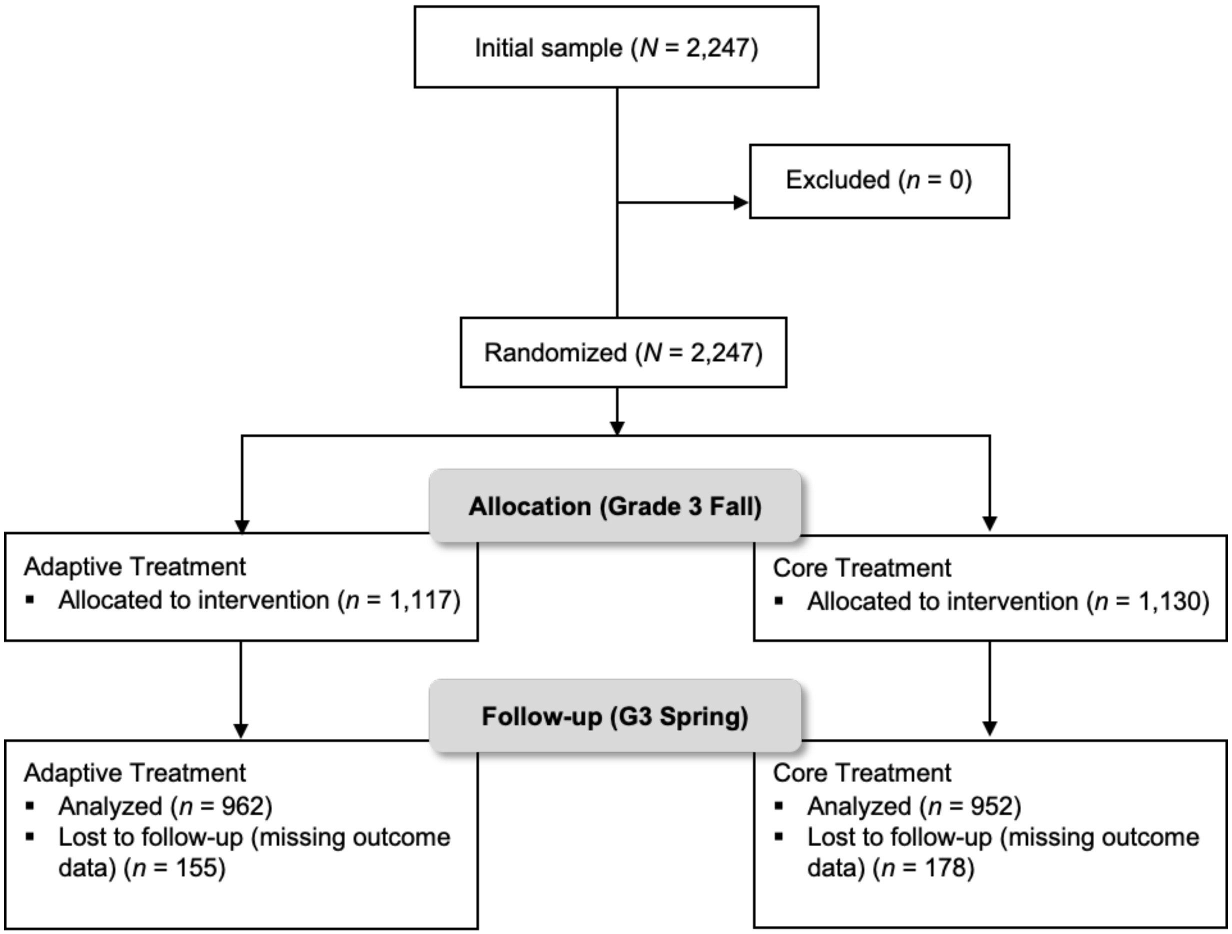
A Quadrant Framework for Teacher Adaptations and Fidelity in Interventions



Note. Source from J. Kim & Mosher (2023)

Figure 2

Consort Diagram for the Randomization Process



References

Ahmadi, A., Noetel, M., Parker, P., Ryan, R. M., Ntoumanis, N., Reeve, J., Beauchamp, M., Dicke, T., Yeung, A., Ahmadi, M., Bartholomew, K., Chiu, T. K. F., Curran, T., Erturan, G., Flunger, B., Frederick, C., Froiland, J. M., González-Cutre, D., Haerens, L., . . . Lonsdale, C. (2023). A classification system for teachers' motivational behaviors recommended in self-determination theory interventions. *Journal of Educational Psychology, 115*(8), 1158–1176. https://doi.org/10.1037/edu0000783

Alexander, R. J. (2008). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). Dialogos.

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

Authors. (2017a, 2017b, 2017c, 2019, 2021, 2022, 2023a, 2023b, 2024a, 2024b, 2024c).

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

Bridwell-Mitchell, E. N. (2015). Theorizing teacher agency and reform: How institutionalized instructional practices change and persist. *Sociology of Education, 88*(2), 140-159. 1 https://doi.org/10.1177/0038040715575559

Brown, M. W. (2009). The teacher-tool relationship: Theorizing the design and use of curriculum materials. In J. T. Remillard, B. A. Herbel-Eisenmann, & G. M. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction* (pp. 17–36). Routledge.

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.

Burkhauser, M. A., & Lesaux, N. K. (2017). Exercising a bounded autonomy: Novice and experienced teachers' adaptations to curriculum materials in an age of accountability. *Journal of Curriculum Studies, 49*(3), 291-312.

Butler, D. L., & Schnellert, L. (2012). Collaborative inquiry in teacher professional development. *Teaching and Teacher Education, 28*(8), 1206-1220. https://doi.org/10.1016/j.tate.2012.07.009

Cabell, S. Q., Kim, J. S., White, T. G., Gale, C. J., Edwards, A. A., Hwang, H., Petscher, Y., & Raines, R. M. (2025). Impact of a content-rich literacy curriculum on kindergarteners' vocabulary, listening comprehension, and content knowledge. *Journal of Educational Psychology, 117*(2), 153–175. https://doi.org/10.1037/edu0000916

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

Century, J., & Cassata, A. (2016). Implementation research: Finding common ground on what, how, why, where, and who. *Review of Research in Education, 40*(1), 169-215. https://doi.org/10.3102/0091732X16665332

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

Coburn, C. E., & Woulfin, S. L. (2012). Reading coaches and the relationship between policy and practice. *Reading Research Quarterly, 47*(1), 5-30. http://dx.doi.org/10.1002/RRQ.008

Cohen, D. K., & Ball, D. L. (2001). Making change: Instruction and its improvement. *Phi Delta Kappan, 83*(1), 73-77. https://doi.org/10.1177/003172170108300115

Cohen, J., Boguslav, A., Wyckoff, J., Katz, V., Sadowski, K., & Wiseman, E. A. (2025). Core requirements, structured flexibility, and local judgment: balancing adherence and adaptation in the design and implementation of district-wide professional development. *Educational Evaluation and Policy Analysis, 47*(1), 263-291. https://doi.org/10.3102/01623737231210285

Collie, R. J., & Martin, A. J. (2017). Teachers' sense of adaptability: Examining links with perceived autonomy support, teachers' psychological functioning, and students' numeracy achievement. *Learning and Individual Differences, 55*, 29–39. https://doi.org/10.1016/j.lindif.2017.03.003

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes and development* (pp. 43–77). Lawrence Erlbaum Associates, Inc.

Correnti, R., & Rowan, B. (2007). Opening up the black box: Literacy instruction in schools participating in three comprehensive school reform programs. *American Educational Research Journal, 44*(2), 298-339. https://doi.org/10.3102/0002831207302501

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

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19*(2), 109–134. https://doi.org/10.1016/0092-6566(85)90023-6

Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

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.

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

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

Francis, D. J., Kulesz, P. A., Khalaf, S., Walczak, M., & Vaughn, S. R. (2022). Is the treatment weak or the test insensitive: Interrogating item difficulties to elucidate the nature of reading intervention effects. *Learning and Individual Differences, 97*, 102167. https://doi.org/10.1016/j.lindif.2022.102167

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Fries, L., Son, J. Y., Givvin, K. B., & Stigler, J. W. (2021). Practicing connections: A framework to guide instructional design for developing understanding in complex domains. *Educational Psychology Review, 33*(2), 739–762. https://doi.org/10.1007/s10648-020-09561-x

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

Gallucci, C. (2008). Districtwide instructional reform: Using sociocultural theory to link professional learning to organizational support. *American Journal of Education, 114*(4), 541-581. https://doi.org/10.1086/589314

Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). Reading First impact study. Final report. NCEE 2009-4038. *National Center for Education Evaluation and Regional Assistance*.

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice* (3rd ed.). Teachers College Press.

Gilbert, J. B., Kim, J. S., & Miratrix, L. W. (2023). Modeling item-level heterogeneous treatment effects with the explanatory item response model: Leveraging large-scale online assessments to pinpoint the impact of educational interventions. *Journal of Educational and Behavioral Statistics, 48*(6), 889-913. https://doi.org/10.3102/10769986231171710

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., 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

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

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

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

Honig, M. I. (Ed.). (2006). *New directions in education policy implementation: Confronting complexity*. State University of New York Press.

Hord, S.M. (1997). *Professional learning communities: Communities of continuous inquiry and improvement.* Austin: Southwest Educational Development Laboratory.

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, 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)

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research, 86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Kim J. S., & Mosher, D. (2023). Structured adaptations for scaling up evidence-based literacy interventions. In S. Cabell, S. B. Neuman & N. P. Terry (Eds.), *Handbook on the science of early literacy* (pp. 253-268). Guilford.

Kim, J. S. (2019). Making every study count: Learning from replication failure to improve intervention research. *Educational Researcher, 48*(9), 599–607. https://doi.org/10.3102/0013189X19891428

Kim, J. S., Burkhauser, M. A., Mesite, L. M., Asher, C. A., Relyea, J. E., Fitzgerald, J., & Elmore, J. (2021). Improving reading comprehension, science domain knowledge, and reading engagement through a first-grade content literacy intervention. *Journal of Educational Psychology, 113*(1), 3-26. https://doi.org/10.1037/edu0000465

Kim, J. S., Burkhauser, M. A., Quinn, D. M., Guryan, J., Kingston, H. C., & Aleman, K. (2017). Effectiveness of structured teacher adaptations to an evidence-based summer literacy program. *Reading Research Quarterly, 52*(4), 443–467. https://doi.org/10.1002/rrq.178

Kim, J. S., Gilbert, J. B., Relyea, J. E., Rich, P., Scherer, E., Burkhauser, M. A., & Tvedt, J. N. (2024). Time to transfer: Long-term effects of a sustained and spiraled content literacy intervention in the elementary grades. *Developmental Psychology. 60*(7), 1279–1297. https://doi.org/10.1037/dev0001710

Kim, J. S., Relyea, J. E., Burkhauser, M. A., Scherer, E., & Rich, P. (2021). Improving elementary grade students' science and social studies vocabulary knowledge depth, reading comprehension, and argumentative writing: A conceptual replication. *Educational Psychology Review, 33*, 1935-1964. https://doi.org/10.1007/s10648-021-09609-6

Kim, Y.-S. G., & Snow, C. (2025). Advancing the science of teaching reading: Introduction to the special issue. *Scientific Studies of Reading, 29*(1), 1-6. https://doi.org/10.1080/10888438.2024.2438630

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

Krechevsky, M., Mardell, B., Rivard, M., & Wilson, D. (2013). *Visible learners: Promoting Reggio-inspired approaches in all schools.* John Wiley & Sons.

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

Ladison-Billings, G. (2014). Culturally relevant pedagogy 2.0: A. K. A. the remix. *Harvard Educational Review, 84*(1), 74–84. https://doi.org/10.17763/haer.84.1.p2rj131485484751

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

Levinson, M., Geron, T., & Brighouse, H. (2022). Conceptions of educational equity. *AERA Open, 8*. https://doi.org/10.1177/23328584221121344

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.

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

Mehta, J., & Fine, S. (2019). *In search of deeper learning: The quest to remake the American high school.* Harvard University Press.

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

Mosher, D. M., Burkhauser, M. A., & Kim, J. S. (2024). Improving second-grade reading comprehension through a sustained content literacy intervention: A mixed-methods study examining the mediating role of domain-specific vocabulary. *Journal of Educational Psychology, 116(*4), 550–568. https://doi.org/10.1037/edu0000868

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). U.S. Department of Health and Human Services, National Institutes of Health. https://www.nichd.nih.gov/publications/pubs/nrp/smallbook

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

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education, 7*(2), 133-144. https://doi.org/10.1177/1477878509104318

North Carolina Department of Public Instruction. (2020). *The North Carolina Department of Public Instruction Mathematics 3–8 End of Grade (EOG) NC Math 1 and NC Math 3 End of Course (EOC) technical report 2018–2019*. Retrieved from https://www.dpi.nc.gov/media/10219/open

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

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. PISA, OECD Publishing, Paris. https://doi.org/10.1787/9789264073234-en

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

Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., Farrell, C., & Allen, A.-R. (2017). How school and district leaders access, perceive, and use research. *AERA Open, 3*(2). https://doi.org/10.1177/2332858417705370

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

Quinn, D. M., & Kim, J. S. (2017). Scaffolding fidelity and adaptation in educational program implementation: Experimental evidence from a literacy intervention. *American Educational Research Journal, 54*(6), 1187-1220. https://doi.org/10.3102/0002831217717692

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., Cheon, S. H., & Yu, T. H. (2020). An autonomy-supportive intervention to develop students' resilience by boosting agentic engagement. *International Journal of Behavioral Development, 44*(4), 325–338. https://doi.org/10.1177/0165025420911103

Relyea, J. E., Kim, J. S., Rich, P., & Fitzgerald, J. (2024). Effects of tier 1 content literacy intervention on early-grade English learners' reading and writing: Exploring the mediating roles of domain-specific vocabulary and oral language proficiency. *Journal of Educational Psychology, 116*(7), 1172–1195. https://doi.org/10.1037/edu0000882

Relyea, J. E., Rich, P., Kim, J. S., & Gilbert, J. B. (2022). The COVID-19 impact on reading achievement growth of Grade 3–5 students in a U.S. urban school district: Variation across student characteristics and instructional modalities. *Reading and Writing, 36*(2), 317–346. https://doi.org/10.1007/s11145-022-10387-y

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

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

Shulman, L. S. (1981). Disciplines of inquiry in education: An overview. *Educational Researcher, 10*(6), 5–23. https://doi.org/10.3102/0013189X010006005

Shulman, L. S. (1986). Those who understand a conception of teacher knowledge. *American Educator, 10*, 9-15.

Skinner, E., & Pitzer, J. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of* *research on student engagement* (pp. 21–44). Springer. https://doi.org/10.1007/978-1-4614-2018-7_2

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge University Press. https://doi.org/10.1017/CBO9780511609664.008

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

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387–431. https://doi.org/10.3102/00346543072003387

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

Su, Y.-L., & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review, 23*(2), 159–188. https://doi.org/10.1007/s10648-010-9142-7

Swanson, E., McCulley, L. V., Osman, D. J., Scammacca Lewis, N., & Solis, M. (2019). The effect of team-based learning on content knowledge: A meta-analysis. *Active Learning in Higher Education, 20*(1), 39–50. https://doi.org/10.1177/1469787417731201

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