ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 24-1001

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

Jackie E. Relyea
North Carolina State University

Joshua B. Gilbert
Harvard University

Mary Burkhauser
Harvard University

Ethan Scherer
Harvard University

Douglas Mosher
Harvard University

Zhongyu Wei
Harvard University

Johanna Tvedt
Harvard University

James S. Kim
Harvard University

Scaling up evidence-based educational interventions to improve student outcomes presents challenges, particularly in adapting to new contexts while maintaining fidelity. Structured teacher adaptations that integrate the strengths of experimental science (high fidelity) and improvement science (high adaptation) offer a viable solution to bridge the research-practice divide. This preregistered randomized controlled trial study examines the effectiveness of structured teacher adaptations in a Tier 1 content literacy intervention delivered through asynchronous and synchronous methods during COVID-19 on Grade 3 students' (N = 1,914) engagement in digital app and print-based reading activities, student-teacher interactions, and learning outcomes. Our structured teacher adaptations achieved higher average outcomes and minimal treatment heterogeneity across schools, thereby enhancing the effectiveness of the intervention rather than undermining it.

VERSION: July 2024

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. (2024). Effectiveness of Structured Teacher Adaptations to an Online Content Literacy Intervention for Third Graders: A Randomized Controlled Trial During COVID-19. (EdWorkingPaper: 24-1001). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/ztcp-v322

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

Jackie E. Relyea[1]
https://orcid.org/0000-0002-7560-7136

Joshua B. Gilbert[2]
https://orcid.org/0000-0003-3496-2710

Mary Burkhauser[2]

Ethan Scherer[2]
https://orcid.org/0000-0002-6548-7724

Douglas Mosher[2]
https://orcid.org/0000-0001-5722-9218

Zhongyu Wei[2]

Johanna Tvedt[2]

James S. Kim[2]
https://orcid.org/0000-0002-6415-5496

[1]College of Education, North Carolina State University

[2]Harvard Graduate School of Education

Corresponding Author:
Jackie E. Relyea (jrelyea@ncsu.edu)

Senior Author:
James S. Kim (james_kim@gse.harvard.edu)

Funding source:
This project was supported by the Chan Zuckerberg Initiative (2017000341). The opinions expressed are those of the authors and do not represent the views of the funding agency.

Abstract

Scaling up evidence-based educational interventions to improve student outcomes presents challenges, particularly in adapting to new contexts while maintaining fidelity. Structured teacher adaptations that integrate the strengths of experimental science (high fidelity) and improvement science (high adaptation) offer a viable solution to bridge the research-practice divide. This preregistered randomized controlled trial study examines the effectiveness of structured teacher adaptations in a Tier 1 content literacy intervention delivered through asynchronous and synchronous methods during COVID-19 on Grade 3 students' (N = 1,914) engagement in digital app and print-based reading activities, student-teacher interactions, and learning outcomes. Our structured teacher adaptations achieved higher average outcomes and minimal treatment heterogeneity across schools, thereby enhancing the effectiveness of the intervention rather than undermining it.

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

Scaling up evidence-based educational interventions is a formidable challenge because it requires schools and teachers to implement them with fidelity in new settings (Coburn, 2003). However, the necessity to adapt interventions to diverse educational settings often leads to tensions between fidelity and adaptation in educational research. Fidelity, rooted in the experimental science paradigm, emphasizes strict adherence to established protocols to ensure that observed outcomes can be reliably attributed to interventions (Bos et al., 2022; Dane & Schneider, 1998). In contrast, the improvement science paradigm advocates adapting interventions to the unique demands of specific contexts, enhancing responsiveness and practical utility for educators and learners (Domitrovich et al., 2010; Gutiérrez & Penuel, 2014; Lewis, 2015).

Historically, fidelity and adaptation in educational interventions have been viewed as mutually exclusive frameworks. However, the COVID-19 pandemic dramatically underscored the need to reconcile these approaches and find integrative solutions that capture the benefits of both experimental and improvement science. The abrupt shift to remote learning compelled educators to maintain instructional fidelity while simultaneously adapting their teaching methods to accommodate the diverse needs of students, who experienced varying levels of access to technology and different home learning environments. This situation exposed the necessity for a deeper understanding of how intervention research can effectively balance adherence to prescribed curricula with adaptation to leverage teachers' local knowledge under unexpected circumstances (Durlak & DuPre, 2008). Furthermore, causal research on the effectiveness of structuring adaptations that promote program implementation fidelity and flexibility is rare.

In this study, we examine the effectiveness of structured teacher adaptations in a randomized controlled trial (RCT) of the classroom-based (Tier 1) content literacy intervention, targeting Grade 3 students during the 2020-2021 school year. Due to COVID-19, we adjusted the original RCT design to implement the core treatment intervention to all participants, eliminating a traditional control group. Instead, we randomly assigned 95 classrooms to one of two conditions: *Core Treatment*, where teachers faithfully replicated the procedures and content used in previous studies (e.g., Kim et al., 2017; Quinn & Kim, 2017), or *Adaptive Treatment*, which included all elements of Core Treatment with the addition of structured teacher adaptations. In Adaptive Treatment, teachers participated in Team-Based Learning (TBL) groups and had opportunities to use their professional craft knowledge and generalizable research knowledge to customize the intervention to meet their students' needs without compromising foundational principles. Our study conceptually replicated the implementation procedures in a fully digital and remote education context to further explore how structured teacher adaptations affected students' engagement and achievement outcomes.

Fidelity and Adaptation

Two dimensions of program implementation—fidelity and adaptation—place specific demands on teachers in evidence-based interventions. The fidelity-focused model expects that teachers deliver the program exactly as designed, treating any deviation as a compromise to the program's integrity (Domitrovich et al., 2010; Sherin & Drake, 2009). This approach restricts teachers' autonomy, judgment, innovation, and adaptation, aligning with a tightly controlled instructional management model where teachers' roles are prescribed, and administrators ensure compliance (Rowan, 1990). While fidelity is essential for evaluating program outcomes, its success in scaling is rare and inconsistent. Promising programs in initial trials frequently struggle to replicate their positive outcomes in broader applications due to variability in effectiveness across diverse populations and settings, exacerbated by the absence of collaborative work structures necessary for professional growth (Kim, 2019).

In contrast, the adaptative approach allows for modifications that preserve core intervention components while adapting to local needs and contexts (Bryk et al., 2015). It focuses on deepening teachers' understanding of the intervention, fostering changes in practice, and promoting ownership of the adaptation process—essential factors for successful scalability (Coburn, 2003). The adaptability of evidence-based interventions to specific school contexts is crucial, as opposed to a uniform application across varied environments (McDonald et al., 2006). Effective adaptations integrate scientific knowledge with local insights, enhancing the program's suitability and effectiveness (Stanovich, 2003). Teachers, leveraging their local knowledge, play a vital role in bridging the gap between research and practice, thus improving outcomes and relevance across diverse settings (Goldenberg & Gallimore, 1991).

Fidelity and adaptation, when viewed as synergistic rather than competing priorities, enhance program implementation and impact (Durlak & DuPre, 2008; Kim et al., 2017). Maintaining fidelity to core components provides a consistent framework, ensuring the integrity and desired outcomes of the program. Concurrently, adaptations that respect these core components while tailoring delivery to local conditions can augment program effectiveness and participant engagement. This dual approach provides a more nuanced understanding of how interventions can be faithfully implemented and flexibly adapted. Therefore, fidelity and adaptation are complementary approaches that, when effectively integrated, support broader outcomes and optimize the scalability and suitability of interventions.

Structured Teacher Adaptations

When scaling evidence-based interventions across diverse local contexts, fidelity and adaptation are inherently connected (Kim & Mosher, 2023). Figure 1 illustrates this balance, positioning structured teacher adaptations in the high-fidelity, high-adaptation quadrant. This approach integrates the strengths of both experimental science (high fidelity) and improvement science (high adaptation), bridging the research-practice divide by ensuring interventions are tailored to specific educational settings while maintaining their core components.

Structured teacher adaptations offer a systematic approach that guides teachers in modifying programs while maintaining their core principles, thus increasing relevance, applicability, and effectiveness. Beyond merely granting autonomy to teachers, this approach involves developing management systems that foster collaborative and peer learning environments, promoting innovation, and encouraging collective problem-solving among teachers (Elmore, 1996). By equipping teachers with the necessary skills and knowledge, these frameworks cultivate a collaborative culture that enhances students' learning experience (Frank et al., 2011). While empirical research on structured adaptations is still evolving, existing studies highlight their role in preserving fidelity while improving student outcomes (e.g., Lemons et al., 2014; Neuman et al., 2021).

Effective implementation of structured teacher adaptations requires a clear framework delineating adaptable and non-negotiable components. Researchers and teachers should collaboratively develop this framework, respecting both the program's theoretical underpinnings and practical educational realities. The process includes initial training, ongoing support, and feedback mechanisms for teachers to share their experiences and outcomes. For example, structured flexibility in our intervention allowed tailoring to local needs, strengthening teachers' adaptive expertise and capacity to implement evidence-based literacy instruction effectively.

Previous evidence (Kim et al., 2017) shows that students in the adaptations condition scored 0.12 standard deviations higher in reading comprehension than those in the core condition. Building on this research, our study compared the content literacy intervention's fidelity-focused core condition with the condition that included both the fidelity-focused core and structured teacher adaptations in a virtual setting during the COVID-19 pandemic, focusing on how adaptability sustains program effectiveness amid schooling disruptions.

The Present Study

The Intervention

The content literacy intervention in this study focuses on developing knowledge structures, or schemas, by integrating domain knowledge building and transfer into language and literacy instruction (Kim et al., 2023, 2024). It emphasizes activating and constructing schemas to help learners interpret and synthesize knowledge from texts (Anderson & Pearson, 1984; Kintsch & van Dijk, 1978). The construction-integration model (Kintsch, 1993) describes reading comprehension as a dynamic cognitive process that integrates preliminary text understanding with prior knowledge to form a coherent mental representation. Effective schema development is essential for Grade 3 students to comprehend and transfer knowledge across related topics, thus improving overall learning outcomes.

The intervention curriculum, structured around a thematic science unit, delivers complex, conceptually related science learning content through the asynchronous digital app and print-based books/trifolds activities and synchronous online lessons. The app activities promote metalinguistic awareness by encouraging reflection on language structures and engagement with linguistic ambiguities (Nagy & Anderson, 1999), supporting language development in an active, engaged, meaningful, and socially interactive context (Hirsh-Pasek et al., 2015). Additionally, students accessed science texts to foster a deep understanding of the content and build coherent text representations (McNamara et al., 1996). Synchronous teacher-directed lessons focused on fostering schema awareness, building academic vocabulary networks, and making connections between known and new topics through analogical mapping activities (Gick & Holyoak, 1983), including reading, writing, and discussion activities about different types of living and nonliving systems (e.g., how is the anatomy of a skyscraper like a human body, how is my watch an example of a system that has parts that move or work together). Systematic exposure to aligned books and semantically connected vocabulary strengthens their semantic networks and constructs a cohesive schema of interconnected concepts (Fitzgerald et al., 2020).

Study Design and Research Questions

We examine the effects of the structured teacher adaptations of our content literacy intervention, implemented online during the COVID-19 school closures. Classroom teachers, blocked by schools, were randomly assigned to either *Core Treatment* directly replicating previous implementation procedures (e.g., Kim et al., 2017) or *Adaptive Treatment*, which augmented the core treatment with structured teacher adaptations. The core components of the intervention included asynchronous activities using a digital app and print-based reading, and synchronous online lessons led by classroom teachers. This study uniquely contributes to the literature by exploring the causal impact of structured teacher adaptations during the pandemic on third graders' engagement, student-teacher interactions, and learning outcomes. The primary research questions (RQs) are: Compared to Core Treatment, what are the effects of Adaptive Treatment on student engagement in asynchronous activities (RQ1), student-teacher interactions during synchronous online lessons (RQ2), and student science vocabulary knowledge, background knowledge, and reading comprehension (RQ3)?

Methods[1]

Study Participants

Participants included 2,247 students and their 95 classroom teachers from 26 elementary schools in an urban district in the southeastern United States. Teachers were randomly assigned to either the Adaptive or Core Treatment conditions within schools. Figure A1 provides a consort flow diagram showing the randomization process and attrition rate by condition. Overall attrition was 14.8%, yielding a final sample of 1,914 students; differential attrition between the groups was not statistically significant. Table A1 details student characteristics by treatment condition. Balance tests indicate minimal baseline differences between the groups, supporting the internal validity of the study design and a causal interpretation of the results.

Intervention Program

Intervention Curriculum and Implementation

In early January 2021, all teachers participated in a synchronous online training session where the research team introduced the intervention's theory of change, empirical basis, curriculum, and lesson plans. Teachers then received an online curriculum for a five-week "human body system" unit, including asynchronous activities (app and books/trifolds) and 10-day synchronous Zoom lesson plans (two 35-minute sessions weekly).

In mid-January, students engaged in asynchronous flipped classroom activities at home, using the app, five science books, and trifolds to familiarize them with domain-specific vocabulary and foundational content before synchronous sessions. Teachers conducted synchronous Zoom lessons from February to March, focusing on goal setting, celebrating

[1]Additional details about the study procedures, measures, and analytic approaches can be found in the Supplementary Online Material (SOM) on the journal website. Replication data and code is available at the following URL: [BLINDED]. Our preregistration materials are available at the following URL: [BLINDED]

progress, fostering collaboration, enhancing science vocabulary through interactive concept maps and morphological analysis, and deepening understanding through collaborative research (Appendix A details the activity sequence and descriptions).

Structured Teacher Adaptations

Both Core and Adaptive Treatment teachers delivered the same core components of the intervention through asynchronous and synchronous lesson activities, with Adaptive Treatment teachers further implementing structured adaptations. Table 1 outlines the operationalization of both conditions and acceptable adaptations for Adaptive Treatment. For asynchronous activities, both treatments used the app and print-based books/trifolds. Adaptive Treatment teachers, however, collaborated regularly with literacy coaches and the research team to tailor and implement strategies, enhancing student engagement. Core and Adaptive Treatment teachers delivered the 10-day intervention lessons in synchronous online lessons. Additionally, Adaptive Treatment teachers delivered a 35-minute extension lesson (Appendix B) to deepen students' understanding of the schema for a *system*, to foster language extensions using the vocabulary related to the lessons of the human body, and to encourage abstract thinking.

Moreover, Adaptive Treatment teachers engaged in a TBL framework (Michaelsen & Sweet, 2008) to enhance their knowledge and application of the intervention. They completed four asynchronous modules to deepen their understanding of the intervention's rationale, lesson components, and effective instructional strategies. Then, they attended a 90-minute synchronous session led by the research team, collaborating to apply the knowledge gained and implement structured procedural and content-based adaptations tailored to their local contexts. This process aimed to improve students' abilities to recall, retrieve, and transfer knowledge of science concept words, promote engagement with the app and books/trifolds, and maintain the intervention's integrity.

Measures

Fidelity of Implementation

We administered a 13-item teacher survey (Table 2) to assess teacher involvement in student engagement with asynchronous reading activities. Teachers rated items on a 5-point Likert scale from 1 (*never*) to 5 (*every day*), reflecting the frequency of strategy engagement (Cronbach's $\alpha = .83$).

Quality of Student-Teacher Interaction

To measure the quality of student-teacher interaction in synchronous Zoom classes, we developed a rubric (Appendix C) to evaluate three dimensions: (a) engagement, assessing the frequency and depth of dialogues, (b) questioning, gauging the elicitation of higher-order thinking through open-ended questions, and (c) feedback, evaluating cognitive process encouragement, detailed feedback, and positive reinforcement. Each dimension was scored on a 4-point Likert scale (1 = *low*, 4 = *high*) based on audio-recorded lessons. Interrater reliability was .91.

Science Vocabulary Knowledge Depth

We developed a 36-item semantic association task measure (Stahl & Fairbanks, 1986; Appendix D) to assess students' science vocabulary knowledge depth and their ability to identify semantic relationships ($\alpha = .90$). Each item presented a word with four options, directing students to "circle two words that go with" the target word. Scoring was 1 for correct-only selections and 0 otherwise.

Science Content Reading Comprehension and Background Knowledge

Students completed a 29-item multiple choice assessing their ability to read near-, mid-, and far-transfer passages. Near- and mid-transfer passages included progressively fewer taught words, while the far-transfer passage included none. These passages focused on the skeletal, muscular, and nervous systems of living entities (monkeys and birds) and non-living structures (skyscrapers). Each passage included multiple-choice questions to assess main idea identification, word or phrase meaning, scientific concept description, and knowledge integration, and was scored dichotomously ($\alpha = .86$; Appendix E). Additionally, three items testing background knowledge preceded each reading comprehension passage. This assessment was administered to the whole class, with passages and questions read aloud to students ($\alpha = .55$; Appendix F).

Domain-General Reading Comprehension

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

MAP. The MAP reading assessment (NWEA, 2019), administered before and after the intervention implementation, evaluated students' independent reading comprehension, inferential and predictive skills, ability to draw conclusions, text structure analysis, and evaluation of the author's craft and purpose in narrative and informational texts. Test-retest reliabilities range from .79 to .86.

EOG Statewide Assessment. The EOG assessed students' reading performance on literary and informational texts from the [BLINDED] Standard Course of Study and their comprehension question responses. Internal consistencies were approximately .90 across demographic subgroups.

Student Engagement

We evaluated student engagement in digital app activities, print-based books/trifolds, motivational engagement, and curriculum activities. For digital and print reading activities, we analyzed backend data on app library access, science book completions, target words accessed, activity accuracy, time spent, and trifold returns. Motivational engagement was measured through an in-app survey assessing students' enjoyment, self-competence beliefs, and perceived task difficulty. Curriculum activity engagement was assessed by lesson completion rates, including Zoom and interactive read-aloud sessions, and accuracy in related activities and quizzes.

Data Analysis

We specified a hierarchical linear model with students at level 1, teachers at level 2, and fixed effects for schools at level 3 to account for the multi-site cluster-randomized design. The reduced-form model is specified as follows:

$$Y_{ijk} = \alpha_k + \beta_1 Adaptive_{jk} + \sum_{p=2}^{11} \beta_p X_{p_{ijk}} + \zeta_{jk} + \epsilon_{ijk}$$

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

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

where Y_{ijk} is the outcome for student i in teacher j's class in school k, α_k is the school fixed effect, β_1 is the adjusted intention-to-treat (ITT) effect of Adaptive Treatment, and $\beta_p X_{p_{ijk}}$ is a set of student-level covariates (baseline MAP reading scores, gender, race/ethnicity, English learner status, home language, individualized education plan status, and neighborhood poverty). The model includes a teacher random effect, ζ_{jk}, and student-level residual, ϵ_{ijk}. We fit analogous

Ordinary Least Squares regression models for teacher-level outcomes, including school fixed effects and covariates[2].

Results

Fidelity of Implementation

Overall, both Core and Adaptive Treatment teachers similarly facilitated students' app and trifold activities. However, as shown in Table 2, Adaptive Treatment teachers more frequently followed up with students who did not complete activities, modeled trifold use, set clearer expectations for the number of trifolds, and communicated with families about trifolds ($ps < .05$), indicating a more proactive and involved approach to engaging students and families.

RQ1: The Impact on Student Engagement in Asynchronous Activities

Table 3 shows that the Adaptive Treatment caused significantly higher engagement in app library access, science book completion, and target word engagement. There were no significant effects on curriculum activity completion rates (see Table A2). Additionally, Adaptive Treatment improved motivational engagement, with students showing greater lesson enjoyment and self-competence. Perceived task difficulties were similar between the groups.

RQ2: The Impact of on Student-Teacher Interactions during Synchronous Lessons

Adaptive Treatment significantly improved quality in student-teacher interactions across all three dimensions: engagement, feedback, and questioning ($ps < .001$; Table A3).

RQ3: The Impact on Student Learning Outcomes

As shown in Table 4, Adaptive Treatment significantly improved students' science background knowledge (effect size [ES] = .09) and science content comprehension (ES = .07), particularly on the near-transfer reading passage (ES = .11). There were no significant effects on

[2] Teacher-level covariates include prior intervention experience, years of teaching, national board certification, and state reading course status.

science vocabulary knowledge depth, mid- and far-transfer science content reading comprehension, and domain-general reading comprehension measured by MAP and EOG tests.

Discussion

We explored the effectiveness of structured teacher adaptations in the online Tier 1 content literacy intervention for third graders during COVID-19. These structured teacher adaptations were designed to enhance student engagement with digital and print-based reading activities, thus strengthening the research-practice connection. Consistent with prior research (e.g., Kim et al., 2017), our findings revealed a positive impact of Adaptive Treatment on science background knowledge (ES = .09), science content reading comprehension (ES = .07), and engagement compared to Core Treatment. These adaptations also led to higher quality and frequency of student-teacher interactions during synchronous online sessions, facilitating personalized feedback and responsiveness to student needs. Importantly, we found no significant adverse effects of Adaptive Treatment on any assessed outcomes. Supplemental analyses examining treatment heterogeneity show that effects on reading comprehension are equal across schools (SD of treatment effects = .04, $p = .61$). Thus, adaptations appear to consistently mitigate remote learning challenges for young children by balancing fidelity with contextual and student needs, thus enhancing the program's effectiveness (Goldenberg & Gallimore, 1991).

Our findings have broad relevance for curriculum developers, instructional designers, improvement scientists, and effectiveness researchers. Fidelity and adaption should be viewed as complementary forces that leverage structured teacher adaptations, capturing the benefits of both experimental and improvement science paradigms. When program developers and researchers scale up evidence-based educational interventions to many new contexts, the variability of treatment effects can equal the average treatment effect, highlighting the need for approaches that address this variability (Bryk, 2015). Our structured teacher adaptations achieved higher average student performance and negligible treatment heterogeneity across schools, promoting achievement for all students.

This study demonstrates that structured teacher adaptations enable effective modification and extension of evidence-based programs, enhancing student and teacher engagement and improving learning outcomes. This approach has shown efficacy in other evidence-based programs (e.g., Reading Enhances Achievement During Summer [READS; Kim et al., 2017], Kindergarten Peer-Assisted Learning Strategies [KPALS; Lemons et al., 2014], and World of Words [WOW; Neuman et al., 2011]) in bridging the research-practice divide at scale-up, even in fully remote settings.

The effectiveness of structured teacher adaptations in amplifying student motivation and engagement is evident. While print-based reading and curriculum activity completion rates were similar between the groups, Adaptive Treatment students were more likely to access and complete digital app activities. Adaptive Treatment teachers facilitated student engagement by extending lesson activities, monitoring and encouraging completion, setting clear expectations, and communicating with families. These efforts likely contributed to improvements in measures that capture readers' subjective task values and self-competence beliefs in reading among Adaptive Treatment students (Marinak et al., 2015). Consistent with self-determination theory (Ryan & Deci, 2000), affording greater autonomy, competence, and belongingness support in Adaptive Treatment may foster teachers' intrinsic motivation to modify and improve evidence-based practices that also support their students' motivations to read. These findings align with Kim et al. (2017) and emphasize the importance of defining core components and permissible adaptations, fostering teacher-researcher collaboration.

Adaptive Treatment positively impacted the quality of student-teacher interactions in virtual classrooms. Teachers in this group more frequently and effectively elicited higher-order thinking, engaged in deeper dialogues, provided constructive feedback, and fostered cognitive engagement. This support helped students process complex science concepts, deepen understanding, and build confidence in the virtual environment.

The differential outcomes in interactions likely stem from the TBL-based teacher training provided before the intervention. This training equipped teachers to adapt methods strategically, ensuring alignment with the intervention's core principles and the unique dynamics of virtual classrooms. The TBL approach also fostered peer collaboration among Adaptive Treatment teachers, bolstering their capacity for dynamic and responsive interactions. These findings highlight the critical role of teachers in utilizing local knowledge to bridge research and practice (Kim & Mosher, 2023). Through strategic and scaffolded adaptations of evidence-based practices, teachers improved the intervention's applicability and relevance in virtual settings (Goldenberg & Gallimore, 1991).

The practical significance of the effect sizes in science background knowledge and reading comprehension is particularly noteworthy, given substantial gains among children, 40% from low SES backgrounds, during COVID-19. These improvements are attributed to the evidence-based instructional approach emphasizing schema building and transfer through semantic vocabulary networks and conceptually coherent texts (Fitzgerald et al., 2020; McNamara et al., 1996). Systematic and productive adaptations by teachers, developed with researchers, likely enhanced student outcomes (Vaughn et al., 2020) by promoting active engagement with digital and print-based reading activities.

In contrast to science reading comprehension, Adaptive Treatment did not positively affect domain-general reading comprehension. The reduced instructional hours during the pandemic school closures likely impeded the depth and consistency of reading instruction necessary to improve general reading comprehension measured on standardized assessments (Kuhfeld et al., 2022). Furthermore, we observed no significant effects of Adaptive Treatment on science vocabulary knowledge, indicating that the adaptations did not provide additional gains, as both groups benefited equally from the core intervention components. Moreover, structured teacher adaptations did not result in treatment effect heterogeneity; the approach did not produce varied impacts across schools. In other words, Adaptive Treatment consistently improved students' background knowledge and science content reading comprehension without any discernible negative effects in any school.

Implications and Conclusion

This study replicates findings from other intervention research indicating that structured teacher adaptations can enhance program effectiveness (e.g., 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. Furthermore, professional development programs should equip teachers with the skills needed to design and implement these adaptations.

In conclusion, this study highlights the crucial synergy between fidelity and structured teacher adaptations in educational interventions. Educators can bridge the gap between research-based interventions and classroom dynamics by fostering an educational ecosystem that values this integration. This approach broadens the impact of programs, enhancing their relevance and effectiveness to meet the needs of learners and educators alike.

Table 1

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 books/trifolds	• Teachers introduce the app and books/trifolds to students and parents through a video and letter.	• Core Treatment PLUS • Teachers and literacy coaches at the same school meet synchronously with the research team to set goals, envision student outcomes, and address potential obstacles and solutions. • Teachers employ researcher-guided strategies to increase student engagement[a] • Teachers regularly receive updates on students' app participation from the research team. • Teachers receive summarized goals and adaptations from other schools from the research team for reference.	• 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 deliver 10-day scripted lessons.	• Core Treatment PLUS • Teachers deliver a Day 11 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 2

*Fidelity of Implementation of Teachers in Adaptive and Core Treatment (*N *= 95) for Asynchronous Activities*

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

Note. Items were rated on a 5-point Likert scale: 1 = *never* to 5 = *every day.*
[a]Adaptive-Core treatment group differences are regression coefficients from OLS regression models with fixed effects for school and controls for prior intervention experience, years of teaching, national board certification, and state reading course status.
†$p < 0.10$, *$p < 0.05$, **$p < 0.01$.

Table 3

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

Variable	Adaptive treatment			Core treatment			Adaptive treatment effect	
	M	*SD*	*n*	M	*SD*	*n*	β	*SE*
Digital app activities								
Ever access app library	0.65	0.48	942	0.60	0.49	972	0.05*	0.03
Number of science books completed on app library	0.82	1.60	942	0.47	1.16	972	0.29***	0.08
Number of target words	2.12	1.79	942	1.84	1.74	972	0.24*	0.10
Total time spent on app library activities (min.)	1.72	1.77	942	1.54	1.73	972	0.17†	0.10
Overall app activity accuracy (std)	0.02	1.03	543	0.10	1.00	516	-0.02	0.05
Print-based books/trifolds activities								
Number of trifolds returned	0.97	3.22	942	0.78	2.85	972	0.17	0.14
Return any trifolds	0.11	0.31	942	0.09	0.29	972	0.01	0.01
Motivational engagement[a]								
Enjoyment of lesson activities	2.92	0.69	932	2.86	0.71	966	0.10*	0.05
Reader self-competence beliefs	2.19	0.55	932	2.16	0.55	966	0.11*	0.05
How difficult was the task	2.01	0.39	932	2.01	0.38	966	-0.01	0.05

Note. Point estimates derived from multilevel models, including the Adaptive Treatment indicator, school fixed effects, student demographics, and baseline Measure of Academic Progress (MAP) reading scores.

[a]Students rated each survey item on a Likert scale. Enjoyment: 1 = *I didn't like it*, 2 = *It was ok*, 3 = *I liked it*, and 4 = *I liked it a lot*. Self-competence: 1 = *ok reader*, 2 = *good reader*, 3 = *great reader*. Difficulty: 1 = *too easy*, 2 = *just right*, and 3 = *too hard*.

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

Table 4

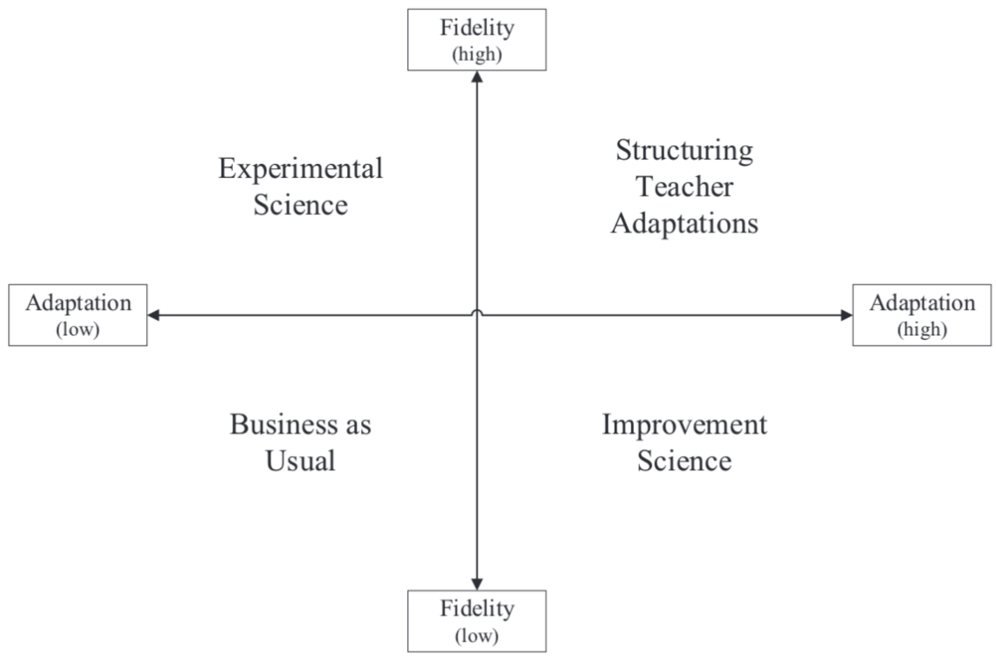
Descriptive Statistics and the Effects of Adaptive Treatment on Student Learning Outcomes

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

Note. MAP = Measure of Academic Progress. EOG = End-of-grade. For researcher-developed measures, we report sum scores for the descriptive statistics and use 2PL IRT-based latent trait estimates for the regression models, standardized to mean 0 variance 1 in the full sample.
* $p < 0.05$, **$p < 0.01$.

Figure 1

A Quadrant Framework for Teacher Adaptations and Fidelity in Interventions



References

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

Bos, S. E., Powell, S. R., Maddox, S. A., & Doabler, C. T. (2023). A synthesis of the conceptualization and measurement of implementation fidelity in mathematics intervention research. *Journal of Learning Disabilities, 56*(2), 95-115. https://doi.org/10.1177/00222194211065498

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.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12

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.

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

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3– 4), 327–350.

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.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*(1), 1–26.

Fitzgerald, J., Elmore, J., Relyea, J. E., & Stenner, A. J. (2020). Domain-specific academic vocabulary network development in elementary grades core disciplinary textbooks. *Journal of Educational Psychology, 112*(5), 855–879. https://doi.org/10.1037/edu0000386

Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N., & Porter, S. (2011). Focus, fiddle, and friends: Experiences that transform knowledge for the implementation of innovations. *Sociology of Education, 84*(2), 138–156.

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

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

Gutiérrez, K. D., & Penuel, W. R. (2014). Relevance to practice as a criterion for rigor. *Educational researcher*, *43*(1), 19-23. https://doi.org/10.3102/0013189X13520289

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest, 16*(1), 3-34. https://doi.org/10.1177/1529100615569721

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*. https://doi.org/10.1037/dev0001710

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

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

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

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

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54-61.

Marinak, B. A., Malloy, J. B., Gambrell, L. B., & Mazzoni, S. A. (2015). Me and my reading profile: A tool for assessing early reading motivation. *The Reading Teacher*, *69*(1), 51-62.

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

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

Michaelsen, L. K., & Sweet, M. (2008). The essential elements of team-based learning. *New Directions for Teaching and Learning*, *2008*(116), 7-27.

Nagy, W. E., & Anderson, R. C. (1999). Metalinguistic awareness and the acquisition of literacy in different languages. In D. Wagner, R. Venezky, & B. Street (Eds.), *Literacy: An international handbook* (pp. 155–160). Westminster, CO: Westview Press.

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

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

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

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

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

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

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

Stanovich, K.E. (2003). Understanding the styles of science in the study of reading. *Scientific Studies of Reading, 7*(2), 105–126. https://doi.org/10.1207/S1532799XSSR0702_1

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

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

Supplemental Online Material

Table A1

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

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

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

[a]Proportion of each demographic category.

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

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

Table A2

Descriptive Statistics and the Effects of Adaptive Treatment on Student Engagement in Intervention Curriculum Activities

Variable	Overall			Adaptive treatment			Core treatment			Adaptive treatment effect	
	M	*SD*	*N*	M	*SD*	*n*	M	*SD*	*n*	*b*	*SE*
Curriculum activities											
Ever access curriculum	0.84	0.37	1,914	0.83	0.38	942	0.85	0.36	972	-0.002	0.02
Proportion of all curriculum lessons completed	0.58	0.40	1,914	0.57	0.40	942	0.59	0.39	972	0.01	0.02
Proportion of Zoom lessons completed	0.64	0.39	1,914	0.62	0.40	942	0.65	0.39	972	0.01	0.02
Proportion of interactive read aloud lessons completed	0.61	0.39	1,914	0.59	0.39	942	0.62	0.39	972	0.01	0.02
Proportion of word sleuthing lessons completed	0.57	0.39	1,914	0.56	0.39	942	0.59	0.38	972	0.01	0.02
Accuracy on curriculum activities (std)	0.03	1.01	1,594	-0.01	1.02	774	0.06	1.00	820	0.001	0.04
Accuracy on end-of-unit quizzes (std)	0.01	1.01	1,176	0.00	1.01	573	0.02	1.01	603	0.03	0.05
Total time spent on curriculum (min.)	4.21	2.06	1,914	4.15	2.10	942	4.27	2.01	972	0.01	0.14

Note. Point estimates derived from multilevel models, including the Adaptive Treatment indicator, school fixed effects, student demographics, and baseline Measure of Academic Progress (MAP) reading scores.

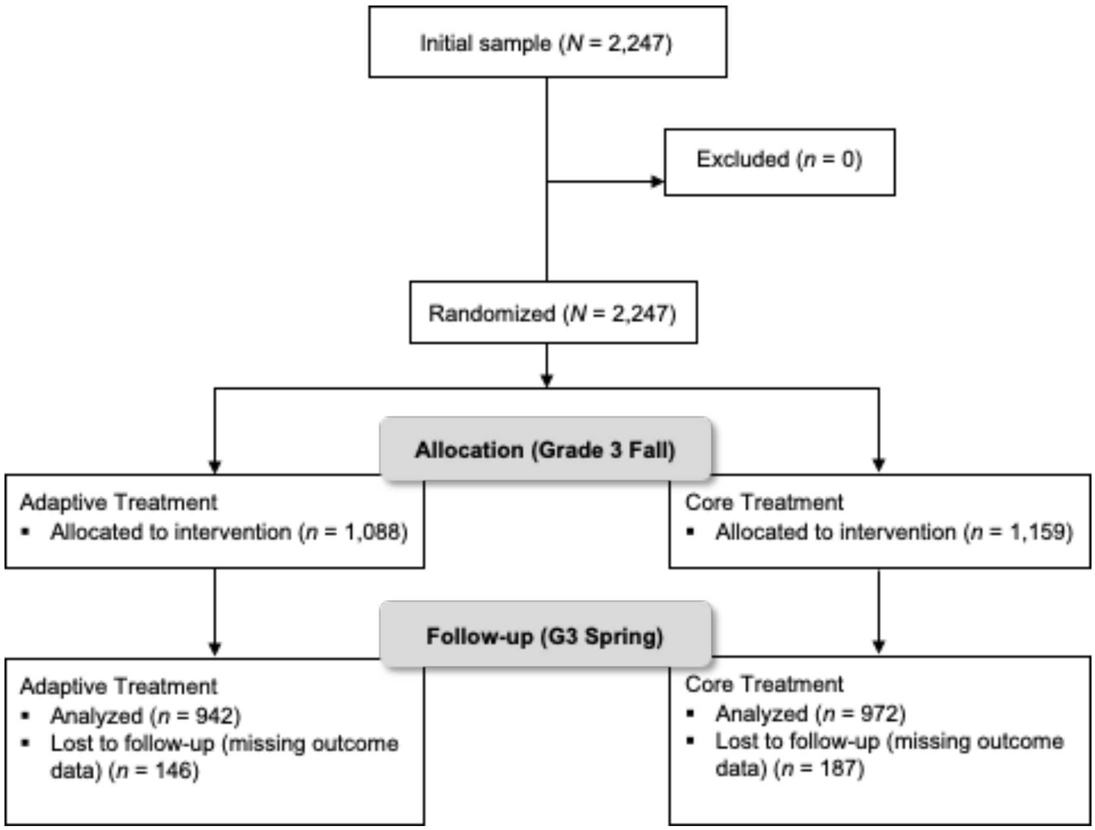
Table A3
Descriptive Statistics and Adaptive Treatment Effect on the Quality of Student-Teacher Interactions

Student-teacher interaction dimensions	Adaptive treatment			Core treatment			Adaptive treatment effect	
	M[a]	*SD*	*n*	M[a]	*SD*	*n*	*b*	*SE*
Engagement	2.95	0.67	23	2.25	0.60	27	0.77**	0.22
Feedback	3.46	0.45	23	2.71	0.68	27	0.80***	0.17
Questioning	3.07	0.74	23	2.26	0.40	27	1.00***	0.20

Note. We employed the following fixed-effects regression model: $Y_{jk} = \alpha_k + \beta_1 Adaptive_{jk} + \sum_{p=2}^{5} \beta_p X_{jk} + \zeta_{jk}, \zeta_{jk} \sim N(0, \sigma_{\zeta}^2)$, where Y_{jk} represents the outcome for teacher *j* in school *k*, α_k denotes the school fixed effect, and β_1 is the adjusted ITT effect of Adaptive Treatment. $\beta_p X_{jk}$ is a vector of teacher-level covariates, including prior intervention experience, years of teaching, national board certification, and state reading course status. ζ_{jk} is the teacher-level residual.
[a]Each dimension was rated on a 4-point Likert scale: 1 = *low*, 2 = *medium low*, 3 = *medium high*, and 4 = *high*. Two researchers, blinded to treatment condition, achieved 94% concordance on four preliminary lessons. One researcher then rated the remaining 149 lessons, and the other independently assessed a random 20%, achieving an interrater reliability of .91.
***$p < 0.001$.

Figure A1

Consort Diagram for the Randomization Process



Note: Disaggregating the attrition rates by experimental condition showed a rate of 13.4% for treatment students (1,088 to 942) and 16.1% for control students (1,159 to 972). A test on differential attrition showed that there was no statistically significant differential attrition between the two conditions ($p = 0.07$).

Appendix A

[BLINDED] Intervention Science Unit Lesson Sequence and Description

10-Day Lesson Sequence of a Science Unit on *Human Body System*

	10 Day Lesson Sequence of a Science Unit on *Human Body System*									
min	1	2	3	4	5	6	7	8	9	10
5	Everybody Guess	Everybody Notice and Wonder	Everybody Guess	Everybody Notice and Wonder	Everybody Guess	Everybody Notice and Wonder	Everybody Guess	Everybody Notice and Wonder	Everybody Guess	Everybody Notice and Wonder
10	Let's Celebrate		Let's Celebrate		Let's Celebrate		Let's Celebrate		Let's Celebrate	
15	Investigate the Word	Let's Celebrate	Investigate the Word	Let's Celebrate	Investigate the Word	Let's Celebrate	Investigate the Word	Let's Celebrate	Investigate the Word	Let's Celebrate
20		Investigate to Apply		Investigate to Apply		Investigate to Apply		Investigate to Apply		Investigate to Apply
25	Pictures to Life		Pictures to Life		Pictures to Life		Pictures to Life		Pictures to Life	
30										
35										

Lesson Activity Description:

- **Activity 1: "Everybody Guess"** served as an introduction to pique students' interest in the content by encouraging them to make a guess related to the topic of "human body system." This activity aimed to foster engagement and inclusivity, encouraging every student to contribute their predictions and setting the stage for further exploration in subsequent activities.
- **Activity 2: "Let's Celebrate"** was designed to recognize and affirm students' efforts in completing App activities related to the content, fostering a sense of achievement and motivation. This activity also provided an opportunity for students who needed assistance to identify themselves, facilitating targeted support in subsequent interactions without disrupting the lesson's flow.
- **Activity 3: "Investigate the Word"** was a vocabulary-focused activity where teachers introduced target words, emphasizing their meaning and form. Through interactive discussions, games, and visual exercises, such as examining images and categorizing words based on their morphological components, this activity aimed to deepen students' understanding of the word's meaning and structure. Teachers utilized Jamboard to co-create concept maps with new words by facilitating collaborative learning and engagement.
- **Activity 4: "Picture to Life"** was a small-group activity conducted in breakout sessions on Zoom. Students applied their knowledge of specific word parts (e.g., micro-) by collaboratively creating speech or thought bubbles for given images. This activity aimed to enhance students' understanding of vocabulary and encourage creative expression.
- **Activity 5: "Everybody Notice and Wonder"** was a warm-up activity where students were encouraged to observe a photograph closely and share their observations and questions. This activity aimed to spark curiosity about the content and lay the groundwork for deeper inquiry in the upcoming lessons, ensuring active participation from every student.

- **Activity 6: “Investigate to Apply”** involved students in collaborative research to explore specific questions related to a topic (e.g., the importance and maintenance of muscles). In this activity, students independently made predictions, engaged in resource-based research, and then convened in small groups to discuss and identify key findings to share with the class, reinforcing their understanding and application of the content.

Appendix B

Extension Lesson Activity Outline

Purpose: Adaptive Treatment teachers provide their students with additional exposure to and increase their knowledge of the word "system."

Goals: The primary goals of this adaptation are:

- To help students laugh and love language
- To get students thinking about the word "system" beyond the immediate context they are studying (i.e., more abstractly)

Activity: Adaptive Treatment teachers will use a "mad libs" activity to get students thinking and writing about different kinds of *systems*. Students will each present on a different kind of system. A "mad lib" activity is a word game where students fill in blanks in a paragraph with different parts of speech (nouns, verbs, adjectives) to create a humorous and engaging paragraph.

Types of Adaptations:

- Adaptive Treatment Teachers implemented the extension lesson at a chosen time within the allocated instructional window, as per their discretion.
- Students' presentation format was flexible, allowing them to choose how they preferred to share their work, with guidance from creative ideas provided in the modules.

Suggested Sequence for Extension Lesson Activity (30-35 minutes)

1. **Engage students:** Begin with a shared "mad libs" activity to engage students, repeating as needed for understanding.
2. **(Re)introduce "System" concept:** Teacher reintroduces the concept of a *system* with an example from the science unit. Teachers make the explanation interactive by involving students in a physical activity.
3. **Mad Libs creation:** Students complete a mad libs exercise about a *system*. Teacher supports English learners by providing word banks and visuals.
4. **Review and prepare presentations:** Students review and edit their paragraphs, add details and multimedia, and prepare for peer presentations, considering fun and engaging presentation methods.

Appendix C

Rubric for Quality of Student-Teacher Interaction

	1 **Low**	**2** **Medium Low**	**3** **Medium High**	**4** **High**
Quality of Engagement • Back-and-forth exchanges • Dialogic interactions • Follow-up questions	Teacher rarely engages in student-teacher interactions.	There are occasional feedback loops—back-and-forth exchanges—between the teacher and students; other times, however, Initiation-Response-Evaluation (IRE) cycles prevail.	There is a mix of back-and-forth exchanges and IRE between students and teachers.	There are frequent feedback loops—back-and-forth exchanges—between the teacher and students, approximating the ideal dialogic interaction model.
Quality of Questioning • Prompting more than one-word response • Prompting higher order thinking	The teacher does not follow the script and ask primarily closed-ended/low-level questions.	The teacher sticks to the script and asks a mix of closed-ended and open-ended questions.	The teacher goes a little beyond the script and asks a few (equal to or fewer than 2) extra-script higher-order questions.	The teacher goes beyond the script and frequently asks open-ended questions and questions that prompt students' higher order thinking skills such analysis and reasoning.
Quality of Feedback • Prompting thought processes • Expansions and clarifications • Connections and integrations • Praise and reinforcement	a) the teacher *rarely* prompts students to explain their thinking for responses; b) the teacher *rarely* provides additional information to expand on students' understanding; c) the teacher *rarely* connects to previously learned concepts; d) the teacher *rarely* praises students' efforts that increases student's involvement.	a) the teacher *rarely* prompts students to explain their thinking for responses; b) the teacher *rarely* provides additional information to expand on students' understanding; c) the teacher *rarely* connects to previously learned concepts; d) the teacher *occasionally* praises students' efforts that increases student's involvement.	a) the teacher *occasionally* prompts students to explain their thinking for responses; b) the teacher *occasionally* provides additional information to expand on students' understanding; c) the teacher *occasionally* connects to previously learned concepts; d) the teacher *occasionally* praises students' efforts that increases student's involvement.	a) the teacher *frequently* prompts students to explain their thinking for responses; b) the teacher *frequently* provides additional information to expand on students' understanding; c) the teacher *frequently* connects to previously learned concepts; d) the teacher *frequently* praises students' efforts that increases students' involvement.

Appendix D

Science Vocabulary Knowledge Depth

It is very important that you listen closely to me read each question and the answer choices. Your job is to choose the two words that best go with the underlined word. I will read all four answer choices, so listen closely. Let's practice!

Practice Question 1: Choose two words that best go with the word **student**.

book fox classroom boat

Choose two words that best go with the word **[target word]**.

Target word	Option 1	Option 2	Option 3	Option 4
1. signal	metal	messenger	transmit	similar
2. skeletal	protection	hair	support	eye
3. repair	surgery	heal	harm	motor
4. organ	sweat	heart	thought	lungs
5. muscular	spine	intelligence	contract	tendon
6. nervous	neuron	ribs	toes	brain
7. diagnosis	treatment	symptom	toothbrush	exercise
8. fracture	crack	attach	solve	fragment
9. structure	wheel	building-block	framework	symbol
10. system	direction	network	cold	interconnect
11. sensory	interact	detect	sleep	visual
12. function	perform	teeth	mystery	operate
13. carnivore	fruit	care	meat	prey
14. extinct	gone forever	living	existing	lost
15. fossil	utensil	bone	folks	footprint
16. hypothesis	thought	poem	guess	hippo
17. brutal	harsh	brave	cruel	friendly
18. evidence	choosing	proving	showing	feeling
19. theory	fairy	theme	argument	idea
20. hunter	handout	catcher	search	conclusion
21. organism	creature	animal	order	practice
22. trait	characteristic	truth	exam	feature
23. paleontologist	paint	history	dinosaur	pair
24. reptile	mobile	lizard	report	cold-blooded
25. potential	future	bones	ability	report
26. unique	characteristic	terrible	careful	different
27. survive	food	dirt	alive	music
28. species	type	family	furry	not real
29. behavior	fur	does	acts	grass
30. camouflage	protect	tell	hide	air
31. advantage	power	finds	follows	helps
32. diversity	large pieces	mix	variety	riches
33. adaptation	freeze	fit	scientist	change
34. habitat	wing	forest	place	weather
35. physical characteristics	claws	looks like	sleeps	acts like
36. complex	hard	quick	upset	problem

Appendix E

Science Content Reading Comprehension Test

“Monkeys” ©[BLINDED]

1) A human is a primate. In primates, the heart sends oxygen in the blood all around the body. The body must have oxygen to function properly. The strong heart muscle contracts to pump blood all around the body. A healthy heart never rests.

2) But what happens when a person’s heart gets weaker? People with weak hearts need help. Scientists have to study a medical mystery. Then they might be able to help people.

3) First, scientists need to diagnose the reasons that a human heart gets weak. They need to do tests. But sometimes it would be hard on a sick person to have tests done. So, scientists study animals that have body systems like humans. Scientists knew that the macaque monkey’s heart is similar to the human heart. So, they did tests with macaque monkeys. The scientists learned what happens when a monkey has a heart attack. After a heart attack, the monkey’s heart muscle has scar tissue where it was damaged. The scar tissue cannot contract like a strong heart muscle can. Then the monkey’s heart cannot pump enough blood to give the body the oxygen it needs. Without enough oxygen, the monkey’s body cannot function properly.

4) Second, scientists wanted to try out new solutions that would keep the heart working. Sometimes the new solutions might fail. So, scientists didn’t want to try the solutions with people right away. They tried a solution with macaque monkeys.

5) Scientists developed an idea. Then, they tested an idea. They injected human stem cells from the human heart into the monkey’s heart. After four weeks the human heart cells grew where the monkey’s scar tissue was. After three months, the monkey’s heart got stronger.

6) Now scientists are using what they found in this experiment to help humans who have had heart attacks.

1. According to the passage, what is an important function of the heart muscle?
    a. To be a part of the nervous system
    b. To help scientists study mysteries
    c. To help primates heal scar tissue
    d. To pump oxygen to the brain

2. According to paragraph 3, why do scientists study macaque monkeys?
    a. Their heart muscles are similar to humans
    b. Their heart muscles are always strong
    c. Their heart muscles pump blood
    d. Their heart muscles have scar tissue

3. What medical mystery are scientists trying to study in this passage?
    a. How to help macaque monkeys stay strong
    b. How to help people after a heart attack
    c. How to find scar tissue in a weak heart
    d. How to find macaque monkeys with weak hearts

4. Why do scientists study monkey hearts?
    a. Because they want to help humans who have heart attacks
    b. Because they want to cover up scars
    c. Because they want to provide a home for monkeys
    d. Because they want to diagnose monkey problems

5. What is the purpose of this passage?
    a. To describe what happens after a heart attack
    b. To help people keep their bodies strong
    c. To tell why it is hard to help sick monkeys
    d. To show how scientists study a medical mystery

6. A heart muscle that functions properly is
    a. not getting oxygen
    b. showing scar tissue
    c. working well
    d. helping scientists

7. The author of this article mainly wants you to learn
    a. How difficult it is to find healthy macaque monkeys
    b. How scientists are learning to help humans with weak heart muscles
    c. How the scientists feel about injected human heart stem cells
    d. How scientists work together to give tests to a sick person

8. Structure means
    a. Someone is strict and follows rules
    b. A heart muscle that is weak
    c. How something is put together
    d. How a heart muscle gets oxygen

9. How did scientists know that they successfully found a way to help humans from a macaque money's heart?
    a. The scar tissue on the monkey's heart grew bigger
    b. The monkey got a heart attack after three months
    c. The monkey's heart with human stem cells became stronger
    d. The monkey's heart did not pump blood to give the body oxygen

10. What is this passage mostly about?
    a. How the heart muscle provides oxygen to the brain
    b. Why scientists want to heal the hearts of monkeys
    c. Who invented the idea of using stem cells
    d. How scientists study other animals to help human hearts

“Birds” ©[BLINDED]

1) Some birds that are here today are smaller than they were many years ago. Why is this happening? It is a mystery. How do scientists study a mystery?

2) First, scientists try to understand the bird’s skeletal and muscular systems. Those systems help birds to fly. One thing scientists know is that birds’ bones are hollow inside. So, they don’t weigh much. Birds’ skeletons weigh less than their feathers. They have also learned that a bird wing works like a human arm. Birds contract their chest muscles to make their wings flap.

3) Second, scientists try to diagnose the reasons that the birds’ bodies might have changed over time. Some scientists thought that birds’ skeletal or muscular systems might have changed. So, they studied migratory birds in North America. Migratory birds fly long distances to warmer places for the winter.

4) Third, scientists measured the physical characteristics of the birds. They measured the size of the birds’ bones. They found some of the birds’ bones were shorter than they were 40 years ago. For example, the lower leg bone of many birds is smaller.

5) Fourth, scientists measured the migratory birds’ wings. They found the birds’ wings are longer than they used to be.

6) Many scientists asked, “Why are the birds’ bones smaller than they used be. But their wings are longer?” They thought about the long distance the birds fly in winter. They thought that the birds’ bodies changed to adapt to the long winter flight. They would need a lot of energy for the flight. A smaller body would have smaller muscles and less fat. But smaller bones and longer wings would make flying easier.

7) Over time, these characteristics of birds have become more common.

8) The scientists came up with some very good ideas about why these skeletal changes might be occurring.

1. What is the main question that scientists are trying to answer in this passage?
   a. Why are birds getting smaller?
   b. How do birds fly long distances?
   c. Where do birds find food in the winter?
   d. Why are birds dying in Chicago?

2. Which sentence best summarizes paragraph 2?
   a. Scientists knew where the birds were flying
   b. Scientists learned that a bird wing was like a human arm
   c. Scientists worked to learn about birds’ bones and wings
   d. Scientists thought muscles carry blood

3. Scientists found that when one part of a bird became smaller, another part became?
   a. smaller
   b. larger
   c. heavier
   d. lighter

4. The author said that scientists are studying the mystery of the shrinking birds. What was one thing the scientists discovered?
    a. Birds use more energy now when they fly
    b. It's easier to fly a long way with longer wings
    c. The leg bone of some birds is getting larger
    d. It is hard to measure birds when they are migrating

5. What is the main idea of this passage?
    a. Many birds have been dying
    b. Birds migrate to warmer climates
    c. Birds change over time
    d. Birds have remained the same

6. In paragraph 3, what is another word for diagnose?
    a. migrate
    b. change
    c. identify
    d. distance

7. The author of this article mainly wants you to learn that
    a. All animals including humans change over time
    b. Birds migrate to a warmer climate every winter
    c. Birds' wings are getting smaller and their bodies are getting larger
    d. Scientists study living things to see how they change over time

8. Paragraph 2 states that "scientists had to understand the bird's skeletal and muscular systems." What does the word systems mean?
    a. illnesses
    b. mysteries
    c. networks
    d. bodies

9. What is this passage mostly about?
    a. How migratory birds' skeletal system works
    b. How scientists measure the size of birds' bones
    c. Where scientists found over 50,000 birds to study
    d. Why migratory birds' bones are changing

10. In order to study the mystery of change in birds' body size, what did scientists do first?
    a. Scientists focused on the birds' skeletal system
    b. Scientists flew to North America
    c. Scientists examined the human muscular system
    d. Scientists studied the environment that birds live in

“Skyscrapers” © READS Lab

1) In many ways, skyscrapers are like the human body. They have bones that keep it tall and strong and skin to protect the body. They have a command center that is like a brain.

2) How do engineers design and build strong and smart skyscrapers? It is very challenging to build a skyscraper because there are many parts that have to work together. If one part is broken, the skyscraper will not work properly.

3) First, engineers need a strong and tall frame that supports the floors and walls. This frame is like the bones of a skyscraper. The bones of a skyscraper are made of concrete, steel and other materials. The columns are like the backbone of a human body. The columns go up. The steel columns hold up each floor. There are also steel beams that go across each floor. The steel columns and beams must be strong to resist the force of gravity. The steel columns and beams must work together to form a strong backbone.

4) Second, engineers need to design a command center. The command center is a computer that sends signals through wires in a skyscraper. For example, it sends signals to help bring fresh air into rooms, just like your lungs. The command center also controls the heating and air conditioning. It helps the building cool down when it’s too hot and warm up when it’s too cold.

5) A skyscraper has many different parts that work together. The bones make the building tall and strong. The skin protects the inside of the building. The skin of a skyscraper can be made of metal, stone, or glass. The brain controls the building temperature.

6) The anatomy of a skyscraper is like a human body. All the parts have to work together. That's how a city skyscraper works properly.

1. According to the passage, why does the author think that skyscrapers are like the human body?
    a. A skyscraper has parts that can get old
    b. The skyscraper has parts that are like strong legs
    c. A skyscraper has parts that are like skin, bones, and brain
    d. The skyscraper has parts that need regular check-ups

2. What is the main idea of paragraph 3?
    a. To describe how engineers work together to build skyscrapers
    b. To describe the reasons why a skyscraper is a tall building
    c. To describe why checklists are needed to build a safe skyscraper
    d. To describe what the backbone of a skyscraper is made of

3. What part of the skyscraper helps control the room temperature?
    a. the skeletal frame
    b. the engineers’ design
    c. the skin covering the walls
    d. the command center

4. What does the author think is special about the beams and columns of a skyscraper?
    a. They are like the skeletal system of a human body
    b. They are like the nervous system of a human body

c. They are made of many different materials
d. They can damage parts of the skyscraper

5. According to paragraph 4, a skyscraper's command center is like the human body's
   a. skeletal system
   b. nervous system
   c. muscular system
   d. digestive system

6. To send a signal can also mean to send a
   a. system
   b. turn
   c. conversation
   d. message

7. Why does the author compare the steel and concrete of a skyscraper to a human skeleton?
   a. Tall buildings have fast and safe passenger elevators
   b. These materials are strong and hold up the building
   c. These materials are easily destroyed by wind and earthquakes
   d. The wires inside the skyscraper are like human bones

8. A skyscraper has a command center. Another word for command is
   a. control
   b. check
   c. help
   d. work

9. What is this passage mostly about?
   a. How steel, concrete, and wires are part of skyscrapers
   b. How engineers design and build tall and safe skyscrapers
   c. How a skyscraper's body is like the structure of a human body
   d. How a skyscraper resists the force of gravity and stands tall

Appendix F

Science Background Knowledge Assessment

[Monkey]

<u>Teacher Directions to Read Aloud</u>: "Today, you will read this passage about Monkeys. Let's begin by answering 3 questions about the passage. I'll read each question and the 4 answer choices. Then, you will choose only 1 answer for each question."

1. Which muscles can you control?
   a. the skeletal muscles
   b. the smooth muscles
   c. the heart muscles
   d. the fiber muscles

2. Which muscles in a monkey never rest?
   a. the skeletal muscles
   b. the smooth muscles
   c. the heart muscles
   d. the fiber muscles

3. What does the heart muscle do?
   a. It controls the nervous system
   b. It pumps oxygen to the brain
   c. It moves skeletal muscles
   d. It helps the body fight the flu

[Bird]

<u>Teacher Directions to Read Aloud</u>: "Now, you will read this passage about birds. Let's begin by answering 3 questions about the passage. I'll read each question and the 4 answer choices. Then, you will choose only 1 answer for each question."

1. Which system works with your muscles to help your body move?
   a. Your immune system
   b. Your digestive system
   c. Your nervous system
   d. Your skeletal system

2. A bird wing works like a human
   a. skull
   b. rib
   c. leg
   d. arm

3. Scientists think that birds' bodies can adapt to fly long distances. What does the word <u>adapt</u> mean?
   a. Shape
   b. behave
   c. change
   d. prepare

[Skyscrapers]

<u>Teacher Directions to Read Aloud:</u> "Now you will read this passage about skyscrapers. Let's begin by answering 3 questions about the passage. I'll read each question and the 4 answer choices. Then, you will choose only 1 answer for each question."

1. Which human body system is like a computer?
    a. the muscular system
    b. the skeletal system
    c. the nervous system
    d. the respiratory system

2. Which part of a skyscraper is like the skeletal system of a human body?
    a. the command center
    b. the steel columns
    c. the heating system
    d. the glass windows

3. The human body has many parts that work together. A group of related parts that work together is called a
    a. system
    b. signal
    c. command
    d. skeletal