EDC.ORG

PROPOSAL

Changing Adolescents' Reading Trajectories: Building Futures with High-Impact Tutoring

Mid-Phase Grant Proposal for the Education Innovation and Research Program

SUBMITTED BY:
Education Development Center, Inc. (EDC)
[redacted]

Principal Investigator:
[redacted]
Ph: [redacted] IEm: [redacted]

In collaboration with:
Reading Futures

EDUCATION DEVELOPMENT CENTER TECHNICAL NARRATIVE

TABLE OF CONTENTS

1. Significance. 1
Current Context: U.S. Workforce Readiness at Risk 1
The Reading Futures Model. 2
Innovating on Reading Futures: Reading Futures+..... 4
2. Strategy to Scale. 6
A. Unmet Demand..... 6
B. Management Plan 7
Responsibilities..... 7
Project Management Approach 8
C. Partner Commitment. 11
D. Plan for Effective Scaling..... 12
E. Dissemination for Scaling 15
3. Quality of Project Design 16
A. Logic Model. 16
B. Goals, Objectives, and Outcomes..... 17
4. Project Evaluation 18
A. Impact Study: Evidence of Effectiveness of RF+ on Student Literacy Outcomes 19
B. Evaluation Plan to Guide Implementation and Replication..... 24
C. Fidelity of Implementation..... 25
D. Replication..... 26
References 27

Changing Adolescents' Reading Trajectories: Building Futures with High-Impact Tutoring

1. SIGNIFICANCE

Education Development Center (EDC), in collaboration with Reading Futures (RF) and in partnership with the Bureau of Indian Education (BIE , Maine Department of Education (MDE), and New York State Education Department (NYSED), proposes a five-year, mid-phase Education Innovation and Research EIR; 84.411B) program grant to innovate upon, rigorously evaluate, and scale RF's virtual, field-initiated, high-impact tutoring (HIT) program for high-need middle-grades students-Tier 3 students in a Multi-Tiered System of Supports (MTSS) framework. **The *Reading Futures*+ (*RF*+) project will address Absolute Priorities 1a and 1b and Competitive Preference Priorities 1 and 2** and will scale a promising strategy to improve reading outcomes, particularly in rural and tribal areas.

Current Context: U.S. Workforce Readiness at Risk

The U.S. is at a critical juncture. Although our workforce must be literate for our country to remain competitive in today's global market, a growing number of U.S. adolescents and young adults lack basic reading proficiency. The most recent NAEP results[1] showed Grade 12 students at their lowest average reading score since 1992: 32% of students exiting high school are reading below a Basic level. Over a third (33%) of Grade 8 students scored below Basic in reading, an increase of 6% since 2019.[2] Many students are locked out of grade-level texts, leading to lower graduation rates, poorer health outcomes, and reduced earning potential.[2,3,4,5] These literacy gaps also carry major social and economic costs, including higher incarceration rates and greater reliance on public assistance.[6,7]

Tutoring is an effective strategy to improve struggling students' reading skills. Robust evidence shows that HIT has the potential to accelerate learning, especially for underserved students.[8,9,10] Despite decades of research demonstrating the efficacy of tutoring,[9,11,12] most positive findings center on elementary grades. Evidence-based programs in middle- and high-

school are rare, and the majority focus on math.[13,14] Implementing HIT in middle schools is challenging due to complex schedules, slower growth rates, and long-standing reading gaps that limit vocabulary and background knowledge.[15] Tutors require advanced skills and extended time to meet students' needs,[16] creating demands that often exceed schools' capacity to address them.

The Reading Futures Model

Grounded in the science of reading and aligned with the What Works Clearinghouse (WWC) practice guide on reading interventions for the middle grades,[14] RF has the potential to benefit all students who require extra support for decoding and reading comprehension. Further, it aligns with the MTSS used by districts to support all students' needs through tiered supports and interventions. RF uses an evidence-based intervention model with metacognitive strategies and techniques derived from the Phonological and Strategy Training (PHAST) program.[17] In a multi-site randomized controlled trial (RCT), the PHAST model demonstrated positive effects on students' word reading, with effect sizes for Grade 6-8 students ranging from 0.56 to 0.78 during a year's intervention.[18] RF also documented promising, positive trends in fluency and comprehension outcomes; a small-scale RCT with elementary school students[19] and internal data with middle- and high-school students showed students' word reading growth exceeded that of middle-schoolers in Lovett's study,[17] with an effect size of .26 on the DIBELS composite score.

Tutoring is most effective when it is frequent, sustained, and integrated into the school day,[8] fosters trusting tutor-student relationships,[20] and provides strong tutor preparation and ongoing support, progress monitoring, a high-quality curriculum, small group sizes, and coordination with school staff.[21] The RF model is intentionally designed to align with this research base, incorporating these proven elements to ensure consistent, high-quality tutoring that supports student growth.

The **RF design meets or exceeds all standards for HIT effective tutoring strategies**, as established by the National Student Success Accelerator.[10] See Table J7 in the Appendix for a

map of RF alignment with the evidence. The model has three additional key elements: (1) **RF tutors are experienced, certified teachers** with deep knowledge of how students learn to read and how to address students' decoding and comprehension difficulties. RF tutors are the spouses of U.S. military personnel and, despite frequent physical relocations due to change of duty station, these qualified teachers are always accessible virtually. (2) **RF is high dosage.** Tutors meet with students for 45 minutes, four times per week, over 36 weeks, aligned with research showing that effective tutoring occurs three to five times weekly and that higher dosage leads to stronger outcomes, particularly for students with dyslexia.[8,22] Sessions take place during the school day, but students typically do not miss subject area classes because they participate during already-scheduled intervention blocks. (3) **RF includes explicit strategies for maintaining strong relationships**. Teacher training focuses on relationship building and connecting with students of all backgrounds. RF incorporates structured routines, journaling, personal progress tracking, and other strategies that foster engagement and motivation, even in virtual and hybrid settings. Teachers are matched with the same students for the entire school year. RF also partners closely with in-school staff, including principals, teachers, instructional assistants, and special education coordinators. By providing weekly instructional updates, monthly data reviews, "deep dives" on individual students, and professional learning (PL) on instructional strategies, RF ensures tutoring is relevant for each student. By combining expert instruction, adequate learning time, and strong relationships, RF aims to produce lasting literacy gains for students not reached by traditional interventions. RF supports students with disabilities through Tier 3 or IEP services; half of its teachers are certified in special education. For students with Specific Learning Disabilities, RF collaborates with the school's special education coordinator on goal-setting and progress monitoring. RF also serves English Learners at WIDA Level 3 ("Developing") and above, offering translation services for parents.

Innovating on Reading Futures: *Reading Futures+*

The innovation of *RF+* lies in blending an established, evidence-based, virtual HIT model focused on word-level decoding (RF) with instructional support focused on informational text (IT .*RF+* tutors will help students move from a focus on decoding to apply foundational word-reading skills to complex subject-area material, enhancing their learning across disciplines (e.g., science, social studies). Once students can decode, they require explicit scaffolding and strategy instruction to transition from basic reading to comprehension of IT, which is essential for success across disciplines.[23,24] This integration of decoding, comprehension, and IT is a next-generation approach to adolescent literacy intervention. We will conduct an RCT to evaluate the impact of *RF+* with a large sample of schools and a cost analysis relative to business-as-usual (BAU) services. We then will explore with participating schools and districts how to further scale *RF+* in cost-effective and sustainable ways and will widely disseminate new evidence on how to use virtual HIT to improve literacy outcomes at a national scale.

This proposal focuses on students from rural and tribal schools. For our state education agency (SEA) partners, the BIE supports 157 schools with middle grades, 86% of which are rural, and approximately 81% of MDE districts and 31% of NYS districts also qualify as rural.[2] Students in rural and tribal schools tend to have lower reading achievement after Grade 3 compared to less rural areas[25,26,27] and often face limited access to qualified teachers due to lower salaries, fewer teacher preparation program placements in rural communities, and difficulty recruiting and retaining teachers in specialty areas.[28,29] Tribal communities may have limited access to print and digital materials due to historical and structural lack of access to resources.[30] *RF+* proposes to expand HIT in rural contexts to meet these challenges.

This project targets high-need students: those with significant word-reading and decoding difficulties identified for Tier 3 intervention whose comprehension is limited by their inability to self-teach new words. Many students lack the word reading, vocabulary, and background

knowledge needed to access grade-level content, often struggling with complex words and comprehension. Yet they also bring curiosity, diverse interests, and, for Native students, rich oral traditions, storytelling, and nonverbal communication that strengthen vocabulary and meaning-making.[31,32,33] These strengths provide a foundation for adding subject-area literacy instruction.

Emerging research has identified the *decoding threshold* as the lynchpin for adolescent reading success.[32] The decoding threshold is the level of word reading below which a student fails to develop their comprehension skills over time[33,34] and that limits students' ability to engage with subject-area texts. *RF+* will provide HIT in word decoding to boost students over this threshold. As students make progress, the program will shift to scaffolding students' interactions with subject-area materials: The incorporation of informational texts will support the acquisition of relevant grade-level vocabulary and background knowledge, as well as comprehension.[35,36] Existing progress monitoring will be used to modify pacing of the intervention.[37] This approach will support implementation and scaling, as *RF+* will develop the materials to allow tutoring to occur during subject area instruction to create more room in students' schedules for this high-dosage model. This research will expand knowledge on improving adolescent literacy by increasing access to grade-level, accessible texts. Although the decoding threshold is well established, no longitudinal studies have examined the long-term effects of interventions supporting students beyond that point.

Absolute Priority (a)-*Promoting Evidence-Based Literacy.* RF is an evidence-based, field-initiated innovation that builds on decades of research showing positive impacts of HIT on reading outcomes for struggling students,[9,38,39] as well as the effectiveness of interventions that simultaneously target multiple literacy skills (e.g., decoding, vocabulary, and comprehension) that align with the WWC (2022) practice guide[14] on reading interventions for the middle grades within a coordinated framework.[18,40,41,42]

Competitive Preference Priority 1–*Returning Education to the States.* We have partnered with the BIE, which supports 157 schools serving middle grades on tribal lands across the U.S., two state education agencies (SEAs -MDE and NYSED-and 40 NYS districts NYS to study *RF+* across diverse policy contexts, curricula, and a variety of rural communities. (See Appendix C for letters of support.) This multi-state approach ensures replicability and sufficient sample size for Tier 3 rural students. State-level partners will collaborate throughout the development, piloting, evaluation, and scale-up; and an Advisory Board including BIE, MDE, and NYSED administrators will guide alignment to local contexts.

Competitive Preference Priority 2–*Expanding Education Choice (HIT).* *RF+* extends an evidence-based tutoring model to address two unmet needs: (1) supporting high-need students who are reading significantly below grade level; and (2) accelerated exposure to grade-level-relevant vocabulary and background knowledge. *RF+* will provide states, districts, and schools with a new option for supporting struggling adolescent readers with a HIT model aligned with effective approaches to tutoring.

2. STRATEGY TO SCALE

A. Unmet Demand

Low adolescent reading ability is an emerging national crisis. NAEP results suggest a third of students enter and leave high school unable to read complex text with understanding.[2] Approximately 15% of adolescents cannot decode well enough to comprehend and make meaning from the material they read.[33] A growing segment of the U.S. population lacks the reading proficiency required for workforce readiness and success. U.S. policy priorities-and thus interventions-have focused on early literacy and dyslexia screening initiatives in Grades K-3.[43] States are only now beginning to turn to adolescent literacy initiatives. Virginia, Minnesota, and Indiana all recently passed legislation to support literacy screening in the middle

and upper grades.[44,45,46] This new wave of adolescent literacy screening will fuel demand for effective, innovative approaches to meet the needs of older struggling readers.

Although tutoring has demonstrated strong potential to accelerate learning across subjects, few evidence-based tutoring models exist for adolescent literacy at scale. Most programs target elementary students or focus narrowly on comprehension strategies, overlooking the foundational decoding challenges that persist for many adolescents[47]. The proposed project directly addresses this gap by developing and testing a scalable HIT model that integrates intensive decoding instruction with comprehension and informational text reading. In doing so, the project introduces a next-generation approach to adolescent literacy intervention-one designed to help students move beyond the decoding threshold and engage meaningfully with the academic content essential for secondary and postsecondary success.

B. Management Plan

Responsibilities

As lead organization, EDC will coordinate the project, administer the grant, and conduct the independent evaluation. **EDC will not provide implementation support or deliver tutoring services and has no financial interest in the outcome of the evaluation**. The organizational structure includes three teams: (1) the EDC project management team; (2) the RF implementation team; and (3) the EDC research team. The EDC project management team will facilitate robust collaboration between the implementation and research teams, while maintaining the firewall necessary to protect the independence of the evaluation. This division of labor ensures the evaluation is unbiased and credible while also leveraging EDC's expertise in managing large federal grants. The organization chart shown in Table J3 in Appendix J displays the relationships between the organizations, and the timeline in Table J5 in Appendix J shows major milestones, timelines, and who is responsible for each project activity.

External Accountability. To reinforce transparency and independence, the research team will report directly to the U.S. Department of Education (ED), and an external **Advisory Board** of literacy researchers, state partners, and education leaders will provide oversight and guidance. In addition to SEA and BIE advisors, the Advisory Board will include [redacted], a national expert in tutoring research, [redacted], an experienced researcher with extensive EIR experience, and [redacted] and [redacted], who bring expertise in adolescent literacy. We will consult with the advisors individually and as a group through online conference calls. Advisors will spend up to two days per year providing technical advice. They will review project activities and progress toward goals, suggest refinements to the RF resources and research design and methods, review emerging findings, provide input on dissemination strategies, review reports and manuscripts prepared for publication, and help build regional infrastructure for scaling *RF+* after the Impact Study ends. RF will form their own **technical advisory board** from nationally recognized literacy experts including [redacted], [redacted] and [redacted], [redacted], the [redacted] and [redacted], a [redacted] and [redacted]. Advisors will guide refinement of the program design, analyze interim findings and data to identify how *RF+* influences students' progress in decoding and comprehension to inform approaches for more effectively supporting both skills. (See Appendix C for letters of support and biographical information for advisors.)

Project Management Approach

EDC, RF, BIE, MDE, and NYSED will create a robust, collaborative, respectful, and open partnership, using regularly scheduled and as-needed virtual meetings, email, and more formal communication of deliverables, timelines, and findings. A **staffing chart** outlining staff responsibilities and skills can be found in Table J2 in Appendix J. Quarterly **timeline and**

milestones aligned to project objectives can be found in Table J5 in Appendix J. A table outlining project goals, objectives, and outcomes can be found in Section 3 (p. 17).

The EDC project management team will provide overall leadership, coordination, and administrative oversight for the *RF+* project, ensuring that all project deliverables, financial reports, and required documentation are implemented on schedule, within budget, and in compliance with ED requirements. EDC is a nonprofit research organization with decades of managing large-scale federal projects, completing them on time and within budget. Members of the EDC project management team are certified in Project Management for Development Professionals (PMD Pro). The project will use Asana to manage tasks, timelines, and deliverables across partners and will house project documents, meeting notes, and data summaries in SharePoint, a secure collaborative workspace. The team will manage subcontracts, partner agreements, memoranda of understanding (MOUs), and data-sharing agreements with districts and schools; monitor progress toward milestones; convene biweekly project management meetings with RF leadership; coordinate quarterly leadership and Advisory Board meetings; and manage reporting schedules and communication protocols. The management team also will oversee communication and documentation systems to ensure transparency, accountability, and efficient collaboration across all partners.

The EDC research team will lead all aspects of the independent evaluation, encompassing the formative studies, impact study, and cost-effectiveness analyses. The research team will be responsible for: recruiting; conducting random assignment; developing and administering data collection instruments and data management plans; securing Institutional Review Board (IRB) approvals; and overseeing data collection and analysis. Only the EDC research team will have access to raw student-level data. EDC will provide aggregated findings from the formative, impact, and cost studies to RF on a quarterly basis and at key milestones to inform planning and continuous improvement. The EDC management and research teams will

coordinate closely to ensure alignment of evaluation activities with project timelines while maintaining the independence of the evaluation.

The RF implementation team will lead the implementation of the tutoring model, PL, and school partnerships. It will adapt its existing tutoring model and PL structures for scaling to additional sites; integrate an informational text component focused on subject-area texts; recruit, train, and supervise tutors and site staff; coordinate with participating schools to align with school literacy goals, assign children to the intervention, schedule and deliver tutoring sessions; manage a digital platform to support consistent implementation; monitor student progress and tutor supports; and refine the tutoring resources and delivery needed based on both internal monitoring systems and evaluation findings to explore the development of hybrid tutoring supports for school staff as part of the scale-up. To expand its tutoring corps for this project, the implementation team will work with its existing partners, the Department of Defense Military Spouse Employment Program and the military-affiliated nonprofit, VetJobs. (See Appendix C for letters of support). Both organizations have provided access to hundreds of certified teachers with experience teaching structured literacy programs.

The BIE, **MDE**, **and NYSED** will partner with EDC and RF, helping the research team to identify and recruit implementation sites and working with both the research and implementation teams to ensure the project and all project materials are salient to local communities, aligned with state, district, and school priorities, and implemented with fidelity across schools. Representatives from both agencies also will serve on our Advisory Board. All teams and partners will contribute to strategic dissemination activities.

Key Personnel. Project staff are highly qualified to carry out the proposed work and bring extensive experience in literacy learning and instruction, adolescent literacy tutoring, technology and product development, strategic dissemination, qualitative and quantitative research, and managing large-scale collaborative research and scaling efforts. (See Appendix B).

[redacted] (PI; 35%Y1, 40% Y2-Y5 will co-direct the EDC research team, lead the impact study design and randomization, and coordinate the fidelity of implementation data collection, analysis, and sharing sessions. [redacted] (co-PI; 35%Y1, 40% Y2-Y5) will co-direct the EDC research team and oversee data collection and analyses for the quantitative impact analyses and lead the cost effectiveness study. [redacted] (co-PI, 30%) will lead the EDC project management team as Project Director and oversee all project management activities, including high-quality and rigorous activities and budget management. [redacted] (senior methodologist; 15% Y1, 25% Y2-Y4; 20% Y5) will lead the quantitative data analysis, serve as lead methodologist, and jointly oversee the research staff. [redacted] will coordinate the formative research. [redacted] (external liaison, 5%) will oversee the Advisory Board and research, policy, and practice communities on dissemination products. [redacted] (dissemination lead, 10%) will coordinate outreach and dissemination.

Other Personnel. [redacted] ([redacted]) will lead tutoring activities and implementation. [redacted] ([redacted]) will lead program development and PL. [redacted] ([redacted]) will lead RF's tutor cohort and tutor mentors. [redacted] [redacted]) will oversee RF's approach to measurement of the decoding threshold and related progress monitoring. (See Appendix B for details on key personnel.)

C. Partner Commitment

The project brings together partners with deep and complementary expertise in adolescent literacy, tutoring, and educational research. (1) **EDC** brings decades of experience conducting and disseminating rigorous evaluations of large-scale literacy initiatives, including leadership of multiple federally funded RCTs that have advanced the national evidence base for improving learning, including literacy. EDC also contributes extensive expertise in cost and implementation analysis, research-practice partnerships, use of evidence to inform education policy and systems change, federal grant management, and strategic dissemination. (2) **RF**

contributes nationally recognized expertise in designing and implementing evidence-based literacy tutoring programs for middle-school students with severe reading difficulties, integrating structured decoding instruction, supported by Pearl Tutoring Management System, a digital tutoring platform, student progress monitoring systems, and ongoing PL for tutors. (3) Our three **SEA partners** bring policy and systems-level expertise to guide alignment of *RF+* with statewide literacy and tutoring initiatives, ensuring the program's feasibility, relevance, and sustainability in diverse contexts. Finally, in reflection of the need for effective adolescent literacy interventions like *RF+*, 40 school-district partners signed on to support this project, pledging significant in-kind contributions to support implementation of the tutoring at their school sites. Together, EDC, RF, and the SEAs bring a unique combination of rigorous research capacity, deep implementation experience, and strategic leadership to advance the design, evaluation, and scaling of effective tutoring for adolescent readers.

D. Plan for Effective Scaling

Scaling is embedded throughout the project. In Years 1-4, the *RF+* HIT program will expand to 66 schools in NYS, ME, and the BIE with supports for subject-area reading, with each adaptation designed to promote scalability in new subject-area contexts (e.g., science, social studies). Year 5 will then focus on translating these refinements into sustainable, large-scale implementation models. Working in collaboration with our Advisory Board, we will leverage the relationships and partnerships developed with schools as part of implementation to develop capacity in the school systems we serve, setting the stage for affordable and effective scaling practices that are suitable across the participating states, including rural and tribal schools.

Although HIT has a strong evidence basis in education,[9] practices and programs that produced significant results at a small and local scale often failed to replicate those results at significant scale,[48] due to inconsistent fidelity, lack of integration with school structures, and insufficient local capacity.[8,20] Particular barriers to adolescent literacy HIT scalability include a

less relevant and interesting program for older students and less flexibility implementing the model during subject-area classes. The *RF+* scaling plan addresses these challenges using ongoing structured feedback loops to refine the model while ensuring fidelity and aligned adaptation across diverse contexts. Rural schools often lack qualified staff to deliver intensive literacy tutoring. RF currently addresses this gap through virtual tutoring, and the project will examine whether school-based staff can be trained to provide HIT or whether alternatives such as pooling district resources can work. The cost-effectiveness analysis will further inform scalable models, addressing the challenge that tutoring requires external funding.

Our plan to scale RF tutoring for middle school students thus draws on Coburn's four dimensions of scale-**depth, sustainability, spread, and ownership**-to develop strategies to address these challenges directly.[49] Our focus on **depth** will be ensuring that tutors internalize the principles of *RF+*, including decoding support, dynamic correction and individualization, vocabulary instruction, and the use of informational texts. Tutors will be supported by highly qualified and experienced teachers through ongoing PL, coaching, and progress monitoring systems so that they can flexibly apply the model while maintaining fidelity of implementation. RF's virtual tutoring model is well-suited to effective mentoring relationships as tutors and tutor mentors can participate in an online session in real time and, later, review a recording of the session to discuss specific instructional moves. RF will explore using these recordings to create a library of instructional excellence to support initial training and continuous improvement at scale. Finally, we will also assess how school structures and supports contribute to the depth of RF+ implementation.

Spread will involve expanding *RF+* both outward to new tutors, schools, and districts, and inward by strengthening tutors' day-to-day practices. In terms of outward spread, the evaluation's mediator and moderator analyses will generate actionable findings to help RF make decisions about components to scale up or student populations to target. By identifying the

balance between fidelity and flexibility, we will ensure that *RF+* is both effective and adaptable, supporting successful replication across different students, schools, and districts.[49] RF has been scaling slowly, hiring extremely selectively, training carefully, and creating a culture of instructional excellence. RF is **poised to accelerate the pace of scale-up**: Of the 20+ teachers serving as tutors at RF, over half have the potential to serve as mentors in the timeline of this study, each capable of mentoring and supporting an additional ten teachers. Spread will be supported through adaptable training modules, coaching-based feedback systems, partnerships with literacy experts to expand informational text resources, and use of the Pearl digital platform for consistent implementation. This approach will promote shared values, mentorship, and sustained instructional improvement.

To ensure *RF+* can continue to scale **sustainably** and effectively, RF will integrate EDC's research tools for monitoring fidelity and quality into its infrastructure to sustain continuous improvement and accountability to grow the program and ensure long-term adoption at scale. Long-term sustainability will involve embedding *RF+* into partner district systems and funding streams, with SEA partners helping identify scalable staffing models to sustain tutoring. At the school level, RF will collaborate with schools to develop protocols for aligning tutoring with districts' MTSS frameworks, existing schedules and school literacy instruction routines. Collaboration with state and school administrators will root *RF+* in sustaining structures and will ensure that local perspectives shape program refinements.

To support sustainability and scale, the project will explore **transitioning program ownership to districts and states**. RF will develop and adapt *RF+* with district leaders and school-based coordinators in mind, training local staff to monitor fidelity and outcomes, integrate tutoring into existing PL and Tier 3 literacy supports, and embed *RF+* practices into district- and state-level literacy strategies. Our SEA partners on the Advisory Board will help shape

implementation and dissemination of *RF+* to fit their local contexts, positioning them to sustain and replicate the program post-grant.

E. Dissemination for Scaling

EDC and RF will lead dissemination with vital contributions from our SEA partners, including sharing of findings via the Accelerate Literacy community of practice. Our plan has **two goals**: (1) reach and inform education decision-makers and policymakers who can propel replication and scale-up of *RF+* to benefit students nationwide; and (2) Share evaluation findings to accelerate the work of other tutoring programs and researchers. We are dedicated to raising visibility of this important work for ED. If awarded, with ED approval, partners will publish press releases on our websites and spotlight the work on social media.

Key Audiences and Mechanisms. Our primary audiences for dissemination are *RF+* study participants; other education leaders (e.g., SEAs, tribal education leaders, school district leaders, other administrators); education administrators and policymakers at tribal, county, city, state, and national levels; other HIT program developers; and education researchers. RF will lead efforts to share and scale the RF+ model itself, in addition to EDC's focus on sharing research findings. To support adoption by schools and districts, RF will develop practical implementation resources and engage its national network and state partners showcase effective practices and lessons from the project. Our tactics will include: **1) Briefing Materials (*RF+* participants, education decision-makers, and policymakers .** We will share project findings in clear and concise briefs, summaries, and infographics for busy people. **2) Project Website (all audiences).** We will launch a project website that EDC's Digital Design Group and website development team will review to ensure optimal 508 compliance, user experience, and content. **3) Presentations (education decision-makers, policymakers, other tutoring programs, researchers).** EDC, RF, and the SEAs will plan to co-present in Year 2 (one event); Years 3 and 4 (three events per year); and Year 5 (four events) at convenings of the Council of Chief State

School Officers; American Association of School Superintendents; National Indian Education Association; National Student Support Accelerator (HIT Summit); Accelerate; National Rural Education Association; state chapters/affiliates of the International Literacy Association; Campaign for Grade-Level Reading; International Society for Technology in Education; and Society for Research on Educational Effectiveness. **(4) Articles and blog posts (same audiences as 3).** In Year 1, we will publish a project profile on EDC.org. In Years 2 to 5, we will publish one blog post per year. Possible host blogs include *ISTE+ASCD*, *EdSurge*, *K-12 Dive*, *The 74*, *Literacy Today*, and *Edutopia*. By the end of the project, we will publish at least two articles in target journals such as *Educational Leadership*, *The Reading Teacher*, and *Reading Research Quarterly*. **S) Social Media Spotlights (all audiences).** We have drafted a social media strategy and will use EDC's social media platforms to raise visibility of *RF+* and to highlight project milestones through carousels with findings and brief videos.

Open-Source Products. EDC and RF are committed to reducing barriers to accessing research data and findings. All articles will be published open-access, and all blog posts will be readily accessible. RF's curricula, PL, and manuals all are existing intellectual property, though components (e.g., TopicReads) are already freely available for use in schools. RF commits to publishing an open access "instructional library" with video clips from exceptional tutoring sessions, as well as a practice guide on monitoring students' progress against the decoding threshold and relevant instructional decisions.

3. QUALITY OF PROJECT DESIGN

A. Logic Model

Appendix G includes the logic model guiding the project. While the key components, mediators, and outcomes include school, tutor, and student levels, we are focusing only on student-level outcomes in this project. Key components of *RF+* include certified reading teachers who tutor students virtually four times a week for the duration of the school year, using

a curriculum based on the PHAST strategies to build decoding skills and the new informational text curriculum to be developed for this project to bridge the gap between early reading and text comprehension. Pre-service training, as well as coaching, mentoring, and supervision during implementation and regular student progress monitoring informs responsive instruction and a flexible instructional plan. Key short- and medium-term outcomes for students include improved word decoding ability, reading comprehension, and informational text skills; long-term outcomes include higher graduation rates. Key mediators for student-level processes include student engagement in the sessions, time spent on decoding and reading, and the quality of the student-tutor relationship.[21] Key moderators include district and school policies and supports for literacy, school climate, the school's core literacy curriculum and technology infrastructure, tutor experience teaching, and students' grade level, baseline reading ability, and students' IEP status.

B. Goals, Objectives, and Outcomes

The overall purpose of the project is to improve word decoding skills, reading comprehension, and informational text skills for high-need middle-grade students (particularly those in rural and tribal areas) using a HIT program, *RF+*. The project has four primary goals.

Goal	Objective	Primary Outcomes
1. Innovate on *RF+*	Add subject-area supports to the existing RF program *RF+* and pilot in 10 schools (~120 students)	Pilot report with recommended revisions to materials and implementation, particularly improving comprehension and informational text skills beyond the decoding threshold
2. Implement *RF+* at scale	Roll out *RF+* in 22 districts (~800 students) over 2 years	Data on fidelity, cost, feasibility, and effectiveness across contexts; inform Y5 revisions and continued regional scale-up in BIE, MDE, and NYS schools

3. Evaluate effectiveness of *RF+*	Conduct RCT with 66 schools (~800 students) over 2 years	Evidence on efficacy, cost-effectiveness of *RF+* vs. BAU; data on contextual variations-who benefits and under what conditions?
4. Scale *RF+* as alternative model for SEAs	Partner with BIE, NYSED and MDE to explore *RF+* delivery models (e.g., virtual tutoring, in-school mentoring)	State/district adoption of best-fit *RF+* models; dissemination of impact and cost-effectiveness data to guide policy and practice

These goals align with Absolute Priorities 1a and 1b and Competitive Preference Priorities 1 and 2 in their focus on promoting evidence-based literacy instruction, particularly by expanding education choice for states and districts via a high-impact tutoring model. Our scale-up plan is focused on returning education to the states by taking the data from our formative research and impact study and developing alternative models for education agencies to select for and receive appropriate support in implementation.

4. PROJECT EVALUATION

EDC's evaluation will be guided by the program logic model (Appendix G and will have three stages. It will begin with an iterative design-based research approach,[50,51] organized around iterative cycles of design, implementation, evaluation, and refinement. This will be followed by a pilot of the *RF+* modules and resources, and a two-year Impact Study to estimate *RF+*'s effects on student literacy outcomes.

Innovation Design. Our formative evaluation will focus on the development of the informational text resources, synthesizing insights from experts and research evidence about the supports and practices struggling readers need to improve subject-area literacy skills. We will iteratively test the initial intervention with students and tutors, providing feedback to RF to adjust the design based on teacher and student experiences. We will adapt tools from the

Tutoring Survey Instruments[52] and Assessing ELA Curriculum Shifts Toolkit,[53] including teacher, student, and administrator surveys and PL and classroom observation protocols designed to capture curriculum implementation in ELA Grades 3-12.

Pilot Study. In the first project year, researchers will pilot *RF+* with 10 schools (approximately 120 students and four tutors) to refine study instruments, data collection, and communication processes. The pilot will document typical literacy supports in Tier 2 and 3 classrooms to ensure the Impact Study measures capture all aspects of implementation and will test methods for identifying when students move beyond decoding to focus on subject-area comprehension.

A. Impact Study: Evidence of Effectiveness of *RF+* on Student Literacy Outcomes

After the first year of formative research and piloting, and following adjustments to the final intervention design, the summative component will examine the effect of *RF+* on students' literacy outcomes, using an experimental design that meets WWC Standards *without reservations*.[54] The two-year Impact Study will examine whether high-quality tutoring improves literacy outcomes for middle-school students with Tier 3 reading difficulties, and for whom, how, and under what conditions it is most effective. It will provide evidence about quality of implementation at a larger scale, and how implementation mediates the impact of tutoring on achievement.[55] Findings will inform education leaders, policymakers, and researchers and will advance knowledge on designing and scaling effective tutoring for adolescent readers. The following **research questions** will guide the Impact Study.

1. **(Impact)** What is the effect of *RF+* on Tier 3 student achievement in literacy, including decoding, reading comprehension, and informational text skills, compared to a control group that is not receiving high-intensity literacy tutoring?
2. **Implementation)** To what extent are tutors able to implement the *RF+* model across contexts, what adaptations do they make and why, and what additional supports do they need?

3. **(Exploratory/Mediators**) Do initial (one year) and sustained effects (over two years) differ by the tutor implementation of the *RF+* model-specifically, tutor instructional quality, relationship with student, and time spent decoding and reading? Do initial and sustained effects on comprehension differ by students' decoding abilities?
4. **(Exploratory/Moderators)** Do the initial and sustained effects of the intervention differ by student, tutor, or school characteristics, including (a) students' entering literacy scores, grade, demographic characteristics, prior years' attendance, or disability status; (b) tutor teaching experience; and (c) school supports for literacy and climate?
5. **(Cost Effectiveness Analyses)** What is the cost-effectiveness of *RF+* compared to the literacy instruction received by students in the BAU condition?
6. **Scaling)** What are the successes and challenges of the scaling strategies? As *RF+* scales across diverse sites, how do depth, sustainability, spread, and ownership of the tutoring model emerge, and what contextual factors support or hinder these dimensions of scale?

Sample and Statistical Power. Based on our power analyses, assuming a three-level block randomized design blocked by district and randomized at the school level within districts, a sample of 22 districts, 66 schools, and approximately 790 students (accounting for 10% attrition) will allow us to detect a minimum detectable effect sizes (MDES) of approximately 0.30 standard deviations, on par with effect sizes found in recent studies[9,38,39] see Appendix J4 for the power analyses) and of sufficient size to inform scaling strategies, given the focus on students performing several years below grade level and the depth of support. RQ3 and RQ4 are exploratory questions, as we do not anticipate having sufficient power for these sub-group analyses. Researchers will screen all students who have been identified by their school as needing Tier 3 support using the CAPTI ReadBasix decoding subtest[34] (see "Measures" section below) to assess their decoding ability. Students who score below the mastery cut point (identified during the pilot study) will be enrolled in the study with parental option to opt-out.

English Learners at WIDA Level 3 ("Developing") or higher will be eligible. Students with disabilities will be included when the intervention aligns with a student's IEP goals; participation decisions will be made collaboratively with school staff.

Recruitment Criteria. We will use proven procedures to recruit schools in collaboration with our SEA and district partners. Our criteria for schools to participate include the following: (1) public school (including charter and magnet schools) that serves Grades 6 and 7; (2) not already using HIT for literacy, to ensure a sufficient contrast between treatment and comparison conditions; (3) at least six eligible students, as defined below; (4) a commitment to random assignment; and (5) willingness to share student assessment and demographic data. To balance statistical power while also ensuring representation of a range of contexts, recruitment will target districts and schools that vary in key factors influencing tutoring effectiveness: literacy curriculum,[10] literacy policies, such as instructional time and administrative support,[56] and school average achievement. (See Appendix D.) We will recruit diverse students and districts to enhance generalizability and will document school and tutor characteristics to clarify the populations to whom findings apply.

Treatment and Control Conditions. The study design will use a three-level block randomized design, blocked on district (Level 3), randomized at the school level (Level 2), and with students as the unit of analysis (Level 1). Schools will be assigned to a study condition after entering into an MOU developed in collaboration with districts that reflect shared commitment to the study's goals. We will randomly assign half the schools to participate in *RF+* tutoring in Year 1 of the Impact Study and half to tutoring starting in Year 2. All Tier 3 Grade 6 and 7 students who meet screening criteria will participate in data collection over two years, including as they advance to Grades 7 and 8. To preserve a BAU comparison group for each cohort, only Grade 6 will receive the delayed treatment during Year 2, while Grades 7 and 8 continue to serve as the control. The delayed treatment will expand to Grade 7 after the end of the study. Students

in the treatment group will receive *RF+* four times per week for 45 minutes over two academic years. Students in the comparison group will receive BAU literacy instruction and their schools' usual Tier 3 supports. We will preregister this design with the Registry of Efficacy and Effectiveness Studies (REES) prior to study start.

Measures. See Appendix J4 for a detailed description of measures, including psychometric properties. Because participants come from different states and state assessments may not accurately capture gains for low-achieving students,[57] we will measure student outcomes using the CAPTI ReadBasix, an adaptive, web-based assessment of foundational reading skills designed for middle- and high-school students with evidence of reliability and validity[58] consistent with WWC standards for effectiveness studies. The measure provides scores across six subtests: word recognition and decoding, vocabulary, morphology, sentence processing, reading efficiency, and reading comprehension. Because existing measures of informational text skills rely on open-ended questions impractical for large-scale RCTs, we will adapt the Informational Text Skills Instrument (ITSI ,[59] a 15-minute, 37-item assessment of students' ability to use text features to locate information. We will pilot and evaluate the adapted ITSI for middle-grade students. For **mediators**, to understand the relationship between decoding and comprehension, we will use students' scores on the ReadBasix decoding subtest. The fidelity section details the approach for measuring overall RF implementation fidelity; administrator and tutor interviews and surveys will examine school and tutor experiences, including satisfaction, supports, communication, and adaptation needs, challenges, and rationales. To measure tutor instructional quality and time spent decoding and reading, and student engagement during sessions, we will use practice logs, validated through classroom observations, and measure tutors' relationships with students through student post-surveys from the Tutoring Survey Instruments[55] at the end of each academic year. To measure the influence of **moderators** on any impacts, researchers will collect district demographic data (gender, EL status, attendance, IEP

status) and administer pre-surveys to tutors (experience, degree and students (school climate, using the Tripod School Climate Index[TSCI]).[60] School context will be assessed through state test scores and administrator surveys on literacy curriculum, school climate (teacher TSCI ,[61]technology, and literacy schedules.

Data analysis. All analyses estimating student outcomes will follow an intent-to-treat estimate of the average impact of the treatment for each unit offered in the intervention, regardless of whether schools and teachers implemented, or students received, the intervention.[61] Our analysis of **student outcomes (RQ1)**, will use a three-level regression model, with fixed block effects for districts with controls for student- and school-level covariates and tutor experience at baseline aggregated to the school level. To explore the influence of **moderators (RQ4** , we will use the same model as used for the confirmatory RQs, but we will introduce interaction effects for each of our covariates of interest. To explore **mediators (RQ3)**, we will use causal mediation analysis.[62] We will adjust for multiple comparisons for these moderator analyses using the Benjamini-Hochberg method. We will examine **implementation fidelity (RQ2** using descriptive and multivariate analyses, including means, correlations, cross tabulations and ordinary least squares regressions and multilevel analyses. For the **cost-effectiveness analysis (RQS)** we will use the ingredients method[63] to document the costs of *RF+* implementation and how costs vary across sites and numbers of teachers or students served. Analyses for the **approach to scaling (RQ6)** will be examined using a grounded theory[64] qualitative analysis aligned to the logic model and framework.

Attrition. To minimize attrition and maintain internal validity, we will maintain clear communication with administrators and teachers regarding study goals, timelines, and participation requirements. Along with incentives for schools to participate in the research activities and free tutoring, the research team will prioritize relationship-building and flexibility throughout the study. Designated site coordinators will maintain ongoing communication with

schools and tutors, serving as trusted points of contact to troubleshoot challenges and support engagement. The team will adopt an adaptive approach, remaining flexible with data collection windows and responsive to school requirements as feasible. We will assess overall and differential attrition following WWC (2022) standards and will conduct sensitivity analyses using multiple imputation to evaluate the robustness of findings to missing data.

B. Evaluation Plan to Guide Implementation and Replication

Our team is committed to collaborative research and development, continuous improvement, accountability, and transparency. This project will use iterative cycles of design, implementation, evaluation, and refinement to ensure the intervention and research are both evidence-based and useful. The team will document inflection points, decisions, and evaluation results at each stage, and the Advisory Board will serve as an external review panel, providing independent feedback at critical points in the research, and will establish goals and benchmarks.

Initial formative research will guide the adaptation of the existing RF model to incorporate the informational text "expansion pack." A pilot study with 10 schools will evaluate the resulting *RF+* program, driving any necessary improvements and informing procedures, ways to strengthen fidelity, and paths toward scalability. A summary brief will document these findings to serve as potential guidance for future modifications, and any changes to our design or methodology will be noted in our REES pre-registry. The Impact Study then will evaluate the final *RF+* program, with insights from implementation informing adjustments to the model for implementation with the control group and the Year 5 scale-up. Impact study results will generate insights into how *RF+* performs across different contexts and student populations to understand how the program fits within varied school settings, what supports are needed for effective implementation, and how impacts may differ by site or subgroup. Findings will help schools and districts understand how to implement *RF+* most effectively and sustainably.

C. Fidelity of Implementation

For the intervention group, fidelity of implementation-the extent to which implementation is consistent with developers' intensions-to interpreting program impacts.[65,66] We will assess fidelity across adherence, quality of delivery, dosage, participant engagement, and adaptation.[67] Our analysis of fidelity will examine how tutors used the core components[68] of *RF+* as described by our logic model (Appendix G) and will collect data about implementation to (1) document the extent to which inputs and activities (e.g., PL and coaching and tutoring activities) are implemented as intended and adopted with fidelity; (2) provide treatment-control group contrast to determine that groups are sufficiently different; (3) provide examples and illustrations of high-quality tutoring instructional practices and student experiences; and 4) inform refinements of project inputs. Understanding the implementation of intervention components will support the interpretations of findings, inform program adjustments, and lead to insights about scale, including how and where the intervention may be most effective.

FOI Matrix. We will develop a fidelity of implementation (FOI) matrix with RF and state and district partners following the U.S. Department of Education's FOI framework. In the first year we will develop a detailed plan to measure implementation of each of the key components described in the theory of change, including processes, strategies, and materials required to support *RF+* implementation and intervention components and the moderating factors aligned with characteristics of high-quality literacy tutoring, and will set a threshold for acceptable implementation. Core measures will include time students engage with the tutor, use of RF curriculum, tutor instructional quality, content and skills of focus, and time spent decoding and reading. Data sources include weekly instructional logs completed by tutors and teachers, classroom observations of a subset of 16 teachers and tutors using an adapted observation protocol designed to measure special education teachers' literacy instruction with trained raters achieving 2:70% reliability[69] and RF back-end tutoring session log data. We will triangulate

observation and log data to validate self-reports and generate a composite fidelity score for each tutor. Using the FOI matrix, supplemented by the weekly logs as required, we will generate a fidelity of implementation score for each tutor and teacher in the control group. Fidelity scores will be used to document implementation quality, confirm treatment-control contrasts, and explore whether higher fidelity is associated with stronger student outcomes.

D. Replication

This project is guided by the understanding that replication and open science practices are critical to research, evidence-based practice, and policy.[70] We will attend to replication in development, implementation, reporting, and dissemination. First, **we will develop and implement materials and research with replication in mind**. We will prepare structured training and manuals that can be reused and easily adapted to align with contexts. Tutors will use structured lesson protocols that other districts can adopt, also with minimal adaptation. Formative and qualitative data will inform our rationale for adaptations, and we will document the adaptations we have made for different contexts, those that support effective outcomes, and less effective adaptations to avoid. The research design, methods, and analysis will be pre-registered on REES with sufficient detail for replication.

The Impact Study will explore and document fidelity and context. We will collect fidelity logs, tutor training records, and school-level contextual data to identify how the program was implemented and which conditions support success. The study will include a varied sample in terms of school, student context, and background characteristics. The evaluation will yield generalizable, scalable evidence for replication studies, including **valid and reliable estimates of effect sizes** for decoding, comprehension, and informational text skills. Moderator analyses will show which student subgroups benefit most and will inform replication studies of potential contextual effects. At project end, **we will share de-identified data and analytic code** on an open data-sharing repository such as OSF https://osf.io/) to permit others to repeat our analyses

or ask new questions from the same data. Ultimately, these analyses and the project will strengthen the capacity of *RF+* to scale effectively across diverse educational contexts by identifying the practices and conditions that promote consistent, high-quality implementation and improve student reading outcomes in collaboration with state and district partners.

REFERENCES

[1] NCES. *NAEP Reading Grade 12.* 2025.

[2] NCES. *NAEP Reading Grade 8.* 2024.

[3] Carlson CL. Adolescent literacy and economy. *J Educ Hum Dev.* 2013;2(1 :1-8.

[4] Hernandez DJ. *Double Jeopardy: Reading & Poverty.* Annie E. Casey Fdn; 2011.

[5] Krueger PM et al. Mortality and education. *PLoS One.* 2015;10(7):e0131809.

[6] NCES. *Dropout and Completion Trends.* N.d.

[7] Rothwell J. *Eradicating Illiteracy Gains.* Brookings Inst; 2020.

[8] Kraft MA, Falken GT. Scaling tutoring and mentoring schools. *AERA Open.* 2021;7:1-8.

[9] Nickow A et al. Tutoring meta-analysis. *Am Educ Res J.* 2023;61(1 :74-107.

[10] [redacted]

[11] J-PAL Evidence Review. *Tutoring for PreK–12 Outcomes.* 2020.

[12] White S et al. Tutoring implementation review. *Annenberg WP.* 2023;22-652.

[13] [redacted]

[14] What Works Clearinghouse. *Reading Interventions Grades 4–9.* IES; 2022.

[15] Cirino PT et al. Reading skills in struggling readers. *Read Writ.* 2013;26(7 :1059-1086.

[16] National Student Support Accelerator. *Tutoring at Scale.* Stanford Univ; 2023.

[17] Lovett MW et al. Putting struggling readers on the PHAST track. *J of Learn Disab*. 2000;33(5):458-476. doi:10.1177/002221940003300507

[18] Lovett MW et al. Reading interventions. *J Educ Psychol.* 2021;113(4).

[19] Cassady J et al. *RF virtual reading tutoring program:2-year effects.* Ball State Univ; in-press.

[20] Grossman JB, et al. School-based mentoring duration and re-matching effects. *Am J Community Psychol.* 2011;49(1-2):43-54.

[21] Robinson CD et al. *Design Principles: High-Impact Tutoring.* Annenberg Inst; 2021.

[22] Donnelly PM, et al. Intensive summer intervention and reading. *Front Psychol.* 2019;10:1-10.

[23] Patton SA et al. Comprehension of texts. *Learn Disabil Res Pract.* 2022;37(2):124-139.

[24] Ritchey KD et al. Informational text intervention. *Learn Disabil Q.* 2017;40(2):68-80.

[25] Graham SE, Teague, C. Reading levels of rural and urban third graders. *Carsey Inst*; 2011.

[26] Ingersoll RM, Tran, H. Rural teacher turnover. *Educ Adm Q.* 2023;59(2):396-431.

[27] NIEA. *National Indian Education Association Teacher Initiative.* 2018

[28] NASEM. *K–12 STEM in Rural Areas.* NAS Press; 2025.

[29] Showalter D et al. *Why Rural Matters 2023.* NREA; 2023.

[30] NCSL. *Striving to achieve: Helping Native students succeed*. 2008.

[31] August D et al. Native American children and youth. *J Am Indian Educ.* 2006;45(3):24-37.

[32] NCES. *NAEP Reading Grade 12.* 2025. The National Indian Education Study. 2021

[33] Wang Z et al. Decoding threshold test. *J Educ Psychol.* 2019;111(3):387-401.

[34] Wang Z, Sutherland R. *The decoding threshold.* ETS Research Institute. 2024.

[35] [redacted]

[36] Seidenberg MS et al. Language and reading. *Top Lang Disord.* 2018;38(1):66-83.

[37] Filderman MJ et al. Word reading intervention. *J Learn Disabil.* 2022;55(5):393-407.

[38] Barnes MA et al. Tutor vs computer comprehension. *Sci Stud Read.* 2024;28(4):1-29.

[39] Kraft MA et al. Tutoring impacts at scale. *Annenberg WP.* 2024;24-1031.

[40] Lovett MW et al. HS reading remediation. *J Learn Disabil.* 2012;45(2):151-169.

[41] [redacted]

[42] [redacted]

[43] Westall J et al. Early literacy policies. *Educ Policy Innov Collab.* 2023.

[44] Virginia Dept of Education. *Virginia Literacy Act.* n.d.

[45] Minnesota Dept of Education. *READ Act Screening Guidance.* n.d.

[46] Indiana Capital Chronicle. Reading and education laws. 2024 Jun 27.

[47] AdLit. *Adolescent literacy interventions*. n.d. www.adlit.org/adlit-101-overview/

[48] Bhatt MP et al. *Personalized Learning Initiative.* UChicago Educ Lab & MDRC; 2025.

[49] Coburn CE. Rethinking scale. *Educational Researcher*. 2003;32(6):3-12.

[50] Hoadley C, Campos FC. Design-based research. *Ed Psych*. 2022;57(3):207-220.

[51] Sandoval WA, Bell P. Design-based research methods. *Ed Psych*. 2004;39(4):199-201.

[52] NSSA. *Tutoring survey instruments*. 2005. https://nssa.stanford.edu/research/tutoring-surveys

[53] RPPL. Assessing ELA curriculum shifts. https://rpplpartnership.org

[54] What Works Clearinghouse. *Procedures & Standards Handbook v5.0.* IES; 2024.

[55] Hill CJ et al. *Conducting implementation research in impact studies of education interventions*. 2023. http://ies.ed.gov/ncee.

[56] Makori A et al. *Scaling High-Impact Tutoring.* Annenberg Inst; 2024.

[57] Guryan J et al. *Not Too Late: Academic Outcomes.* NBER; 2021.

[58] Sabatini J et al. SARA Reading Components Tests, RISE Forms. *ETS Research Report Series*. 2015(2):1-20.

[59] [redacted]

[60] Phillips SF, Rowley JFS. The Tripod School Climate Index. *Social Work Research*. 2015;40(1):31-39.

[61] Raudenbush SW, Schwartz D. Randomized experiments in education. *Ann Rev Stat Appl*. 2020;7(1):177-208.

[62] Imai K et al. A general approach to causal mediation analysis. *Psych Methods*. 2010;15(4):309-334.

[63] Levin HM et al. *Economic Evaluation in Education.* Sage; 2018.

[64] Strauss AL, Corbin JM. *Basics of qualitative research*. 2nd ed. SAGE Publications; 1998.

[65] Lewis LK, Seibold DR. Innovation modification during intraorganizational adoption. *Acad Manage Review*. 1993;18(2):322.

[66] Ruiz-Primo M. *A multi-method and multi-source approach for studying fidelity of implementation.* 2006.

[67] Bishop DC et al. Measuring fidelity and adaptation. *Eval & Hlth Prof*. 2013;37(2):231-257.

[68] Cordray DS, Pion GM. Treatment strength and integrity. In: *Strengthening research methodology: Psychological measurement and evaluation*. 2006:103-124.

[69] Moylan LA, et al. Developing decoding instruction observation protocol. *Read Writ Q*. 2021;37:1-18.

[70] Balafoutas L et al. Replication crisis in social sciences. *J Behav Exp Econ.* 2025;114:102327.