



FY2025 Education Innovation and Research (EIR) Mid-Phase Grant - ALN 84.411B  
CitySchools Collaborative  
Project Narrative

## TABLE OF CONTENTS

<b>Absolute Priority</b>	<b>1</b>
<b>Competitive Preference Priority 2: Expanding Education Choice (High-Impact Tutoring)</b>	<b>4</b>
<b>A. Significance</b>	<b>5</b>
(A)(i) National Significance	5
(A)(ii) Promising Strategies and Evidence Base	5
(A)(iii) Potential Contributions to Knowledge and the Field	7
<b>B. Strategy to Scale</b>	<b>8</b>
Introduction	8
(B)(i) Demand	8
(B)(ii) Management Plan and Implementation Approach	9
(B)(iii) Project Partners	12
(B)(iv) Scaling and Efficiency	13
(B)(v) Dissemination and Replication	14
<b>C. Project Design</b>	<b>16</b>
Introduction	16
(C)(i) Logic Model and Framework	16
<b>(C)(ii) Goals, Objectives, and Outcomes</b>	<b>18</b>

<b>D. Evaluation</b>	<b>20</b>
(D)(i) WWC Alignment	21
(D)(ii) Guidance for Implementation and Replication	25
(D)(iii) Fidelity of Implementation	27
(D)(iv) Replication Information	29

CitySchools Collaborative (CSC) seeks Mid-phase support to scale a Science-of-Reading (SoR)-aligned literacy tutoring model, meeting Absolute Priority 1(a). High-impact tutoring content, materials, and coaching are explicitly aligned to the SoR, targeting phonemic awareness and phonics/decoding, fluency, vocabulary, and comprehension, meeting Absolute Priority 1(b). This project is supported by Moderate Evidence from an independent quasi-experimental evaluation of CSC’s program called Intermediary Supports for Tutoring (IST).<sup>1</sup> Through this grant, CSC will investigate the impact of these supports on literacy high-impact tutoring (HIT) through a rigorous study, described in the Evaluation section.

CSC’s intervention pairs consistent, small-group or one-to-one in-school tutoring with an IST system that addresses the common reasons tutoring underperforms at scale: weak dosage, variable alignment to core instruction, and insufficient implementation support. Our schools and providers implement tutoring to a defined, evidence-based dosage standard of at least 900 minutes (≈90 minutes per week for ten or more weeks). CSC supplies enabling infrastructure: structured coaching and technical assistance; Communities of Practice; Design Sprints to integrate tutoring into schedules and budgets; and quality-assurance (observations, dosage/progress monitoring, continuous-improvement routines). To safeguard fidelity and cost-effectiveness during scale-up, CSC uses outcomes-based contracting under which providers receive full payment only when at least 80 percent of students reach full dosage.

This approach is responsive to what the field has learned. Syntheses of tutoring research<sup>2</sup> highlight large average effects under the right conditions, while recent district and state experiences have produced uneven results when dosage, consistent student assignment, or curricular alignment are weak, a pattern underscored in the 2025 Stanford “Null Results” study.<sup>3</sup>

CSC's IST model prevents these failures by embedding fidelity support into day-to-day delivery.

In DC, an evaluation by EmpowerK12 using an inverse-probability-weighted quasi-experimental design found that students reaching the full 900-minute dosage achieved the equivalent of 59 additional instructional days in math. Students served by providers highly engaged with the IST model realized larger gains ( $\approx 92$  additional instructional days).<sup>4</sup> A Stanford University study of DC HIT also showed attendance benefits when tutoring is scheduled in-school, with students gaining 2–5 additional days of attendance per year.<sup>5</sup> These findings point to the IST model as a key driver of impact.

This project will embed the same support within SoR-aligned literacy tutoring and evaluate literacy outcomes directly (Absolute Priority 1b). As OSSE's contracted strategic partner, CSC ensures HIT providers also adopt SoR-aligned curricula alongside LEAs. CSC only approves providers that demonstrate use of high-quality curriculum aligned with the SoR and explicitly checks for SoR instructional practices in its observation and evaluation process.

Implementation prioritizes high-need students (Absolute Priority 1a) who are economically disadvantaged, multilingual learners, and students with disabilities, consistent with CSC's goals. By scaling this model across Washington, DC, Maryland, and Virginia, CSC will expand access for all students while demonstrating its relevance across diverse contexts. CSC will reach over 15,000 students with literacy HIT by year 5 of implementation. The research study will focus on grades 4-8, evaluating ELA outcomes for 82 schools serving 5,000 students. Final site counts, grade bands, and enrollment targets will reflect district and partner commitments and will be confirmed in the Work Plan and Management sections. By pairing SoR-aligned tutoring with intermediary supports validated in prior work and holding partners to implementation standards, this project directly satisfies EIR Absolute Priority requirements and

advances evidence on scaling literacy tutoring for students most affected by long-standing inequities.

### **Competitive Preference Priority 1: Returning Education to the States**

We meet CPP-1 by placing state education agencies in the lead and building policy, guidance, and monitoring infrastructure that enables SoR-aligned HIT to operate with fidelity at scale. The treatment studied in this RCT is not simply HIT, but IST services delivered by CSC in partnership with state education agencies, tutoring providers, and LEAs. In DC, the Office of the State Superintendent of Education (OSSE) acts as the SEA; CSC serves as OSSE's contracted strategic partner for IST services. Together, we have built state-level architecture – policy guidance, quality standards, and monitoring routines – that enables schools and providers to deliver tutoring with fidelity during the school day.

Early in the pandemic, OSSE prioritized HIT as an evidence-based recovery strategy and, working with CSC, rapidly mobilized federal, local, and philanthropic funds to stand up implementation at scale. With OSSE leadership and CSC's support, we produced for state use practical artifacts that inform and improve classroom delivery: a HIT Toolkit<sup>6</sup>, model guidance for LEAs<sup>7</sup>, a HIT math curriculum<sup>8</sup> aligned to state standards, a tutoring cost estimator<sup>9</sup>, and an evaluation/QA tool<sup>10</sup>. CSC operates the enabling infrastructure of IST, including Design Sprints, observation and feedback cycles, Communities of Practice, data dashboards, and subcontracts to tutoring providers, so dosage and quality sustain as programs scale. This scale and the growing evidence of academic results have led DC to sustain HIT with local investment.

Through this project, CSC will replicate and tailor the IST model throughout the mid-Atlantic region. Maryland has signed a letter of support for this project, given our alignment with Maryland's commitment to the SoR and the state's goal to rank among the top ten states in

NAEP reading by 2027. Virginia's literacy legislation mandates aligned curricula, teacher training, and consistent professional development in evidence-based reading instruction core to the SoR. CSC will provide partner SEAs with technical assistance, research, and fidelity monitoring, solving core implementation challenges as they develop state policies and funding guidance for HIT. Through these efforts, CSC will ensure new states own the capacity to launch, monitor, and sustain HIT.

### **Competitive Preference Priority 2: Expanding Education Choice (High-Impact Tutoring)**

We meet CPP-2 by expanding access to high-quality tutoring options that schools can select and schedule during the school day and, where available and standards-aligned, via vetted virtual models. This makes choice meaningful, consistent with research, and practical to implement. In DC, OSSE and CSC set (and exceeded early) a goal to reach 10,000 students by FY24; the network now includes more than 130 schools and 22 tutoring providers in DC and eight schools in Maryland and Virginia, with more than 30,000 students served as of SY2025-26. Schools choose among qualified providers that meet the seven research-backed standards co-developed by OSSE and CSC<sup>11</sup> grounded in 200 tutoring research studies. Providers sign outcomes-based contracts in which  $\geq 80\%$  of students must reach full dosage for full payment. This protects fidelity and ensures that choice is accompanied by quality.

Momentum to scale HIT continues to accelerate across districts and states, as a core strategy for academic recovery and improvement. CSC's demonstrated impact, including the 2024 Accelerate States Leading Recovery award and EmpowerK12 evaluation showing 59 additional instructional days of math growth, positions us to expand literacy tutoring at scale through the IST model. CSC has proven resources, state partnerships, and a coalition of providers and schools ready to expand education choice for families across the region.

**(A)(i) National Significance**

Across the United States, reading proficiency has declined and recovery from the pandemic remains incomplete. In 2022, only 31 percent of eighth-grade students reached the Proficient level on the National Assessment of Educational Progress (NAEP), the lowest rate in two decades.<sup>12</sup> Students from low-income families, English learners, and students with disabilities experienced the steepest losses, with average reading scores nearly two grade levels below those of higher-income peers. These inequities have persisted despite historic federal investments in learning recovery and a nationwide surge in HIT initiatives.

Research and field experience show that tutoring’s effectiveness depends not simply on the availability of tutors but on the fidelity, dosage, and instructional quality.<sup>13</sup> Many districts lack the infrastructure to monitor these factors or to align tutoring with classroom instruction and SoR. The result is widespread variation in quality and uneven impact.

CSC addresses this systemic gap. As the backbone organization for DC’s HIT effort, we established a coordinated infrastructure for implementing and evaluating HIT at scale. Our model builds state and district capacity to deliver consistent, SoR-aligned, and accessible HIT for all students. This project will extend this tested approach across jurisdictions to strengthen literacy outcomes and generate national evidence about how IST drives effectiveness in HIT.

**(A)(ii) Promising Strategies and Evidence Base**

CSC’s project introduces innovation in several ways that align with the Department’s priorities for evidence-building and replication in various choice-based education settings. First, it extends an existing evidence-based model into new states and contexts, testing whether IST maintains effectiveness when serving different populations and geographies. Second, it

represents a unique composition of interventions. While some of these have been the subject of individual studies, they have not been tested together rigorously at scale. Third, it incorporates emerging, highly-promising components for further testing, including Sophia, an AI coaching tool for tutors, and CSC's Tutor-to-Teacher program.

This intervention couples high-quality literacy tutoring with comprehensive IST that ensures fidelity, builds workforce capacity, and enables continuous improvement.

- **High-Impact Tutoring (HIT):** Through our flagship program, CityTutor DC, students receive  $\geq 900$  minutes (90 minutes per week for ten weeks) of small-group or one-to-one literacy tutoring embedded in school, both in-person and virtually. Sessions use SoR-aligned materials emphasizing phonological awareness, decoding, fluency, vocabulary, and comprehension.
- **Intermediary Supports for Tutoring (IST):** CSC provides Design Sprints to help schools integrate HIT into schedules and budgets; ongoing technical assistance and coaching for tutoring providers; Communities of Practice that facilitate peer learning; a quality-assurance rubric; and data dashboards that track dosage, fidelity, and outcomes.
- **Innovation for Scale:** Sophia, CSC's AI-supported tutor coach, delivers real-time pedagogical and content guidance to tutors using CSC's curriculum, increasing coaching reach without proportional cost increases. A parallel Tutor-to-Teacher pipeline builds a pipeline of future educators, strengthening instructional capacity and workforce development as tutoring expands. These innovations leverage the latest technology and align to the Department's priority to Advance Artificial Intelligence in Education.

### **Existing evidence of effectiveness**

CSC's model qualifies for Moderate Evidence under the EIR criteria. An independent

evaluation by EmpowerK12 (2025) used an inverse-probability-of-treatment weighting design to compare 5,810 tutored students across 92 schools with matched peers. Students who received the full 900-minute dosage gained the equivalent of 59 additional instructional days in math, with even larger gains for economically disadvantaged students. Providers highly engaged with CSC supports saw 90 additional instructional days of growth compared with less-engaged providers.<sup>14</sup>

The prior, quasi-experimental study demonstrates that CSC's IST model drives higher fidelity and stronger outcomes when tutoring reaches full dosage. In this project, we pair the same supports with Science-of-Reading-aligned literacy tutoring and test their impact on reading growth via an independent, WWC-aligned RCT across multiple states. By expanding to Maryland and Virginia partners, including College Park Academy (PG County) and multiple Virginia school divisions, CSC will test the model's validity across states and populations.

### **(A)(iii) Potential Contributions to Knowledge and the Field**

If successful, this project will generate strong causal evidence about which implementation supports are needed to ensure return on investment in literacy HIT, information that can inform investments at all levels. While research shows HIT accelerates learning, few studies have isolated the causal effects of the intermediary structures that make scaling possible.

Through a randomized-controlled trial led by Bellwether, CSC will test whether schools receiving IST support achieve greater literacy gains than comparison schools implementing tutoring with standard local supports. Outcomes will include student reading growth on state assessments and nationally normed interim assessments (e.g., i-Ready, MAP, mCLASS) and implementation fidelity measures such as dosage and quality-rubric ratings. The evaluation will meet What Works Clearinghouse (WWC) Standards Without Reservations and will yield generalizable findings about:

1. The extent to which IST yields stronger literacy achievement gains compared to business as usual, and whether impacts vary for different contexts or subgroups
2. Whether IST supports implementation fidelity and higher tutoring dosage for students
3. The effect of intermediary support for HIT on student engagement and confidence in reading, on tutors' confidence in implementing SoR-aligned curricula, and on students, teachers, and caregivers' satisfaction with tutors
4. The cost-effectiveness of providing IST support

By rigorously testing an adaptable, SoR-aligned HIT model across multiple states, we will contribute actionable evidence to the field, informing district, state, and federal decisions on how to sustain literacy acceleration for students most in need of support.

## **B. Strategy to Scale**

### **Introduction**

This project will expand access to high-quality, SoR-aligned literacy HIT by leveraging CSC's proven IST model, rigorous evaluation, and state partnerships. Over the five-year grant period, CSC will support at least 15,000 students with literacy HIT; the research study will focus specifically on grades 4-8, to evaluate ELA outcomes for ≈80 schools serving 5,000 students.

### **(B)(i) Demand**

On NAEP, only 33% of fourth graders and 31% of eighth graders nationwide are proficient in reading, dropping to 21% and 19% for students from low-income families.<sup>15</sup> District and state leaders are increasingly adopting HIT to accelerate literacy, but without an implementation infrastructure. Maryland has committed to SoR and MSDE is targeting a top-ten NAEP reading rank by 2027. Virginia law requires SoR-aligned curricula, teacher training, and ongoing professional learning in evidence-based reading instruction. Both states have identified

Parents and caregivers are also seeking this support: a 50CAN/Edge Research study (2024) found that 77% of Maryland families and 57% of DC families reported their child did not receive tutoring in SY23-24, with cost cited as a barrier by 36% and 29% respectively, revealing need for publicly-funded, in-school models. On-the-ground, demand also confirms the need for IST. In SY25, CSC began expansion efforts from DC into Maryland and Virginia, with demand from schools outpacing capacity. We have received formal requests to access CSC's IST model from twelve school leaders in Maryland, the MSDE, and eleven Virginia schools.

Our project addresses documented needs at an achievable scale given CSC's existing infrastructure and partner capacity. What all jurisdictions need most is the intermediary infrastructure to ensure that scaled tutoring maintains the fidelity standards necessary for impact. This unmet demand, combined with state policy momentum and family need, creates the conditions for rapid, sustainable scale of SoR-aligned literacy HIT.

### **(B)(ii) Management Plan and Implementation Approach**

CSC will serve as lead organization and fiscal agent, responsible for project oversight, partner coordination, state relationship management, and quality assurance. Reading Partners, Reading Assist, and EduTutor VA will provide tutoring with their established SoR-aligned curricula and school partners, while Bellwether will conduct the independent RCT evaluation.

In DC, CSC will build on our strategic partnership with OSSE to manage site recruitment and program operations (see Table 3). Existing contracts provide in-kind infrastructure support, reducing startup costs, while provider partners bring established networks enabling rapid site recruitment. Through CSC's contract with EmpowerK12, our data-sharing agreements cover the vast majority of DC Public Schools and public charter schools. CSC maintains an established

application to collect session-level tutoring attendance (see Table 2). These partnerships and existing infrastructure make the project feasible within the proposed budget.

In Maryland, CSC has begun recruiting schools with MSDE support and through Reading Partners' and Reading Assist's existing district relationships. In Virginia, we will continue cultivating relationships with EduTutorVA and Richmond Public Schools to expand further. State-specific engagements will proceed in coordination with the cohort schedule (Tables 1-2).

Following randomization by Bellwether, CSC will manage IST delivery in treatment schools, facilitating Design Sprints, convening Communities of Practice, and holding monthly implementation meetings with school teams and expert consultants (Table 2). Responsibilities are delineated as follows: CSC manages site recruitment, Design Sprints, Communities of Practice, and state relationships; providers deliver tutoring and participate in quality assurance protocols; Bellwether leads randomization, data collection, and evaluation; EmpowerK12 manages data aggregation. Tables 2 and 3 detail timelines, milestones, and responsibilities.

CSC leadership will manage regular check-ins with state leadership to share learnings, review data, and ensure resources reflect each state's needs and policy context. Our timeline allows Cohort 1 (Years 1-2) to inform refinements before scaling to Cohorts 2-3 (Years 2-5), enabling continuous improvement while maintaining research integrity. This phased approach is feasible given CSC's existing DC infrastructure and partners' established networks. Where timing considerations arise (e.g., testing windows, calendar shifts), CSC will sequence activities to keep tutoring embedded during the school day without disrupting instruction (Table 2).

**Table 1: Project Summary by Cohort**

	Year 1-2	Year 2-3	Year 3-4	Year 4-5
<b>Studied Schools</b> (get CSC supports; data collected)	Cohort 1A 15 schools	Cohort 2A 15 schools	Cohort 3A 15 schools	
<b>Control Schools</b> (no CSC supports; data collected)	Cohort 1B 15 schools	Cohort 2B 15 schools	Cohort 3B 15 schools	
<b>Delayed Treatment Schools</b> (get CSC supports; no data collected)		Cohort 1B 15 schools	Cohort 2B 15 schools	Cohort 3B 15 schools

**Table 2: Project Activities & Milestones**

Project Activities and Milestones	Owner	Y1	Y2	Y3	Y4	Y5
<b>Planning and Management</b>						
Hiring of additional CSC staff and selection of CSC contractors and tutoring providers	CSC	Q1				
Kick-off Meetings with key partners from SEA and relevant LEA; Kick off with providers in DC and MD	CSC	Q1-2				
Secure letter of support and plan with VOE for Cohort 2 recruitment	CSC/ EduTutor VA	Q2-3				
Recruit Schools and complete data sharing agreements and MOUs	CSC/RP/RA Bellwether	Q1-2 (C1)	Q1-2 (C2)	Q1-2 (C3)		
Kick-Off Meeting with school partners	CSC	Q2 (C1A)	Q2 (C2A+1B)	Q2 (C3A+2B)	Q2 (C3B)	
<b>Research &amp; Evaluation</b>						
Implement randomized school selection	Bellwether	Q2 (C1)	Q2 (C2)	Q2 (C3)		
Data collection, analysis, and reporting	Bellwether	Q3-4	Q3-4	Q3-4	Q3-4	Q3-4
<b>HIT Scaling &amp; Supports</b>						
Conduct Design Sprints	CSC	Q3 (C1)	Q3 (C2A+1B)	Q3 (C3A+2B)	Q3 (C3B)	
Conduct Observation & Feedback Cycle Conduct Communities of Practice, technical assistance, and coaching	CSC	Q3-4 (C1A)	Q1-2 (C1A) Q3-4 (C2A+1B)	Q1-2 (C2A+1B) Q3-4 (C3A+2B)	Q1-2 (C3A+2B) Q3-4 (C3B)	Q1-2 (C3B)
<b>SEA Support &amp; Capacity Building</b>						
Capacity-building with DC and MD state leaders (monthly coaching meetings, artifact delivery)	CSC	Q1-4	Q1-4	Q1-4	Q1-4	Q1-4
Survey state leaders; Annual summary report of data, learnings, and best practices	CSC/ Bellwether		Q3	Q3	Q3	Q3
Finalization of state-level toolkits						Q1-4

Below are key leaders of this project team. Additional staff will also support implementation.

**Table 3: Project Team Overview**

Position	Responsibilities
	Provide leadership; facilitate strong partnerships with state policymakers at OSSE and MSDE; manage Bellwether collaboration; ensure steady pipeline of philanthropic dollars for this project.
	Manage the project team; oversee the research and evaluation team against the timeline and deliverables committed.
	Manage implementation of programming, including supports and observation cycles with school partners.

Project Narrative

[REDACTED]	Manage HIT tutoring provider partnerships, including matchmaking of schools and providers and ongoing coaching and support.
[REDACTED]	Oversee all aspects of randomization, data collection, and evaluation.
[REDACTED]	Manage the expansion of Reading Assist tutoring with intermediary support for tutoring across DC and MD.
[REDACTED]	Manage the expansion of Reading Partners tutoring with intermediary support for tutoring across DC, MD, and VA.
[REDACTED]	Manage the expansion of Reading Partners tutoring with intermediary support for tutoring across VA.

**(B)(iii) Project Partners**

To carry out this project, CSC will partner with an SEA, LEAs, tutoring providers, and an independent evaluator. Partners have provided formal letters of commitment or formal contracts.

*State and LEA Partners.* We maintain a multi-year partnership with OSSE through a standing contract (see Other Attachments) and a joint award from Accelerate’s States Leading Recovery, under which we provide strategic HIT support to tutoring providers, local education agencies, and schools. Established in 2021, this partnership has been supported by \$7.7M in combined philanthropic and federal funding. OSSE will provide student-level assessment data through existing data-sharing agreements with DC Public Schools and public charter schools.

In Maryland, we have launched partnership work, including designing and piloting a HIT program at College Park Academy. MSDE has provided a letter of support (Appendix C) affirming state interest in SoR-aligned tutoring and committing to policy coordination and potential data-sharing. This positions us to replicate our IST model in Maryland. In Virginia, we have established a partnership with Richmond Public Schools via EduTutorVA and are poised to expand HIT in additional elementary and middle schools. We have also joined a proposal to serve six middle schools in Virginia Beach City Public Schools, with a decision pending.

LEA and delivery partners across all three states have committed to participating in randomization, scheduling tutoring during the school day, providing rosters and assessment data,

and facilitating observation visits. These LEA partnerships are essential for site access, data collection, and ensuring tutoring integrates with classroom instruction.

*Tutoring Provider Partners.* Across the region, we have established multi-year partnerships with providers who bring proven SoR expertise and a commitment to participating in quality assurance protocols, including randomization, observation cycles, dosage tracking, and coaching. Key partners include: **Reading Partners**, who will serve as a delivery partner for HIT with expertise in leveled curriculum design integrating all five SoR components (Letter, Appendix C.); **Reading Assist**, who will provide Orton-Gillingham-based literacy instruction with focus on serving students below proficient, English learners, and students eligible for free or reduced-price lunch. (Letter, Appendix C.); and **EduTutorVA**, who will provide SoR-aligned HIT for K–8 students in Virginia, delivered by trained college tutors (Letter, Appendix C).

*Evaluation Partner.* Bellwether Education Partners will serve as independent evaluator, conducting an RCT designed to meet WWC Standards Without Reservations (Letter, Appendix C). [REDACTED] brings extensive expertise in large-scale experimental and quasi-experimental evaluations, including Charter Schools Program, Regional Comprehensive Center, and Teacher Quality Partnership grants. Bellwether will lead school randomization; coordinate data collection from state ELA testing systems and interim assessments; conduct site visits and interviews; analyze CSC quality assurance observations, EmpowerK12 surveys, and program records; assess student literacy outcomes and stakeholder confidence and satisfaction; and produce annual reports and a final evaluation. Bellwether has committed to coordinating with the cohort schedule to preserve instructional time and minimize school burden.

#### **(B)(iv) Scaling and Efficiency**

States have invested heavily to expand HIT, yet research has documented uneven results

when programs lack dosage, consistent assignment, or curricular alignment.<sup>16</sup> This project is designed to increase states' return on investment by funding the IST that protects fidelity. With these supports in place, the per-student cost of achieving literacy gains decreases, and a greater share of tutored students meet growth targets and improve on state assessments.

CSC's own cost trajectory illustrates this efficiency. The per-student cost of CSC's IST support to OSSE declined from ≈\$500 per student in FY22 to < \$100 per student in FY26 – an ~80% reduction – while the number of students supported increased over the same period. Student outcomes remained strong throughout this scale-up, demonstrating that cost reductions did not compromise program effectiveness. Student reach, dosage, and test schools have all improved since the 21-22 school year. This improvement reflects the power of our standardized tools, which allow us to maintain the quality of tutoring as the number of schools and providers grows. New tools like Sophia will further increase quality while driving scale. Likewise, expanding to Maryland and Virginia will generate additional efficiencies through shared tools and training materials that can be adapted rather than built from scratch.

As schools see measurable results and internalize quality routines, they require less intensive support, creating a virtuous cycle: CSC's per-student costs decline, schools build internal capacity, and states realize stronger returns on their tutoring investments.

### **(B)(v) Dissemination and Replication**

Dissemination is core to CSC's mission and practice. We have already co-developed public resources with OSSE, referenced previously, with each designed for direct use by schools and providers. This grant will scale that commitment, producing practical, adaptable materials grounded in school-day implementation, dosage  $\geq 900$  minutes, SoR alignment, and observation-driven QA. All materials will be openly licensed (Creative Commons) to enable free

Our strategy is structured to reach distinct audiences with fit-for-purpose products. Policymakers and funders will be served through ongoing meetings with state leaders (DC, MD, VA), state-level toolkits adapted to local policy and data contexts, model contract language incorporating outcomes-based thresholds and SoR requirements, and Bellwether's evaluation reports and state-level briefings. Educators and practitioners will be reached through IST support with artifacts meant for immediate implementation. To reach beyond direct project participants, CSC will present at national conferences (e.g., Accelerate, NewSchools Venture Fund, and National Student Support Accelerator), host public webinars on implementing SoR-aligned HIT with fidelity, and maintain a publicly accessible online repository of all tools and resources.

Dissemination will occur throughout the project: annual reports and updated toolkits will be released each year, with the final evaluation and implementation guide published in Year 5. Upon completion, CSC will submit the final evaluation report to the Education Resources Information Center (ERIC) to ensure broad public access and support national replication.

This approach supports replication and IST durability beyond the grant. Resources will be designed for adaptation: toolkits will include customization guidance; observation rubrics will specify minimum fidelity thresholds while allowing state-specific adjustments; cost estimators will enable districts to model implementation under different scenarios. By building evidence of external validity across DC, Maryland, and Virginia and ensuring state education agencies possess codified resources to launch, monitor, and sustain SoR-aligned tutoring, the model can be carried forward by states and districts over time.

## **Introduction**

CSC is tackling one of the most persistent challenges in education: the wide and unacceptable variability in literacy achievement across states and student groups, often driven by inconsistent quality in tutoring programs.<sup>17</sup> This project will test whether SoR-aligned HIT delivered with IST, including Design Sprints, technical assistance, Communities of Practice, and quality assurance systems, produces stronger literacy outcomes than tutoring delivered without structured implementation infrastructure. By demonstrating causal impact through a multi-state randomized controlled trial, CSC will generate actionable evidence for state policymakers to maximize return on tutoring investments. Findings will inform sustainable policy and funding frameworks so states can maintain high-quality HIT interventions beyond federal grant support.

### **(C)(i) Logic Model and Framework**

The logic model below illustrates how CSC's Intermediary Support for Tutoring (IST) model drives student literacy gains. The model represents the treatment condition for the randomized controlled trial: schools receiving CSC's full IST package will be compared to schools implementing literacy tutoring without it, testing whether IST produces significantly greater literacy gains and implementation fidelity.

The Project inputs include federal EIR funding, SoR-aligned curricular resources, state partnerships, and data systems enabling session-level attendance tracking and academic assessment monitoring. These resources address the primary implementation barriers that cause tutoring programs to underperform. Design Sprints solve scheduling and resource constraints, integrating tutoring into master schedules and budgets. Individualized technical assistance and coaching close provider capacity gaps, ensuring tutors deliver content addressing all five SoR

components. Communities of Practice overcome isolation among educators through peer learning and troubleshooting. Standardized observation rubrics and data dashboards make implementation quality visible and enable rapid intervention.

By systematically removing these barriers, CSC's supports increase the percentage of students receiving high-fidelity tutoring at target dosage. Improved implementation produces measurable student outcomes: more students meet or exceed expected literacy growth on interim assessments, demonstrate gains across the five SoR components, and report increased confidence and engagement in reading. As full-dosage, high-fidelity HIT reaches more students, particularly students furthest from opportunity, cumulative exposure drives increases in grade-level reading proficiency on state assessments and narrowing of achievement gaps. Simultaneously, SEA staff develop capacity to sustain high-quality HIT through coaching, codified resources, and validated models, creating durable policy and funding structures beyond federal support.

**Table 4. Logic Model** (The full, detailed version of the logic model is in Appendix G)

Inputs	Key Actions/Outputs	Short-Term Outcomes	Long-Term Outcomes
<ul style="list-style-type: none"> <li>● Federal, local, and philanthropic funding</li> <li>● Program team</li> <li>● Research and evaluation team</li> <li>● SoR-aligned curricular resources, state policies, and state investments</li> <li>● Existing research on HIT and SoR</li> <li>● State contract with OSSE and state support from Maryland</li> <li>● Active engagement and human capital from schools and providers, trained in SoR</li> <li>● Data Systems at SEA, LEA, and school levels; HIT attendance tracking</li> </ul>	<ul style="list-style-type: none"> <li>● Support and funding of literacy HIT</li> <li>● Design sprints for schools to develop launch plans</li> <li>● Individualized technical assistance &amp; coaching</li> <li>● Standardized observation, evaluation, and feedback</li> <li>● Communities of Practice</li> <li>● Dissemination of SoR-aligned curricula, tools, and training</li> <li>● In-depth data analysis on dosage, surveys, and assessments</li> <li>● SEA-level coaching and resource development</li> <li>● State Data partnerships</li> <li>● Resource dissemination</li> </ul>	<ul style="list-style-type: none"> <li>● Increase the number of students receiving high-fidelity SoR-aligned literacy HIT</li> <li>● Higher rates of students meeting or exceeding expected growth in literacy on nationally normed assessments (iReady, MAP, mCLASS)</li> <li>● Increased student confidence and engagement in reading, as measured by student surveys</li> <li>● High confidence of tutors in implementing SoR-aligned lessons</li> <li>● Increased SEA capacity</li> </ul>	<ul style="list-style-type: none"> <li>● Increased percentage of students reading at or above grade level on state assessments</li> <li>● Reduced literacy achievement gaps for all students and subgroups served</li> <li>● A validated, cost-effective, SoR-aligned literacy HIT model ready for statewide adoption</li> <li>● Policy momentum and funding structures in DC and Expansion States to sustain literacy HIT beyond the grant period</li> <li>● Meaningful contribution to the national evidence base on literacy HIT</li> </ul>

**(C)(ii) Goals, Objectives, and Outcomes**

CSC's goals and objectives are ambitious yet achievable within the five-year project period, directly aligned with EIR's purposes of improving literacy outcomes for high-need students, building the evidence base for literacy interventions, and enabling state-led scale.

Goal 1 increases reading proficiency through widespread, high-fidelity implementation of SoR-aligned literacy HIT. Targeting 15,000 students across at least 90 schools in three states is ambitious but achievable given CSC's existing DC infrastructure and established provider partnerships in Maryland and Virginia. Implementation targets – at least 25% more students at treatment schools meeting the 900+ minute dosage target, and higher observation ratings – are grounded in prior experience. The +10 percentage point increase in students meeting expected literacy growth is ambitious but achievable given CSC's track record of improving implementation quality. Additional targets include 80% of students improving across the five SoR components and 85% reporting increased confidence and engagement in reading.

Goal 2 enables SEAs to sustain and expand literacy HIT beyond the grant period, aligning with EIR's emphasis on state-led innovation. Providing MSDE with codified materials and validated models creates conditions for statewide adoption without ongoing federal support. The +10 percentage point increase in students receiving full-dosage tutoring in expansion states of Maryland and Virginia is ambitious but feasible with CSC's IST reducing implementation barriers. SEA leaders in expansion states are expected to report at least a 10% increase in confidence in their abilities to launch, monitor, and report on SoR-aligned literacy HIT based on pre-IST and post-IST surveys. Objective 2.2, producing one research study demonstrating causal impact, fulfills EIR's evidence-building purpose.

All objectives include quantitative targets measurable through existing data systems:

student assessments (iReady, MAP, mCLASS), dosage tracking, observation scores, and stakeholder surveys. CSC will disaggregate outcomes by student subgroups (economically disadvantaged, English learners, students with disabilities) to ensure effectiveness for high-need populations, with annual reporting of subgroup-specific growth rates and dosage attainment.

**Table 5. Objectives and Performance Measures.**

<p><b>Goal:</b> Increase reading proficiency rates through wide-spread, high-fidelity implementation of HIT.  <i>Outcome:</i> CSC-supported students achieve higher literacy proficiency and sustained reading growth on interim and state assessments</p>		
Objective	Measure	Target by Year 5
1. Scale access to literacy HIT across Washington DC, Maryland, and Virginia	<ul style="list-style-type: none"> <li># of students receiving SoR-aligned literacy HIT</li> <li># of schools implementing SoR-aligned literacy HIT</li> </ul>	<ul style="list-style-type: none"> <li>Reach at least 15,000 students with SoR-aligned literacy HIT</li> <li>Partner with 100+ schools to implement SoR-aligned literacy HIT</li> </ul>
2. Improve quality of literacy HIT implementation through IST Supports	<ul style="list-style-type: none"> <li>% of students meeting target dosage (900+ minutes of HIT)</li> <li>Evaluation scores from site observations of schools receiving IST from CSC</li> <li>% of stakeholders reporting satisfaction with their tutor or their student’s tutor</li> </ul>	<ul style="list-style-type: none"> <li>80% of students at Treatment Schools meet the dosage target (900+ minutes).</li> <li>90% of Treatment Schools are rated “Proficient” or “Exemplary”</li> <li>80% of stakeholders at Treatment Schools report satisfaction with HIT</li> </ul>
3. Improve literacy results of students at schools receiving IST	<ul style="list-style-type: none"> <li>% change in students at Treatment Schools who are meeting or exceeding expected growth in literacy on nationally normed assessments (iReady, MAP, mCLASS)</li> <li>% of students at Treatment Schools will show improvements in SoR components on MAP i-Ready, and mCLASS.</li> <li>% of students at Treatment Schools will report increased confidence and engagement in reading</li> </ul>	<ul style="list-style-type: none"> <li>+10% of students at Treatment Schools who are meeting or exceeding expected growth in literacy on nationally normed assessments (iReady, MAP, mCLASS)</li> <li>80% of students at Treatment Schools will show improvements in SoR components on MAP, i-Ready, and mCLASS</li> <li>85% of students at Treatment Schools will report increased confidence and engagement in reading, as measured by student surveys</li> </ul>
<p><b>Goal:</b> Enable a stronger return on investment for states prioritizing literacy HIT. <i>Outcome:</i> CSC-supported states develop policies and increase funding in literacy HIT as a result of seeing a demonstrable link between literacy HIT and student higher literacy proficiency</p>		
4. Provide coaching and tailored resources to SEA staff on HIT implementation best practices	<ul style="list-style-type: none"> <li># of SEAs receiving State-level HIT implementation resources</li> <li>% of SEAs receiving Exemplar HIT models created</li> </ul>	<ul style="list-style-type: none"> <li>At least 2 SEAs will receive codified materials and toolkits for implementing HIT at scale, within local contexts</li> <li>All CSC-supported states receive at least 1</li> </ul>

	<ul style="list-style-type: none"> <li>• % increase in students receiving 900+ minutes of literacy tutoring</li> <li>• % increase in confidence of state leaders' abilities to launch, monitor, and report on SoR-aligned literacy HIT</li> </ul>	<p>validated, cost-effective, SoR-aligned HIT model ready for statewide adoption</p> <ul style="list-style-type: none"> <li>• In expansion SEAs, there will be a +10 percentage point increase in the percent of students receiving 900+ minutes of literacy tutoring</li> <li>• Expansion SEA leaders will report at least a 10% increase in confidence in their abilities to launch, monitor, and report on SoR-aligned literacy HIT</li> </ul>
5. Meaningfully contribute to the national evidence base on literacy tutoring	<ul style="list-style-type: none"> <li>• # of research studies on the impact of IST at the school level</li> </ul>	<ul style="list-style-type: none"> <li>• 1 research study demonstrating causal impact of IST at the school level</li> </ul>

#### D. Evaluation

Bellwether will lead an independent evaluation of the *CitySchools Collaborative Intermediary Supports for Tutoring (IST) program* including implementation and rigorous experimental impact analyses. Bellwether will conduct a school-randomized controlled trial (RCT) to estimate the effect of CSC on student literacy outcomes on state standardized achievement exams and assess CSC's scaling strategy. The implementation study will investigate how IST is implemented relative to implementation fidelity standards and will provide timely formative information to support program improvement and future replication. The evaluation will measure and report on all components, mediators, and outcomes described in the program logic model (Table 5). The impact evaluation will address the research questions (RQ) in Table 6.

**Table 6. Evaluation Timeline**

	Y1 (2026)				Y2 (2027)				Y3 (2028)				Y4 (2029)				Y5 (2030)			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
<b>Implementation Study</b>																				
Interviews						X				X				X				X		
Surveys						X				X				X				X		
Project Data						X				X				X				X		
Memo							X				X				X				X	
<b>Impact Study</b>																				
Randomization		X				X				X										
Analysis								X				X				X				
Surveys						X				X				X						
Impact Report								X				X				X				

**(D)(i) WWC Alignment**

Table 7 summarizes impact study research questions, program outcomes, and associated data sources. Research questions include the mediating and moderating effects to be studied.

**Table 7. Research questions, program outcomes, and data sources**

<b>Impact of IST on Students</b>	<b>Data Source</b>
RQ 1. Confirmatory: What is the effect of IST on ELA achievement for students in grades 4-8?	Student-level scores from state ELA assessment (DC-CAPE, MD-MCAP; VA-SOL)
RQ 2. Exploratory: Which, if any, student subgroups benefit the most from the intervention?	
RQ 3. Exploratory: Do students in IST schools report increased confidence and engagement as readers compared to students in comparison schools?	EmpowerK12 Student Survey
RQ 4. Exploratory: What is the effect of IST on fluency, decoding, and comprehension skills for treatment students in grades 1-3?	Student-level scores from early literacy assessment (NWEA MAP, I-ready, M-Class)
RQ 5. Mediation: What is the effect of IST on tutoring dosage and implementation fidelity? To what extent are dosage and implementation fidelity related to IST student outcomes?	Implementation fidelity scores (Spring 2026); Dosage data per school; Student-level scores from state ELA assessment
<b>Impact of IST on Adults</b>	<b>Data Source</b>
RQ 6. Exploratory: Do tutors who receive IST report improved confidence in implementing SoR-aligned lessons compared to those at non-IST schools?	EmpowerK12 Tutor Survey
RQ 7. Exploratory: Do students, teachers, and parents in IST schools report higher satisfaction with tutors compared to those in non-IST schools?	EmpowerK12 Student, Teacher, and Parent Surveys
<b>Implementation of IST</b>	<b>Data Source</b>
RQ 8. Is IST implemented as intended? What factors facilitate or hinder implementation and replication?	IST administrative data (Summer, 2025, 2026, 2027) EmpowerK12 Student, Teacher, Tutor, and Parent Surveys

RQ 9. How do students, tutors, teachers, school leaders, and state agency leaders perceive their experience and the impacts of IST?	Interviews with teachers, tutors, and state agency leaders
---	--

**Confirmatory Impact Analysis (RQ1).** Bellwether will estimate one confirmatory contrast using a multisite (blocked) cluster randomized trial (MSCRT): the impact of IST on student literacy achievement after 1 year of program exposure. The study will compare student outcomes in grades 4-8 between schools implementing CSC’s IST program and a Business-as-Usual (BaU) control group. The evaluation will include at least 5,000 students across 82 schools in three states (approx. 50-65 students per school). Each new cohort of schools will be randomly assigned within blocks, with each having a 50% probability of assignment to the treatment group. Randomization will be blocked by state, school type (charter or traditional public school), and school level to provide balance across grade level, school, and contextual factors. To ensure that the study does not include joiners, Bellwether will collect student rosters prior to school randomization; these rosters will be used to determine the student and teacher samples for the duration of the study (i.e., in year 2, student joiners will not be included in analysis). During the RCT control schools will provide tutoring to students as they normally would (which includes not providing any HIT); they will be given CSC’s IST program when the study concludes (“business-as-usual” with delayed treatment control condition).

Bellwether will use recruitment and data collection strategies to minimize both overall and differential attrition (e.g., clear communication prior to randomization, financial incentives for all data collection activities, local (district) research coordinators to support data collection.<sup>18</sup> Outcome measures are valid, reliable, appropriately aligned to the intervention and will be collected in the same way across treatment and control schools. No outcome data will be imputed and no significant confounds exist.

We will use Hierarchical Linear Model (HLM) models to adjust standard errors

associated with the clustering of observations within schools, thus minimizing Type I error associated with nested models.<sup>19</sup> For all outcome analyses, Bellwether will test for baseline equivalence in the analytic sample using ELA achievement scores on state standardized assessments prior to randomization. The predicted literacy outcome for student  $i$ , in school  $j$ , in randomization block  $k$  as a function of attending a school assigned to treatment is modeled using as the following Hierarchical Linear Model:

$$Y_{ijk} = \beta_0 + \beta_1(CSC\ Support_{jk}) + \beta_2(Baseline_{ijk}) + \delta_k + \mu_{jk} + e_{ijk}$$

Random effects  $e_{ijk}$  and  $u_{jk}$  account for residual variation at the student and school levels, respectively. The vector  $\delta_k$  represents fixed effects for school-type by school-level by state randomization blocks. Student baseline achievement scores will be included as covariates to improve the precision of the impact estimate. The coefficient  $\beta_1$  provides the estimated Intent-to-Treat (ITT) effect of school assignment to IST on student literacy achievement. Missing data will not be imputed; analyses will be conducted using all available cases.

**Power.** Bellwether will identify ~90 schools for the study, aiming to retain 90% (82) after attrition. We calculated the MDES using a two-level model, assuming the top level N is 82 schools, with 50 students per school. The MDES for ELA achievement outcomes is .15, a medium effect for educational interventions.<sup>20 21</sup> Bellwether will use the Benjamini-Hochberg correction to account for two confirmatory comparisons within the same domain.

**Exploratory Analysis.** Bellwether will conduct exploratory analysis to generate information about the contexts in which IST is most effective. To address RQ2, we will run the above-described HLM models separately for subgroups of interest (e.g., by English learner status, special education status, and gender), though with a non-centered indicator for the subgroup of interest and an interaction term between the subgroup indicator and the treatment

indicator. To address RQ3, we will run the above-described HLM models with student confidence and engagement as the outcomes, rather than ELA achievement. To address RQ4, we will use OLS regression to model growth from beginning to end of the year on fluency, decoding, and comprehension controlling for student and school characteristics (including student baseline achievement and demographics, as well as school state, level, and type).

Mediation models are described in greater detail below.

**Instrumentation.** The rigorous evaluation design will be bolstered by objective and reliable outcome measures that capture the key constructs in the logic model. *Student ELA Achievement (RQ1)*. Bellwether will collect student-level state-administered standardized ELA (grades 3-8) test score data in spring 2026, 2027, 2028 (C1, C2, C3 baseline) and spring 2027, 2028, and 2029 (C1, C2, C3 outcome), for all treatment and comparison students. Standardized ELA data typically include the student's scale scores, which are measured as continuous variables. We will standardize these scores using the state means and standard deviations (where available) or the sample mean and standard deviation (if state values are not available). This approach will allow us to combine achievement scores across states whose assessments are on a different scale from one another. Bellwether staff will also work with each state to obtain student-level demographic data (e.g., race/ethnicity, socio-economic status, gender) for treatment and comparison students. What Works Clearinghouse standards recognize state assessments as valid and reliable.<sup>22</sup> *Student phonological awareness, phonics, & fluency.* Bellwether will create data sharing agreements with participating LEAs to obtain student data from early grades standardized literacy assessments, including: I-Ready, M-Class, and NWEA MAP to supplement what CSC has already secured for more than half of DC schools and LEAs. Standardized assessments are recognized by WWC as valid and reliable measures of phonological awareness,

phonics/decoding, and fluency.<sup>23</sup> *Student Confidence & Engagement in Reading; Student,*

*Teacher, Parent Satisfaction with Tutoring; Tutor Confidence Implementing SOR-aligned*

*lessons.* EmpowerK12 has developed surveys for students, teachers, parents, and tutors to collect

data on their experiences with tutoring. Current survey topics include tutors' sense of efficacy

and impact, students' self-confidence and engagement with tutoring and self-confidence, and

teachers' and parents' satisfaction and sense of impact. Bellwether will add or revise items on

these surveys as needed to measure the constructs above. EmpowerK12 will administer the

surveys through its Qualtrics platform and share raw individual level data with Bellwether for

independent analysis. The surveys will be distributed to treatment and comparison schools.

*Implementation Fidelity, including Dosage.* We will use dosage data and the *High-Impact*

*Tutoring (HIT) Evaluation Tool* developed by the CSC and OSSE to assess the fidelity of

program implementation. To collect dosage data, tutors at treatment schools will use CSC's web

application to submit daily attendance data, while control school tutors will track this data in a

shared Excel document for weekly submission. To assess program implementation, CSC's tool

combines structured observations, program interviews, and scoring rubrics aligned to national

HIT standards. The tool provides ratings for tutor effectiveness, curriculum quality, data use, and

collaboration with schools and yields an implementation score for each school. CSC conducts an

annual training for inter-rater reliability and also conducts multi-rater evaluations for any new

raters. CSC staff will administer the tool; Bellwether will analyze the data to assess fidelity.

#### **(D)(ii) Guidance for Implementation and Replication**

Bellwether will document CSC's strategies to provide IST across multiple states, LEAs,

and schools. We will study (a) the resources CSC invests in supporting schools to design and

continuously improve their tutoring programs, (b) schools' use of IST support components, and

(c) state and school leaders' developing knowledge and skill related to implementing and sustaining high-quality HIT implementation. For each of these topics, we will investigate contextual factors that support or inhibit successful replication and document variation as the program is adapted to local needs. Our final analysis will triangulate this school-level contextual data with implementation fidelity data and impact estimates to identify project components that are critical to successful replication and sustained implementation in various contexts.

To conduct this investigation, Bellwether will gather data related to each program component (see Table 8) through artifact analysis, stakeholder surveys, program staff interviews, and site visits. For **Implementation Planning**, we will analyze the HIT launch plans created during Design Sprints by each cohort of expansion schools, noting similarities and differences across sites' goals (e.g., target grade levels and student selection criteria) and implementation details (e.g., location, timing, technology, curriculum alignment, data tracking). For **Instructional Resources for Tutors**, we will analyze tutor survey data to understand tutors' experiences with different curricula and trainings. For **Technical Assistance and Continuous Improvement**, Bellwether will conduct site visits to a purposive sample of treatment and comparison schools (representing a variety of geographies, school types (charter vs. district), and tutoring models), to interview school tutoring program directors and conduct focus groups with tutors and teachers about their experiences with IST and confidence about sustaining high quality HIT. We will also interview CSC and SEA staff annually about their perceptions of contextual factors that support or hinder successful replication and scaling. All interviews will explore how the IST components support replication, the sites' plans for sustainability, and what additional support might be needed to sustain HIT across sites. In addition to the qualitative data, we will work with Empower K12 to add relevant items to the teacher, tutor, parent, and student surveys

for further insight into local needs, variations, and contextual factors that affect replication.

Bellwether will provide annual briefings to CSC to provide formative feedback about implementation and effective scaling of supports. At the conclusion of the evaluation, data will be synthesized and included in dissemination efforts to contribute to the knowledge base about scaling effective supports for HIT.

### **(D)(iii) Fidelity of Implementation**

Bellwether will examine program implementation, as defined by the key project components, and its relationship with the student outcomes described above. These analyses will include implementation fidelity, treatment control contrast, and mediation.

**Implementation Fidelity.** Bellwether will collect and analyze implementation data to assess the fidelity of CSC's IST program. Key components and their annual thresholds are summarized in the implementation framework (Table 8). We will measure reach and scale of school engagement (Component 1) through records of school onboarding, identification of students for tutoring, and dosage. We will measure launch planning (Component 2) and technical assistance and continuous improvement (Component 4) through records of school participation in Design Sprints, monthly coaching meetings, and quarterly community of practice meetings. The HIT Evaluation Tool and tutor survey will yield information about resources for tutors (Component 3). To be considered implemented with fidelity, schools must meet benchmarks for all four components each year. Bellwether will supplement this fidelity analysis with mixed-methods data to understand variation in implementation quality and context. Tutor surveys will capture data on program uptake among participating sites (treatment only), differences between IST and business-as-usual practices (treatment-control contrast), and perceived supports and barriers to implementation (treatment only). Analyses will also examine

variation across LEAs to assess the consistency of replication. Finally, interviews with state partners, school staff, IST staff, and tutors will provide insight into local conditions and contextual factors that may influence implementation quality or potential contamination or crossover (if these occur).

**Table 8. Implementation Framework**

Key components and indicators	Threshold	Data source
<b>Component 1: Reach and Scale</b>		
100% of schools reach site-level threshold on all three indicators		
Indicator 1.1 Schools identify students for high intensity literacy tutoring in alignment with SoR each year	100% of IST schools in each cohort	Program Data
Indicator 1.2. Schools in each new cohort are onboarded to understand the core components of IST and requirements for implementing HIT in alignment with SoR	100% of IST schools in each cohort	Program Data
Indicator 1.3. Students identified for tutoring receive at least 90 minutes a week for at least 10 weeks	80% of students receive full dosage HIT	Dosage Data
<b>Component 2: Implementation Planning</b>		
100% of schools reach site-level threshold		
Indicator 2.1. School participates in four-week intensive Design Sprint	Once per school	Program Data
<b>Component 3: Instructional Resources for Tutors</b>		
100% of sites reach threshold on both indicators		
Indicator 3.1. Tutors use CSC-approved SoR-aligned curricula like UFLI and Reading Partners Connects	100% of tutors	HIT Evaluation Tool data; tutor survey
Indicator 3.2. Tutors participate in CSC or provider training on SoR-aligned curricula	100% of tutors	Program data; tutor survey
<b>Component 4: Technical Assistance and Continuous Improvement Support for Schools</b>		
100% of sites reach site-level threshold on all indicators		
Indicator 4.1. School leadership attends coaching check-ins with CSC	Monthly through each school year	Program data
Indicator 4.2. School leadership attends Community of Practice with CSC	Quarterly	Program data
Indicator 4.3. CSC staff conduct implementation fidelity evaluation with CSC's HIT Evaluation Tool and provide feedback to site manager	Once per semester per site	Program data; Evaluation Tool data
Indicator 4.4. Empower K12 administers surveys to tutors, teachers, students, and caregivers	Once per year per school	EmpowerK12 survey data

**Mediation.** We will conduct mediation analyses to understand how implementation quality is related to IST's effectiveness. We hypothesize that the following elements of IST

implementation are most likely to mediate the effects of IST on student achievement: dosage (minutes of tutoring) and overall implementation quality. We will use structural equation modeling (SEM) to test mediation effects of the IST elements because of its ability to adequately address the presence of measurement error within a statistical model over regression models.<sup>24</sup> These models will estimate the proportion of any estimated IST impacts that are mediated through the facets of instruction to which IST is most closely aligned, allowing for a better understanding of which program components are most critical to IST success.

#### **(D)(iv) Replication Information**

The evaluation will result in guidance for effective scaling and replication of IST in other settings through the implementation and impact studies; codification of strategies; reporting; cost-effectiveness analysis; and publication and dissemination. **Studies.** The implementation study will triangulate data on contextual factors, implementation fidelity, and stakeholder perceptions to generate insights about program components most critical to effective replication and adaptation. The impact study will assess whether the program impact differs across subgroups and randomization blocks, which will help indicate areas for IST to improve to better support diverse students in various contexts and different states. **Codification.** Bellwether will collaborate with CSC to create toolkits and templates that codify insights about effective implementation and scaling for new states looking to replicate. **Reporting.** Bellwether will report each year's implementation findings in briefings with CSC and state leaders, providing feedback that will support continuous improvement of replication efforts in each state. Insights from these annual reflection sessions will inform final implementation resources and evaluation reports. **Cost Effectiveness.** Bellwether will conduct a cost-effectiveness analysis using the Resource Cost Model (RCM) framework, which identifies and values the resources required to implement

an intervention.<sup>25</sup> The analysis will employ the CostOut Tool to estimate program costs and link them to the corresponding impact estimates, enabling calculation of cost-effectiveness ratios and comparisons across outcomes. ***Public Dissemination.*** Finally, we will publicly disseminate findings from the implementation and impact studies to inform the field about effective replication strategies identified through the evaluation. ***Pre-registration.*** Bellwether will pre-register the IST impact study in the Registry of Efficacy and Effectiveness Studies (REES), including research and analysis activities and sample information, and update the registry if changes occur.