

The End of ESSER: Lessons about Sustaining Tutoring at Scale from Guilford County Schools

Ayesha K. Hashim

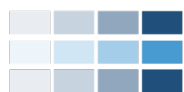
Sofia Postell

Jimmy Leak

Kara Hamilton

April 2026

WORKING PAPER No. 338-0426



CALDER

National Center for Analysis of
Longitudinal Data in Education Research



**The End of ESSER:
Lessons about
Sustaining Tutoring
at Scale from
Guilford County
Schools**

Ayesha K. Hashim
NWEA

Sofia Postell
NWEA

Jimmy Leak
Guilford County Schools

Kara Hamilton
Guilford County Schools

Contents

Contents.....	i
Acknowledgments	ii
Abstract.....	iii
1. Introduction	1
2. Literature Review on HDT Implementation and Scale	3
3. Home-Grown HDT in GCS.....	8
4. Conceptual Framework for Understanding GCS's Tutoring Implementation.....	11
5. Data and Methods.....	14
6. Findings	19
7. Discussion.....	39
References	46

Acknowledgments

We are grateful to district leaders in Guilford County Schools for their support of this research–practice partnership, as well as to the educators and school leaders who generously gave their time to participate in interviews, focus groups, and other data collection activities. We also thank Dan Goldhaber, Michael DeArmond, Emily Morton, and Amanda Lu for their thoughtful feedback on earlier drafts of this paper and guidance on revisions. We appreciate the contributions of former members of the NWEA research team, including Jazmin Isaacs and Miles Davison, who supported data collection and analysis. This research was funded by the AIR Equity Grants Initiative, the Belk Foundation, and The Duke Endowment.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any errors are attributable to the authors.

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

The End of ESSER: Lessons about Sustaining Tutoring at Scale from Guilford County Schools

Ayesha K. Hashim, Sofia Postell, Jimmy Leak, & Kara Hamilton

CALDER Working Paper No. 338-0426

April 2026

Abstract

We examine Guilford County Schools' efforts to scale and sustain a home-grown high-dosage tutoring (HDT) program using a multi-level case study design drawing on district documents, observations, and interviews with district and school leaders, educators, and tutors. We trace how the district rapidly expanded tutoring under Elementary and Secondary School Emergency Relief (ESSER) funding and subsequently downsized the program after those funds expired. With external pressure easing, the district shifted attention from reach to quality, investing in systems intended to strengthen tutor hiring, supervision, and implementation fidelity. Despite these system-level investments, persistent challenges remained in middle school math, hard-to-staff schools, and schools serving large proportions of students below grade level. These patterns highlight the structural limits of systems building, even as districts transition to more stable, post-ESSER operations.

1. Introduction

In response to pandemic-era disruptions to schooling, the federal government distributed billions of dollars to states and school districts through the Elementary and Secondary School Emergency Relief fund (ESSER) (Boughton et al., 2021; Dusseault & Pillow, 2021). Some of this money targeted evidence-based interventions for student academic recovery (other money went to facilities improvement etc.). Among the academic interventions funded by ESSER, high-dosage tutoring (HDT) was viewed to be particularly promising, a strategy that policymakers and researchers alike promoted (Jacobson, 2023; Stanford, 2022).

Districts used ESSER funding to pursue different HDT models with varied success. Some contracted with third-party tutor vendors (Cortes et al., 2024; Kraft et al., 2022; Kraft & Lovison, 2024; Robinson et al., 2024), but these programs typically operated at a small scale, faced high turnover, and have been costly to sustain without supplemental funding. Opt-in online programs, meanwhile, have shown low participation and negligible impact (Robinson et al., 2022; Schueler & Rodriguez-Segura, 2021). Other districts built “home-grown” tutoring programs. Compared to other approaches, home-grown models offer districts more direct control over tutor hiring and costs, as well as flexibility to adapt to local contexts and programs (Kraft, Edwards, et al., 2024). Despite these seeming advantages, evidence suggests that home-grown tutoring programs have also had limited or null impacts on student achievement at scale (Bhatt et al., 2025; Carbonari, Davison, et al., 2024; Kraft, Edwards, et al., 2024). Regardless of the model, districts have faced a myriad of implementation challenges, such as difficulties with tailoring tutoring to diverse student needs and logistical hurdles like staffing and scheduling (Carbonari, Davison, et al., 2024; Huffaker et al., 2025; Kraft, Edwards, et al., 2024).

With pandemic-related gaps in student achievement persisting, many districts and states are continuing HDT initiatives even after they have stopped receiving pandemic relief from the federal government (DiMarco, 2025). Given the prior track record of these programs, their continued use and implementation raise important questions about how, if at all, school districts are building systems to strengthen and sustain tutoring programs over time. In this paper, we

examine how district leaders in Guilford County Schools (GCS), North Carolina, adapted their home-grown tutoring program after the end of pandemic relief funding. Rather than frame tutoring scale-up as a one-time rollout associated with ESSER, GCS's experience allows us to study tutoring implementation as an iterative process unfolding in a dynamic context. Specifically, GCS's experience allows us to observe how GCS leaders responded to implementation barriers and the steps they took to build a sustainable program.

GCS offers a compelling case for understanding tutoring implementation under a range of conditions. The district used ESSER funds to rapidly scale its tutoring program, which went from serving 5,251 students in the 2021-22 school year to 17,348 students across grade-levels and subject areas by the 2023-24 year. But by the 2024-25 school year, when the district's ESSER funding had ended, GCS strategically downsized its tutoring program. During this phase, the program focused only on students in grades K-3 for early literacy and students in grades 6-8 for math (only 7,700 students total). Although the downsized program was much smaller than the ESSER-driven expansion, it is worth noting that GCS's reduced tutoring program still stands out as one of the largest home-grown initiatives in the nation. Following GCS's trajectory across these phases and conditions offers an opportunity to better understand how school districts address implementation barriers and how those barriers shift under different conditions and pressures.

To deepen our understanding of how GCS responded to these changes, we ask three main research questions: (1) What implementation barriers and supports did district leaders in GCS encounter as they scaled tutoring with ESSER funds in the 2021-22 to 2023-24 school years? (2) How, if at all, did district leaders respond to implementation barriers? What conditions enabled and/or constrained district action? (3) How, if at all, did district and school implementation shift during the 2024-25 school year, the first school year without ESSER funding?

We investigate these questions using a multi-level case study design. We primarily use qualitative methods to analyze public documents, observations of district and RPP meetings, and interviews and focus groups with district leaders and school staff. We find that GCS leaders

encountered well-established implementation barriers during the ESSER-funded expansion of tutoring and attempted to respond to these barriers by investing in system-level supports for the program. These supports included: (1) a more robust personnel infrastructure designed to improve tutor hiring and quality; (2) increasing tutor supervision and support in schools around data use and curriculum planning; and (3) leveraging its Multi-Tiered System of Supports and data dashboards to support greater alignment with high-dosage practices such as small group sizes, session frequency, duration, and pull-out instructional delivery. Importantly, improvements in these areas were easier to consolidate as program scale decreased post-ESSER. However, while district leaders and educators perceived these shifts as strengthening HDT implementation in elementary literacy by the 2024-25 school year, they noted persistent challenges in middle school math, hard-to-staff schools, and schools where large numbers of students were behind grade-level learning, highlighting the limits of systems building and program improvement.

2. Literature Review on HDT Implementation and Scale

Evidence prior to the pandemic created a general consensus that HDT was a promising intervention strategy for improving student achievement post-COVID. A well-cited meta-analysis of pre-pandemic randomized control trials, for example, found a pooled effect size of 0.37 standard deviation units (Nickow et al., 2020), translating into three to 15 months of additional learning (Robinson et al., 2021). This literature also identified the characteristics of more effective HDT programs, which typically involve frequent sessions, small group sizes, skilled tutors, and high-quality instructional materials aligned to grade-level content (Morton & Hashim, 2023; Robinson et al., 2021).

However, a more recent study suggests that these effects may attenuate when HDT is scaled, as it was during the ESSER-funded expansion in GCS. In large-scale programs serving 400 or more students, this newer study found that the pooled effect size drops significantly, leading researchers to conclude that “large effects for tutoring at scale are far from guaranteed”

(Kraft, Schueler, et al., 2024, p. 5). This gap between research and practice resonates with implementation findings during COVID that emphasize implementation barriers related to funding constraints, staffing shortages, scheduling barriers, and fidelity (Carbonari, DeArmond, et al., 2024, 2024; Kraft, Edwards, et al., 2024; Makori et al., 2024). We summarize the evidence on the implementation challenges associated with delivering HDT at scale during the pandemic.

2.1 Time, Space, and Funding

Tutoring creates a host of practical demands. Evidence from the pandemic, for example, suggest that—especially at scale, and on a short timeline—finding classroom space and coordinating schedules can compromise implementation fidelity, leading tutors to deviate from HDT practices by increasing small group sizes, reducing the frequency or duration of tutoring sessions, or using tutors for push-in support during core instruction instead of pulling-out students for targeted lessons in a separate and dedicated space (Carbonari, Davison, et al., 2024; Makori et al., 2024). While some districts have dealt with scheduling issues by offering tutoring before or after school hours, this approach involves transportation and staffing challenges (Makori et al., 2024), and can limit the targeting of tutoring at lower achieving students who may not attend programming outside of school hours (Kraft, Edwards, et al., 2024). It should be said that hiring and scheduling tutors can be time-intensive under any circumstances. But these issues could be amplified under ESSER when districts faced a limited window to spend funds and increased pressures to staff up quickly (Carbonari, Davison, et al., 2024; Hashim et al., 2024).

Finally, although sometimes overlooked during high-funding periods like ESSER, HDT creates substantial costs at scale, including tutor compensation, tutor training and supervision, and the costs of maintaining low tutor-student ratios. Kraft and Falken (2021) estimate that a

HDT program with a 2:1 student-to-teacher ratio could cost districts from \$2,500 to \$3,800 per pupil; an expense that most districts cannot sustain without federal or philanthropic funding.

2.2 *Aligning Tutoring with Diverse Student Needs*

HDT's effectiveness, in theory and practice, depends on the intervention's ability to offer instructional support that is tailored to meet individual student needs (Robinson et al., 2021).

Tailored instruction can, however, be especially challenging to maintain at scale across diverse schools and student populations (Banerjee et al., 2017). This ties back to implementation conditions that compromise model fidelity. Tutors are able to provide individualized instruction via small tutor-student ratios, and frequent tutoring sessions (Robinson et al., 2021). When districts scale HDT across large numbers of students, achieving these conditions is more difficult, with studies suggesting tutoring at scale is less likely to deliver the recommended dosage of instructional hours and weeks (Bhatt et al., 2025; Carbonari, DeArmond, et al., 2024; Kraft, Schueler, et al., 2024). Indeed, post-pandemic evaluations of tutoring suggest districts face a trade-off between *participation rates* and *program intensity* as they sacrifice tutoring hours and small group sizes to serve more students.

HDT's effectiveness also rests on serving lower-achieving and low-income students, who often cannot access extra support on their own (Fryer & Howard-Noveck, 2020; Kraft, 2015; Robinson et al., 2024). When districts grow their HDT programs, targeting can become more challenging (e.g., higher achieving or those less likely to benefit from tutoring may also receive services). If districts *intend* to reach a wider group of students as they scale, they face a different problem: differentiating their program to meet a wider range of student need (Kraft, Schueler, et al., 2024; Nickow et al., 2020). When district-provided curricula and materials are designed for

the average student, tutors need to use their own expertise and familiarity with materials to further personalize content (Kraft, Edwards, et al., 2024).

2.3 *Tutor Hiring, Quality and Turnover*

Studies on tutoring frequently identify staffing as a key implementation challenge. Tutor recruitment, attendance, preparation, and retention all present problems for districts when they are building a tutor workforce, whether from local volunteers or college students (White et al., 2023). While certified teachers can be more qualified than volunteers, and might not require as much training, districts employing teachers as tutors can still face challenges because of limited schedule availability, educator burnout, and limited interest among teachers in assuming additional work responsibilities (Kraft, Edwards, et al., 2024). As districts expanded HDT to support large numbers of students during the pandemic, many faced challenges hiring qualified tutors to meet diverse student needs who were willing to work in schools with greater academic challenges (Carbonari, Davison, et al., 2024; Heinrich et al., 2014; Kraft, Edwards, et al., 2024; Makori et al., 2024). District efforts to hire, onboard, and place tutors are also resource-intensive and time consuming, potentially detracting from other dimensions of the tutor pipeline such as training, supervision, and retention (Carbonari, Davison, et al., 2024).

2.4 *Will and Motivation of Implementation Agents*

Decades of implementation research suggest that district efforts to scale tutoring depend on the will and motivation of front-line workers tasked with delivering and supporting the intervention (McLaughlin, 1987). While teachers may buy into the idea of students receiving tutoring, they still may not perceive tutors as professional peers or have the flexibility in their schedules to coordinate tutoring activities (Kraft, Edwards, et al., 2024; Makori et al., 2024). School principals can also be gatekeepers who enable and/or constrain access to critical inputs

for tutoring, such as student and staff time, school space, and student data (White et al., 2023).

When school principals take ownership of tutoring, they may be more willing to find solutions to curriculum, space, time and funding challenges (Hallgren et al., 2017; Makori et al., 2024).

2.5 Tutor Supervision

Even if schools are able to hire tutors, their effectiveness hinges in part on oversight procedures that can monitor tutoring practices and implementation fidelity (Fryer & Howard-Noveck, 2020; Nickow et al., 2020). In small scale experiments, these procedures include observing tutors with an intervention-aligned rubric and hiring supervisors who regularly check-in and provide tutors with support (e.g., Fuchs et al., 2008). To implement these procedures at scale, districts need a more robust infrastructure of coordinators and data systems (Makori et al., 2024; White et al., 2023). Having a central office team dedicated to tutoring has the potential to improve supervision and communication with tutors in ways that help schools implement tutoring practices more consistently (Kraft, Edwards, et al., 2024). However, districts may underinvest in oversight capacity because of funding uncertainty (Hashim et al., 2024) or political pushback against administrative spending (Kraft & Falken, 2021; White et al., 2023).

While the prior literature identifies the above barriers to scaling HDT, few studies have examined how districts try to address these barriers. This paper foregrounds how districts invest in *systems* to address these barriers and how the viability of those systems varies across circumstances. This perspective builds on existing research that highlights the importance of system structures, norms, and resources (Kraft & Falken, 2021) and extends it by examining how districts develop systems over time. Our findings showcase how one district used its ESSER funding to supercharge its HDT program and then, post-ESSER, refined and downsized that program to continue targeting ongoing student needs for academic recovery. This trajectory

highlights the potential value of scaling back and refining HDT programs to better support implementation processes and school-level ownership of tutoring activities, rather than prioritizing scale and reach.

3. Home-Grown HDT in GCS

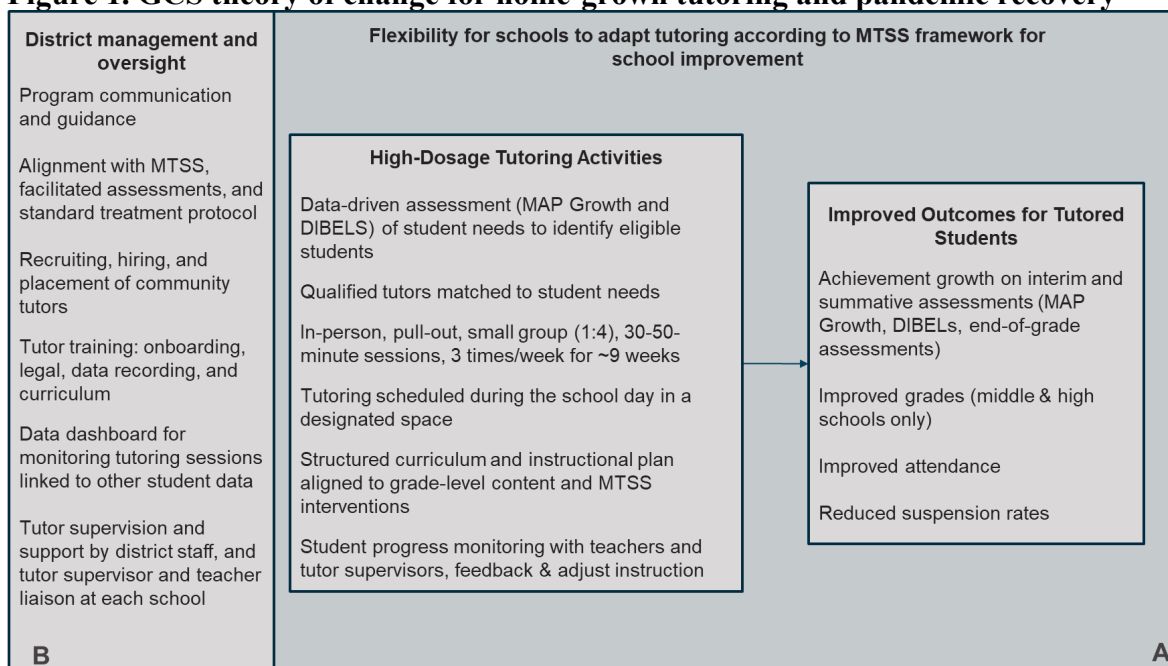
GCS launched its home-grown tutoring program in the 2021-22 school year and rapidly scaled the program to reach almost all of its 126 schools by the end of the 2023-24 school year.¹ During this time, GCS hired upwards of 600 tutors, and more than tripled the number of students participating in a district-provided tutoring session, from 5,251 students in 2021-22 to 17,348 students in 2023-24. GCS also broadened the scope of tutoring, from an initial focus on middle and high school math in Title 1 schools to reading and/or math in all grade-levels and schools. During this phase, GCS relied on ESSER to fund its program budget of \$8-9 million per year.

District leaders tied HDT to GCS' Multi-Tiered System of Supports (MTSS) framework for providing intervention support to students, which was mandated by the state of North Carolina in the 2020-21 school year. MTSS is a school improvement framework that uses a tiered and data-driven approach to meet students' academic and behavioral needs. Tier 1 academic support consists of differentiated grade-level instruction or "core" instruction that reaches most students in a school, Tier 2 consists of supplemental support for students behind grade-level, and Tier 3 consists of intensive support for students far behind grade-level. At its onset, GCS leaders framed their HDT program as an intervention strategy designed to meet the academic needs of Tier 2 and 3 students.

¹ Throughout the pandemic, GCS had participated in a multi-district evaluation of recovery efforts as part of the COVID Road to Recovery project, a research collaboration between NWEA, AIR-CALDER, Harvard University and a group of school districts (Carbonari et al., 2024a, 2024b). The current paper is based on a second, extended research partnership between NWEA, AIR-CALDER, and GCS focused on both program impacts and implementation.

Figure 1 outlines GCS’ initial theory of change for tutoring. Schools were expected to engage in HDT activities that would improve student achievement, grades, and attendance, while lowering suspensions (Box A). GCS asked schools to use DIBELS and MAP Growth *assessment data to identify eligible students* and then *match students with qualified tutors*. The district planned for tutors to provide *in-person, pull-out instruction in small groups of up to four students*, for *30–50 minutes during the school day*, in a *dedicated space, three times per week*, for a *duration of nine weeks*. Tutoring lessons were to follow a *structured curriculum and instructional plan* that was *aligned with grade-level content and MTSS interventions*, with tutors collaborating with tutor supervisors and teachers to *monitor student progress* and *adjust instruction as needed*.

Figure 1. GCS theory of change for home-grown tutoring and pandemic recovery



GCS intended schools to target tutoring at students scoring in the bottom 20th percentile nationally on interim assessments (i.e., Tier 2 or 3). However, tutors were restricted to providing differentiated core (Tier 1) support because North Carolina law requires certified teachers to

deliver Tier 2 and 3 interventions, and not all tutors were certified. Additionally, district leaders granted schools flexibility to adapt HDT to their local context and needs, as reflected in each school's MTSS plan for school improvement. This decentralized approach stems in part from tutoring being tied to differentiated core or Tier 1 support, which can vary across schools, grades, subjects, and classrooms. At the time, district leaders further argued that a decentralized approach was a realistic approach given the wide range of student needs stemming from the pandemic, differences in staffing and scheduling across school sites, and as a mechanism to motivate school actors to integrate tutoring with existing school goals and operations.

District-led management and oversight of the tutoring program (Box B) was intended to balance school-level discretion with fidelity to HDT practices. This involved the district providing guidance on HDT activities and how to align tutoring with MTSS. GCS also established a central tutoring department to recruit, hire, train, and place qualified tutors in schools, drawing on partnerships with NC Ed Corps, local universities, and community volunteers to build a pipeline of tutors. Tutor placement was guided by an allocation formula that prioritized Title I status, average daily attendance, and prior academic performance, ensuring that larger and lower-performing schools received more support.

The tutoring department also provided tutors with onboarding, legal, data recording, and curriculum training to align instruction with district priorities, HDT practices, and adopted curricula and MTSS-aligned interventions. To oversee implementation, the tutoring department worked with designated school-based supervisors and teacher liaisons at each school to provide ongoing supervision and support tutors. GCS also used a data dashboard to monitor student attendance in and dosage of tutoring.

3.1 *Scaling Down and Refinement in 2024-25*

When ESSER funding ended, GCS recalibrated its tutoring strategy (2024–25 school year). This involved both scaling down the overall size of the tutoring program and several other changes. First, district leaders targeted the downsized tutoring effort at specific grades and subjects: early literacy in grades K–3 and math in grades 6–8. This shift reduced the total number of students served in the tutoring program to 7,700 students from 17,348 students served in 2023-24. With an established pipeline of tutors, GCS also prioritized hiring certified teachers and addressed staffing inequities across schools, retaining approximately 175 of the previously 600 hired tutors based on certification status and perceived effectiveness and fit with each school.

With a smaller and established tutoring program, the district more closely integrated tutoring with MTSS by requiring schools to use its standard treatment protocol in MTSS for assigning students to tutoring, identifying curriculum resources and interventions, and monitoring student progress. With a large increase in the proportion of tutors who were certified teachers, GCS now permitted schools to use tutors to deliver Tier 2 and 3 interventions in addition to Tier 1 support and provided more extensive training on its standard treatment protocol and Tier 2 and 3 interventions to tutors. GCS also increased supervision and support of tutors through bi-annual observations of tutoring sessions, monthly curriculum and instruction trainings, communicating firmer expectations for tutor supervisors and teachers to meet with tutors weekly, and providing tutor supervisors with access to real-time data on scheduled and completed tutoring sessions. In the next section, we provide some conceptual framing to help us make sense of GCS’s implementation trajectory.

4. *Conceptual Framework for Understanding GCS’s Tutoring Implementation*

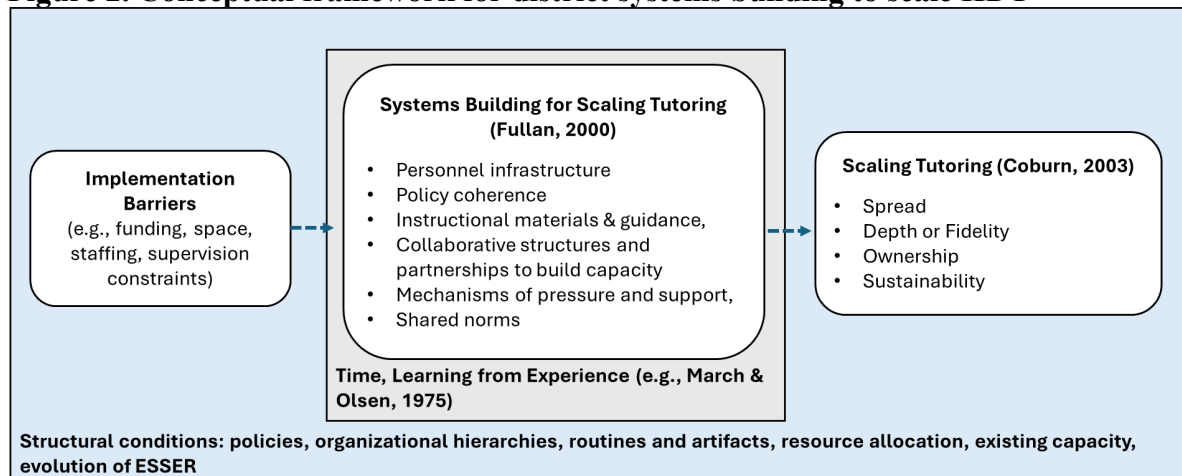
As the above account suggests, the evolution of GCS’s HDT program can largely be understood as *a systems-focused change process* that unfolded gradually in response to early

implementation setbacks, was shaped by existing organizational structures, and was influenced by shifting policy environments such as the sunseting of ESSER.

Consistent with prior work on instructional reform (Coburn, 2003), we understand these system-focused efforts working on several fronts that have been associated with efforts to change system operations *and* target group behaviors at scale: (1) these efforts aim to *spread* practices and beliefs across large numbers of people and organizational levels, from the central office into schools; (2) they emphasize implementation *depth and fidelity* as part of teaching and learning in schools, (3) they balance system moves with an emphasis on *ownership by local implementers*, and (4) and by targeting systems, they aim to support the *long-term sustainability* of the proposed change throughout the district. From this perspective, GCS’s district-level efforts to support implementation can be understood as involving core tasks associated with “scaling” (spreading beliefs; emphasis on fidelity and ownership; investments in central capacity) even as the program contracted. In this way, these dimensions of GCS’s efforts can be thought of more broadly as core elements of efforts to improve implementation across schools. As discussed earlier and illustrated in Figure 2, districts that attempt to scale tutoring can encounter a range of implementation barriers that hinder dimensions of scale. These barriers include challenges with funding, scheduling and space; aligning tutoring with diverse student needs; recruiting and training tutors; maintaining program oversight; and supporting the motivation and will of tutors, teachers, and school leaders. One approach to these challenges is to redesign district-wide systems to strengthen school-level implementation (e.g., Cobb et al., 2018; Fullan, 2000). But importantly, these efforts are not one-shot events; instead, system change to support local implementation can be seen as a process that develops over time and through trial and error (March & Olsen, 1975). As others have noted (Gallagher et al., 2026), district leaders are

uniquely positioned to drive improvement at scale by leveraging the broader system rather than targeting change solely at the individual teacher or school level (intentional district-wide change does not, of course, always happen in practice). Fullan (2000) identifies key district-wide systems that can support school improvement at scale, including building a robust workforce, ensuring coherence between an intervention and existing policies, investing in high-quality instructional materials and guidance, establishing collaborative structures or partnerships to coordinate capacity, developing mechanisms that balance pressure and support for implementation fidelity, and cultivating shared norms across schools and educators.

Figure 2. Conceptual framework for district systems building to scale HDT



Because our ongoing partnership with GCS allows us to observe implementation over time, we observe how system-building efforts are shaped by different structural conditions that might enable or constrain change efforts. A host of structural conditions—such as policies, organizational hierarchies, routines and artifacts, resource allocation, and existing district and school capacities (Giddens, 1979; McDonnell & Elmore, 1987; McLaughlin, 1987)—shape district change and improvement efforts. ESSER provides an interesting—and extreme—case of structural conditions that impact district improvement. ESSER offered unprecedented funding that enabled districts to experiment with new intervention models (e.g., tutoring, small-group

instruction) and launch the early stages of district-wide systems building for sustaining academic recovery. But ESSER also created strong pressures for rapid scaling of interventions, something that arguably made the deliberate, iterative work required for high-quality implementation more difficult (obviously schools faced a host of other pressures during the pandemic as well—politics, safety, etc.). As GCS’s tutoring policy evolved—responding to the initial influx of funding and adapting to the eventual expiration of funds—the structural environment in which tutoring was implemented clearly changed.

Our framework embeds the elements of systems-focused change identified in prior literature within the specific structural shifts that confronted leaders in GCS to better understand how program implementation unfolded. When ESSER funding expired, GCS downsized its tutoring program while making targeted changes to implementation processes associated with fidelity to HDT, and school-level ownership and long-term sustainability. These adjustments did not occur in isolation. Rather, they reflected the district’s response to the policy-driven structural conditions in which it operated, and the ways in which those conditions shifted over time.

5. Data and Methods

This study relies on a rich array of data that was collected as part of an ongoing, mixed methods research practice partnership (RPP). The RPP was born in response to the COVID-19 pandemic and was focused on the implementation and impacts of pandemic-era academic interventions. Following a multi-level case study design (Creswell & Poth, 2016; Yin, 2003), we use multiple data sources across organizational levels in GCS to describe the scaling and implementation of tutoring, and to situate these findings in context. Our document-based data sources include presentations of recovery efforts, inventories of district programs, draft logic models, and district guidance shared with schools. We also collected qualitative data from observations of district-led meetings on the design, adaptation and continuation of COVID

recovery programs, and meetings between RPP researchers and GCS leaders on the design of recovery programs and preliminary research findings. Our primary data source is 79 interviews and four focus groups with district leaders and school leaders and staff from a purposive sample of GCS schools, which we describe in more detail in the following section.

5.1 Interview Sampling and Data Collection

Based on input from district representatives on the RPP team and interview participants, we identified 33 district leaders, administrators and staff, and interviewed 29 participants across levels of authority and roles in the central office. District leader participants included senior leaders (e.g., Chief of Schools, Chief of Academics), directors of recovery programs and related departments (e.g., tutoring, summer school, MTSS), and district staff or intermediary partners who provided direct support to schools (e.g., school principal supervisors, tutoring coordinators). We interviewed 18 district participants in spring to summer of 2023, eight participants in spring 2024, and nine participants in fall 2024. Participants overseeing and supporting tutoring were interviewed at multiple points. To preserve the confidentiality of district participants, we refer to them as “district leaders” or “district administrators” rather than by their specific role.

To document tutoring implementation in schools, we selected five schools in GCS that varied in academic performance (i.e., state accountability letter grades), school size, student demographics, and location. We prioritized sampling from elementary and middle schools to align with the MAP Growth outcome data from grades 3-8 used in the quantitative strand of the RPP. Table 1 shows that our purposeful sample of schools are situated in diverse contexts across GCS, including three elementary and two middle schools, schools with accountability letter grade ratings from A to F, urban and rural settings, and varied student demographics.

Table 1. Case Schools and School-Level Participants

Case	Level	Location	Accountability Grade	Racial Composition	Interview Participants from 23-24 and 24-25 school years
#1	Elem	Midsize City	A	Non-White: >50%	1 School Principal 1 Curriculum Facilitator 5 Teachers (grades 1-5) 4 Tutors
#2	Elem	Rural Fringe	C	Non-White: <25%	1 School Principal 1 Curriculum Facilitator 2 Teacher (grades 3-5) 3 Tutors
#3	Elem	Large City	F	Non-White: >75%	1 School Principal 1 Curriculum Facilitator 2 Teacher (K-3) 2 Tutors
#4	Mid	Large City	C	Non-White: >50%	1 School Principal 2 Curriculum Facilitator 1 Teacher (grade 8) 1 Tutor
#5	Mid	Large City	D	Non-White: >50%	1 School Principal 1 Curriculum Facilitator 2 Teachers (grades 6) 2 Tutors

In the 2023-24 and 2024-25 school years, we interviewed the school principal and tutor supervisors at our study schools and worked with them to identify a sample of teachers and tutors across grade-levels and subject areas. During the 2024–25 school year, one school (School #4) exited the study after merging with another campus and delaying its tutoring implementation beyond our data collection window. In total, we interviewed 26 out of 42 sampled school-level participants in the 2023-24 school year, and 18 out of 27 sampled participants in 2024-25.

To compare our study schools’ findings with a broader sample of schools in GCS, we conducted four focus groups with elementary and middle school tutors and tutor supervisors in spring 2025. We conducted focus groups separately by role and school-level and included between four to five participants each. Tutors and tutor supervisors were sampled from schools

that varied in tutoring participation and dosage rates as monitored by the district.

We hosted interviews and focus groups online (via MS Teams). In both cases we followed a semi-structured protocol and spoke with participants anywhere between 40 to 95 minutes. District leader interviews initially focused on district goals for tutoring, intended inputs and activities, expected student outcomes, and program monitoring. We also asked about guidance and support provided to schools, implementation successes and challenges, contextual factors shaping implementation, district learning on how to scale tutoring, and areas in need of improvement. Follow-up interviews focused on district responses to implementation barriers, actions taken to improve tutoring implementation, and the conditions that enabled and/or constrained iteration and improvement over time. School interviews focused on school goals for academic recovery, procedures for identifying students for tutoring, tutor staffing and qualifications, implementation of HDT procedures (see Figure 1), and perceived outcomes of tutoring. Throughout our data collection we asked probing questions about how tutoring procedures shifted over time, and about how local factors and barriers shaped implementation.

5.2 *Data Analysis*

All interviews were recorded and then transcribed using Otter AI. Transcripts were coded by the research team using Dedoose, a cloud-based application for analyzing qualitative data. Our coding procedure began with a list of preliminary codes of anticipated tutoring inputs, activities, and outcomes based on GCS's theory of change (see Figure 1) and our interview protocols (e.g., implementation barriers, district/school context, district learning and improvement, changes in tutoring program design and implementation). We monitored inter-rater reliability (Creswell & Poth, 2016) to ensure consistent coding across researchers. After reviewing and reaching agreement on code applications across an initial subset of interviews, the

research team double coded all the interview transcripts.

After reviewing coded excerpts, we drafted analytic memos at the district and school-levels to develop claims about GCS's tutoring implementation (Miles et al., 2019). Memos elaborated on each component of Figure 1 and how they shifted over time, and triangulated perspectives across interview participants, documents, and observation notes. For example, our district-level memo described district goals for pandemic recovery, district leaders' intended theory of change, district activities for managing and supervising tutoring, processes and observed trends when scaling up and scaling down tutoring, implementation barriers, and federal, state, and local contextual factors that shaped policy design and implementation. At the school-level, we elaborated on school goals, context, implementation of HDT activities and if/how they aligned with district expectations, contextual factors that enabled and/or constrained implementation, and how dimensions shifted over time. Matrix displays summarized patterns in school implementation over time and across case sites and allowed us to identify key barriers.

This analysis identified three key implementation barriers involved with scaling tutoring in the 2021-22 to 2023-24 school years: (1) hiring and placing qualified tutors at pace with the rapid scaling of tutoring in GCS; (2) training and supervising tutors to ensure instructional alignment with district-approved curriculum resources and individual student needs; and (3) ensuring schools implemented high-dosage activities such as intended session frequency, duration, pull-out instruction, and small group size with fidelity. Once identified, we re-examined excerpts related to these barriers. Specifically, we used ideas from our conceptual framework to examine how barriers identified by district-level leaders were influenced by district-led decisions and responses from schools and tutors (research question 1) and examined how district leaders responded (or failed to respond) to these barriers at the system level. In both

cases, we considered how implementation dynamics were shaped by shifting organizational and structural constraints and affordances (research questions 2).

We then wrote additional memos about the relationship between district efforts at system-building school-level implementation in the 2024-25 school year, with a focus on the spread, depth, ownership, and sustainability of the tutoring program (research question 3). At this stage, we brought in themes and related excerpts from focus groups, which largely confirmed and elaborated the patterns we observed in our study schools. When differences emerged—between district and schools, across case sites, or by subject and grade level—we looked for contextual factors that might explain the divergent patterns. These included variations between elementary and middle schools, and early literacy and math tutoring. Finally, we member-checked our claims and understandings with district leaders (Merriam & Tisdell, 2016).

We should note that while our conceptual framework draws on ideas from in-depth theoretical studies of scaling instructional reform and organizational change, the data in this study were not collected with these theories in mind a priori. Instead, we found concepts from this literature served as useful interpretive tools that helped us make sense of the patterns that we saw emerging from our interviews and observations. In other words, we applied these concepts as part of an iterative, inductive analytical process to better understand how district adaptations of tutoring unfolded in response to implementation challenges.

6. Findings

Like many school districts, GCS leaders reported significant learning gaps among students as a result of the pandemic and framed tutoring as a necessary, evidence-based intervention to accelerate recovery and close these gaps. As one explained, “when our students came back to school [after remote-learning], we saw lots of gaps in math and also literacy and we wanted to make sure that we can act quickly to close gaps, give kids additional opportunities

for tutoring, to be able to meet them where they are and help them grow.” Another added that tutoring was a research-based compliment to teachers’ efforts in the classroom: “instead of the teacher spending 40 minutes going backwards on skills...the tutor [can] provide some of those skills supports...And research has shown us that if they do it at least three to four times a week, for at least 45 minutes, then they get that content faster.”

As GCS began scaling tutoring using an infusion of ESSER funds, leaders quickly realized it demanded district-wide systems for hiring and placing tutors, providing supervision and support, and establishing intervention protocols and data reporting to ensure schools adhered to high-dosage activities such as data-driven assignment, small groups, pull-out instruction, and frequent and sustained tutoring sessions. Below, we document GCS’ efforts to build these systems. For each section of findings, we describe implementation barriers facing the district during the ESSER-funded expansion of the program (from the 2021-22 to 2023-24 school years) and the downsizing phase of the program (the 2024-25 school year), when ESSER funding went away. We also describe how school-level implementation of tutoring shifted in response to the district’s systems-level effort.

6.1 Tutor Hiring and Perceived Quality: Building Efficient and Sustainable Personnel Infrastructure

GCS made several changes to its system-level staffing infrastructure during and after the ESSER funding. Early efforts during ESSER focused on urgent fixes, like streamlining HR processes to speed tutor placement. Over time, GCS gradually built a routine system for hiring tutors into schools, one that was shepherded by proactive leadership from the district’s central office tutoring department and cross-department collaboration with HR, finance, and technology. With the downsizing of the program post-ESSER, district leaders reported they faced a surplus of tutors. Drawing on three years of experience and data from the ESSER-funded era, district

leaders perceived that, post-ESSER, they were better positioned to identify, retain, and deploy tutors they viewed as higher quality—primarily certified educators—from a surplus labor supply.

At the same time, school principals and tutor supervisors reported they had developed their own expertise in recruiting, hiring, and staffing tutors during the ESSER-funded era and, post-ESSER, could work closely with district leaders to retain tutors during the downsizing. Indeed, schools reported assuming greater responsibility for shaping and maintaining their tutoring workforce in the downsizing phase. Prior experience, however, meant that schools also held more shared norms and values about retaining tutors viewed as effective by school leaders (i.e., spread). In general, school leaders preferred to keep tutors who were certified educators and aligned with their instructional routines.

Tutor shortages, hiring inequities, and mismatches in expertise during ESSER. As shown in Figure 1, GCS’s theory of change depended on hiring local, in-person tutors, supported by a newly created tutoring department in the central office. Despite substantial investments and a tutor pool that peaked at 600 tutors in 2023–24, demand continued to outpace supply, especially in secondary, remote, and other hard-to-staff schools. As one district leader said, “there is always a need for more tutors.”

This was not just a generalized problem. Subject specializations faced different staffing challenges. District leaders noted, for example, that tutors overwhelmingly preferred working with younger students and avoided secondary math courses. As one leader explained, “Everyone wants to work with little kids and help them read... but when you're talking about Math 1, Math 2, Math 3...that’s more specialized...” Geographic hiring patterns created other differences: schools generally preferred to hire local tutoring candidates, but the highest-need schools often

faced a small pool of potential tutors. “We need to think about how to funnel the support into our highest need schools,” argued a principal supervisor in the district who noted the problem.

To meet the expanded demand for tutors during the ESSER-funded expansion, GCS turned to university partnerships to find tutors in secondary subjects and other hard-to-staff schools (despite concerns that college-age tutors were not as effective as retired teachers). In a district meeting focused on planning tutoring for the 2024-25 school year, an administrator cautioned against placing “reading buddy trained tutors” with the most fragile kids. Issues with tutor professionalism and reliability, particularly among high school student hires, led some schools to recruit their own tutors for better fit and consistency. “They sent us some high schoolers who showed up sometimes and didn’t show up... then they quit,” one tutor supervisor recalled. “So that’s why we went and recruited our own people.” These barriers—shortages, inequitable hiring, mismatches in tutor expertise, and onboarding inefficiencies—collectively constrained the district’s ability to place qualified tutors where the need was greatest even when they had ESSER resources to expand the program.

Building an efficient and sustainable hiring infrastructure. District leaders responded to hiring challenges during the ESSER-era by streamlining HR processes and placing tutors in schools more quickly. These actions were driven by a sense of urgency. As one leader explained, “our dilemma was that we knew if we stopped [hiring tutors] that meant...students who were not receiving that additional service during the school day, so we didn’t stop, we just kept going.” With a short window to spend ESSER funds and a commitment to connect as many students with tutors as possible, leaders focused on adjusting hiring processes within their control.

They began by expediting HR systems that were overwhelmed by tutor applications. A tutoring department administrator recalled, “we have more applicants than we can process is

really what it is...there's been an overwhelming response from the community, retired educators that want to come back and tutor...And I think that it all met with our current systems and processes that we have now.” To address bottlenecks in onboarding and payroll, the tutoring department worked across HR, finance, and technology departments to track hiring progress, troubleshoot delays, manage payroll, and communicate directly with tutors. One leader noted:

We took on payroll for tutors, so that we knew exactly what tutors were getting paid. And if tutors weren't getting paid, they would reach out to us and we would investigate...we're in a good place as far as thinking creatively on things that we could do differently as far as the hiring process...I just feel like every time you know, we've run into something, it's like you hit a, you hit a brick wall, and you're like, oh, man, so I'm just hoping that we don't hit another brick wall.

Beyond expediting hiring and managing payroll, district leaders realized the need for a more durable infrastructure for tutor hiring. Putting that infrastructure completely in place took until the 2024-25 school year, when the pressure of ESSER expansion had subsided. “I’m happy to report that we have since figured out a process,” a district administrator shared. “It took a lot of communication back and forth and we were told ‘no’ a lot, but we didn’t give up.” The breakthrough came when HR created part-time school-based positions for tutors, allowing onboarding to be automated through standard district processes for hiring staff. This shift underscores how existing district structures for hiring—initially rigid and resistant to change—had to be repeatedly negotiated and reworked to create a sustainable hiring system.

Post-ESSER, strategic down-sizing improved tutor hiring in schools. As they built out a more durable infrastructure, district leaders reported doing more to prioritize tutor quality in 2024–25—when GCS downsized its tutoring program and when ESSER funding was over.

With a strong pool of 600 tutors, district leaders could afford to be more selective. “We’ve got three years of data underneath our belt from ESSER,” one leader noted, underscoring the district’s readiness and capacity to prioritize finding effective tutors among the post-ESSER surplus.

To do this, the tutoring department worked closely with schools to collect feedback and retain tutors that district and school leaders perceived to be effective, based on tutor certification and experience, tutor session data, achievement outcomes of tutored students, and school leaders’ and staff impressions of tutor quality. The department also reassigned certified educators to unfilled roles across the district, especially in Title I and middle schools. One leader described this process as a necessary “shuffling of the deck” to place tutors where they are needed, not just where they want to be. These efforts substantially increased the share of tutors with teaching certification: in 2023-24, less than 22% (out of 600) of hired tutors were certified teachers; by 2024-25, the share of tutors who were certified teachers was almost 80% (of 175 tutors).

After three years of hiring and managing tutors, schools had developed a clearer sense of which tutors were best matched to schools. As a result, school leaders began actively advocating for their staffing needs. This growing school-level ownership reinforced district efforts to improve tutor quality and promote more equitable staffing by the 2024–25 school year. Norms and priorities around tutor selection—especially around the value of teacher certification—reportedly spread and became more consistent across the district.

At one elementary school, for example, a tutor supervisor emphasized the added value that certified educators bring in terms of teaching experience and content expertise over community volunteers. They described their tutor, who had worked at the school for the last two years, as someone who “knows how to identify kids who need to work on phonemic awareness

and how to implement interventions in that area... They definitely know what they're doing. And that's one of the benefits of having a retired teacher versus a community tutor." A middle school tutor supervisor similarly described their tutors as a "perfect fit" for the school due to their grade-level expertise, prior work with multilingual learners, and "faithful" and "consistent" work ethic, qualities they perceived based on the tutors' prior track record at the school.

Beyond educator credentials and grade-level or subject expertise, schools described other qualities that made tutors successful in their context and worked to retain those who demonstrated those qualities over time, such as familiarity with staff and students, collegiality, self-sufficiency, and instructional skills. A tutor supervisor at a hard-to-staff school stressed the importance of finding a tutor who "matches with your school." Describing their tutor as "familiar with the students [and] the teachers," the supervisor shared how they advocated for the district to hire their preferred candidate over existing tutors in the system. "The district wanted to place the tutors that they already...[had but], we don't have anybody that's trying to come over here. So, we advocated for [our candidate], and they gave them to us."

Challenges in middle and hard-to-staff schools. Even as GCS made improvements to tutor hiring in 2024–25, it faced lingering challenges in middle, remote, and hard-to-staff schools. These challenges included tutor turnover and mismatches between tutor preferences, expertise, and school and student needs. Across the study schools, middle schools and elementary schools serving a majority of non-white students either struggled to fill tutoring positions, hired candidates late into the summer, and/or hired candidates that did not have the subject expertise to match tutoring assignments even during the downsizing phase of the program. Middle schools faced acute challenges matching tutors' instructional backgrounds with math assignments. As one district leader explained, even if tutors were certified to teach in

middle school, it did not mean they had secondary-level math expertise. Math tutors were also not experienced with current teaching methods and needed time to learn new ways of solving math to meet teacher expectations. This challenge was compounded by the wide range of student needs at the middle school level, from students who were significantly below grade level to advanced learners. This range of needs required tutors to have expertise across grade-levels and sub-domains in math. As one tutor explained,

I have students that need a number line or that need base-ten blocks...and I'm thinking I taught this in second grade...but then I have students that are doing things that I did in 10th and 11th grade. Those are the nights that I sit in front of those doggone videos...I have had to go back and teach myself, because I've never taught middle school.

6.2 Tutor Supervision: Evolving Oversight and Support Structures

GCS's experience with tutor supervision highlights how the district shifted oversight and support processes to emphasize the use of assessment data and MTSS-aligned curriculum to tailor tutoring to student needs. As with hiring, the district's oversight journey also demonstrates how tutor supervision can take a backseat to rapid scaling and the urgent need to hire tutors during ESSER. By the 2024-25 school year, district leaders' reflections about the lack of district oversight, coupled with reduced demand for tutor hiring and the tighter integration of tutoring with MTSS, paved the way for more supervision, training, and support for early literacy tutors. These strategic changes, combined with the downsizing and smaller tutor caseloads, were perceived by district and school leaders as contributing to more consistent data-driven and curriculum-aligned early literacy tutoring sessions. As with hiring, school-level ownership of tutoring appeared to increase, as tutoring became more tightly integrated with existing core instructional activities. At the same time, gaps remained in middle school math where district

training had not changed in response to tutor needs, MTSS resources were more limited, and tutor supervisors were less available to meet with tutors.

Limited tutor supervision and support during ESSER. According to the district's theory of change (Figure 1), GCS planned for both district-level and school-level actors to guide tutors: district leaders would train and observe tutors, while school-based supervisors and teachers would collaborate on planning, progress monitoring, and ensuring alignment with MTSS. However, the intense pressure to hire tutors during ESSER eclipsed these oversight activities at the district-level. One leader explained that the tutoring team was “mostly in onboarding mode...because tutors are coming on all throughout,” creating “more of an administrative relationship between the tutoring team and the schools.” With only three central office staff supporting over 100 schools, district leaders acknowledged the difficulty of understanding what was occurring instructionally: “It’s hard to know the quality piece...because it’s hard to understand exactly what is going on in schools.”

In the absence of district-level supervision, instructional leaders within schools (principals, tutor supervisors, and teacher leaders) became de facto tutor supervisors. This decentralized approach mirrored broader district initiatives, such as curriculum adoption and MTSS implementation, that relied heavily on school-level capacity. Moreover, district leaders were aware that school supervision and support for tutoring would be uneven across schools. As one district leader put it, “people have in their minds that a flavor [of tutoring] is happening more than it is probably,” but in reality, “schools...just don't have [leadership that is]...intrinsically motivated and has some, like, prior knowledge [of] tutoring.” They went on to share that, in many cases, school staff might see tutoring as “just another thing they have to do...not that they don't want it to succeed, but they don't know how, or they don't have the time.”

Downsizing post-ESSER shifted focus to tutor supervision and support. District leaders recognized that tutors needed closer supervision and support but were constrained in their ability to provide it because of limited staff capacity and pressure to rapidly scale tutoring during the ESSER-funding period. As noted earlier, the tutoring department lacked the manpower to maintain consistent engagement with tutors while building up the district’s pipeline of hired tutors. One leader explained that they could not offer ongoing trainings on topics such as the “mentorship piece...[and] socioemotional learning” because their focus had to remain on getting tutors into schools. Another described the situation as “holding on to a runaway train,” noting that “stopping means that a tutor is not going to get placed at a school and working with a student, so I think that is where we [had] to put our focus.” These reflections illustrate how structural limitations and external policy pressures narrowed the range of actions the district could feasibly take and curtailed early attempts at systems building for scale.

Following the end of ESSER funding, however, the district had more capacity to attend to oversight structures as hiring demands decreased. With more time and stability, leaders reflected on prior gaps and made substantive adjustments to supervision structures, strengthened alignment between tutoring and MTSS, and expanded training—particularly in early literacy—though challenges in math supervision and support continued.

Shifting from administrative to instructional support. With fewer tutors to manage and hiring concentrated in the summer, district leaders reported that they had more capacity to attend to instructional quality. As one leader shared, placing tutors before the school year enabled a focus on “quality of pre-service training and increasing the support for tutors in schools.” The tutoring department also hired a coordinator to observe sessions twice yearly in every school and became more explicit with principals and tutor supervisors about expectations

for weekly planning meetings. Leaders viewed this demand as feasible because school caseloads of tutors had dropped; as one leader noted: “we’re talking about one to three tutors [per school], not 12.” With a small central office staff, the tutoring department was still unable to provide hands-on supervision to tutors in all schools, so school-level oversight remained essential.

Integrating tutoring with MTSS. To sustain tutoring without ESSER funds in 2024-25, GCS relied on other federal school improvement funds, which required integrating tutoring with other improvement strategies, such as MTSS. This funding shift created an opportunity to build coherence between tutoring and MTSS and opened doors for the tutoring department to collaborate with curriculum and MTSS teams. Importantly, the shift allowed schools to use tutors for Tier 2 and 3 interventions if the tutor was a certified teacher. Commenting on the tighter integration between tutoring and MTSS, a district leader shared, “I think that came about when [the tutoring department]...finally got to meet with the MTSS director. And really having them put down on paper exactly how they see tutoring fitting into the MTSS framework.”

The district’s integration of tutoring with MTSS focused heavily on early literacy due to more robust MTSS protocols and state legislation emphasizing the science of reading. One leader noted, training was “more intense because early literacy is a huge focus and it is where all our legislation lives,” and the district had “altered how [tutors] will support students and in turn that alters how we train them.” Under this new approach, certified K–3 literacy tutors received extensive training for delivering Tier 2 and 3 interventions aligned to the district’s standard treatment protocol. In practice, teachers and tutors also used these resources to differentiate core instruction in Tier 1, ensuring alignment in content and instruction across tiers. District materials emphasized that the standard treatment protocol saved time and resources, reduced the need for additional tutor training, and ensured interventions were replicable and evidence based.

In math, district changes to tutor supervision and training were constrained by a weaker MTSS infrastructure and fewer high-quality resources. One leader acknowledged that “in math there is just less resources for Tier 2 and Tier 3 diagnostics...there's less high-quality resources for some of that [for] math [than] English.” Other challenges included tutors’ limited access to intervention materials and the district’s reliance on Zearn—a digital tool designed for independent student use that does not require a tutor’s presence—as its approved Tier 2 and 3 resource in the 2024–25 school year. One tutor supervisor shared, “I didn’t ever really feel like we had adequate resources for math interventions...we [are] kind of flying by the seat of our pants.”

Data-driven conversations and MTSS-aligned curriculum planning. As GCS expanded tutor supervision and support in 2024–25, elementary school supervisors reportedly shifted from basic oversight to more substantive, MTSS-aligned instructional guidance. This shift suggests additional school-level ownership and sustainability of HDT, as tutoring became more deeply integrated with schools’ existing core instructional processes. Specifically, elementary school tutors and supervisors perceived three major changes in HDT implementation in the downsizing period that relied on federal school improvement funds: (1) more consistent use of student data to diagnose needs; (2) curriculum planning explicitly aligned to MTSS resources; and (3) greater integration of tutors into tutor planning and decision-making. Because these shifts coincided with a smaller number of tutors in schools and across the district, the time and space burdens associated with school-level supervision decreased.

As one elementary supervisor explained, data-driven conversations were limited during the last year of ESSER funding “simply because of the [large] number of students they were serving with their schedules.” A literacy tutor recalled “assigning lessons based on what [they]

have observed [which was] kind of middle of the road or a shot in the dark.” By contrast, the smaller caseloads of tutors post-ESSER meant that supervisors are now able to “strategically plan” discussions about students when tutors are onsite “because we need their input as well.” Another supervisor described using regular check-ins to share curriculum aligned to student needs: “I have been giving her [the tutor] resources based on the data from data meetings with the teachers. And what targeted areas they want her to focus on.” These excerpts suggest deliberate efforts by tutor supervisors to integrate tutors into core instructional processes.

Experienced tutors, especially those with teaching certifications who were now over-represented in the tutor pool in the post-ESSER period, were well positioned to act on the increased guidance provided by tutor supervisors. One tutor emphasized how these experienced tutors were already fluent in using MTSS-aligned materials to tailor lessons to student needs according to their data: “I look at the areas that they need to work on, and I use some materials I got from that was recommended by our county office. And I also use materials that I know works...It's things that were already in house.” For tutors with less experience in early literacy, the district expanded professional development, particularly the pre-service and ongoing MTSS-aligned trainings which strengthened their ability to design lessons and be responsive to student needs. Ongoing training also allowed tutors to learn alongside other tutors, fostering professional collaboration that they otherwise had little time for in schools.

Challenges with middle school math. As was the case across a range of issues, the perceived gains in school-level ownership and sustainability of HDT activities in elementary literacy did not materialize in middle school math, where supervision and support remained inconsistent and limited opportunities existed for data-driven or curriculum-aligned planning. In middle schools, tight schedules left supervisors little time for instructional conversations. One

tutor explained that their supervisor “doesn’t have a lot of time to meet with us [but will] try to touch base... but there is no direct [guidance].” Another described check-ins “every like, week or two,” acknowledging the need for more guidance. They went on to explain, “it’s always a little tricky when you’re... administering someone else’s curricular stuff.”

Without routine supervision or teacher input, many middle school tutors relied on their own judgment: “I’m not getting input from teachers, so I have to use quick observations and conversations with kids to figure out what to focus on,” expressing concerns about misalignment with classroom instruction. A tutor supervisor confirmed the reliance on tutor discretion, noting tutors “are able to see where the gaps are and then give [students] resources on that.”

Because school-level guidance was limited, middle school tutors might have turned to the MTSS standard treatment protocol; however, in middle schools this protocol centered on Zearn, which tutors saw as inappropriate for in-person instruction. Many instead recreated classroom materials or developed their own. One tutor explained that they thought pencil-and-paper instruction allowed them to better understand students’ thinking: “I know more about their thought processes when we’re doing it like a pencil paper interaction type [of] thing...” Another added, “I don’t use Zearn either...when I do pull them out, it’s pretty much go to the board and show me what you know.”

6.3 *HDT Implementation: From Flexibility to Fidelity in Targeting and Dosage*

Over several years of expanding and then downsizing its tutoring program, GCS increased the tutoring dosage individual students received to better reflect the program’s original theory of change. For example, the average tutoring hours provided per student rose from 13.88 hours in 2022–23 to 14.44 hours in 2023–24 and 14.86 hours in 2024–25. However, this increase in hours provided still fell short of the program target of three 30-minute sessions for at least

nine weeks (~40 hours). Average *sessions* per student also increased from 26.21 to 28.04 to 30.09 sessions over the same years, gains that approach the program target of at least 27 sessions annually. The increases in hours and sessions post-ESSER are consistent with greater fidelity of implementation (see Figure 1, Box A).

Even as district averages improved, however, substantial variation across schools persisted. The depth of change in teaching and learning conditions remained uneven, prompting additional targeted action in 2024–25. Recognizing the need for still greater standardization, district leaders used real-time data systems to monitor tutor attendance and dosage to more tightly integrate tutoring services with MTSS. Smaller caseloads at the elementary level further enabled higher-dosage activities with greater fidelity, although middle schools continued to face resource and scheduling barriers that made achieving program fidelity more challenging.

Flexible school-level implementation during ESSER. During the ESSER-funded expansion, leaders emphasized school level-autonomy over implementation. “We want to give schools the autonomy to do what works for their school,” said one leader. School-level flexibility was also a practical choice, given the varying staffing, schedules, and resources in the 100 schools implementing the program. “Every school is very different...a lot of it is funding based per school,” a leader noted, explaining that better-resourced schools could dedicate more staffing and time to adhere to high-dosage principles but that under-resourced schools might not be able to do the same. Such remarks hint at the role structural differences across schools play in shaping implementation, including differences in funding levels, scheduling constraints, and variation in grade levels and subjects served.

In practice, school-level flexibility created wide variation in how schools targeted students for HDT, formed groups, and managed session frequency and duration. The result was

limited fidelity to HDT activities and an uneven depth of change to teaching and learning. This variation persisted through 2023–24, as illustrated by variation in student targeting and assignment practices. GCS’s program was supposed to target tutoring at students in the bottom 20th percentile or those eligible for Tier 2/3 interventions. However, state policy required that Tier 2/3 services had to be provided by certified educators. With a shortage of certified tutors during the ESSER expansion, schools provided tutoring to a wider set of students. For example, some schools targeted “bubble” students closer to the 60th percentile. Others included Tier 1 students between the “20th and the 40th percentile.” Some excluded the lowest-performing students because English learners and EC students were already receiving other services: “Some of our lower performing students are already identified as English learner or Exceptional Children students...So how do we fill in those gaps...when the time to do so is so limited?”

Inconsistent targeting had cascading effects on other aspects of HDT, like duration, group size, and instructional delivery. Tutors described groups expanding beyond district expectations, as teacher discretion, student absences, scheduling instability, and limited supervision pushed tutoring groups to larger numbers of students. Commenting on how control over assigning students to tutoring and forming small groups devolved to teachers, one literacy tutor shared: “At the beginning of the year...the [tutor supervisor] would identify students based on their testing data...[but] they [then] handed that process to the grade-level teachers who are now assigning the students...that is why we had an increase in the sizes of the student groups.” Another tutor experienced four scheduling changes in two months, with student assignments driven by teacher input rather than assessment data. A middle school tutor noted that their school tried to maintain tutoring assignments for “six to seven weeks—shorter than GCS’ requirement of nine weeks—but that they would inevitably need to make “adjustments [to groups] over that time period.”

Instructional delivery also diverged from the program intent. Some schools used pull-out sessions as intended by district leaders, others used push-in support, and some abandoned tutoring altogether due to staffing shortages, large class sizes, and chronic absenteeism. At one case school, a tutor supervisor described both pull-out and push-in tutoring models. “Some students who are struggling readers...meet with the same tutor three times,” one tutor supervisor noted, “but others are supported based on data from last week.” In one of the middle school cases, tutors were reassigned to support core instruction. “People are drowning at the school level,” one middle school tutor shared, “and I am being placed where they actually need people.”

In this fragmented landscape, tutoring became reactive, shaped by immediate classroom needs rather than the goal of providing a targeted group of students with frequent, consistent, and sustained support. In some ways, these implementation challenges reflect tensions between pressures to expand with fidelity and reliance on decentralized decision making. At a district meeting, one leader reflected that GCS wants to scale tutoring up as big as possible, to help as many kids as possible, but also get them the proper dosage to see those gains, and that it left it up to schools to navigate the tension between these goals and their related tradeoffs.

Standardization post-ESSER through MTSS and data systems. District leaders viewed inconsistent school-level implementation of HDT activities during the ESSER expansion period as a signal that they needed to build district-wide structures to support implementation fidelity. Leaders emphasized building stronger coherence between MTSS and tutoring to standardize HDT activities such as group size, session frequency, duration, and instructional delivery. In 2024–25, post-ESSER, tighter MTSS integration using federal school improvement funds enabled the district to clarify that certified tutors could serve Tier 2/3 students and encouraged schools to follow standard treatment protocol routines, including grouping students

by data, ensuring minimum thresholds for frequency and duration, and tracking data over time to monitor student progress. “With these guidelines and really defining how tutoring fits into the MTSS framework...we're being much more specific in how tutors can be utilized in the schools,” a district administrator explained. Another highlighted the role of diagnostics in MTSS to guide student progress monitoring and duration of tutoring: “we give guidance on how many data points should be collected...looking at trend lines and making sure that students...are on track to meet their annual goals.”

As GCS scaled its tutoring program, it also contracted with external vendors to build a data infrastructure to monitor student academic progress and tutoring dosage. This emerging data system complemented the district’s prior approach of educating school on high-dosage practices by introducing a layer of accountability. The system had evolved from having tutors manually submit attendance in Word documents, to an online attendance platform in 2022–23, and then to a 2023–24 dashboard that displayed tutoring participation alongside academic performance and other indicators (e.g., attendance, discipline). Collaboration with the data department enabled access to real-time data that allowed the tutoring department to intervene proactively in schools. One administrator explained, “when the data team reaches out... ‘hey, this school was logging a ton of sessions. But we’re not seeing much growth,’ we know to target that school at that moment.” By 2024–25, schools were given access to data, enabling self-monitoring and reinforcement of high-dosage practices.

Greater perceived school-level fidelity in HDT activities in early literacy. In response to district changes, elementary schools perceived an increase in depth and consistency in HDT activities in the downsizing period. Structured pull-out sessions during frequent intervention blocks replaced mixed models of instructional delivery; group sizes reportedly dropped to around

four students grouped by need; students stayed in tutoring for at least nine weeks; and schools maintained a consistent, data-driven process for assigning students to tutoring. “Right now, currently I’m only doing pull out...Last year I did...a little bit of push in,” one tutor shared. “You’re not going to have a kid...working on segmenting phonemes with another kid who’s struggling to read words...The kids [are] grouped based on need,” another tutor supervisor explained. Another tutor echoed the move toward more precise targeting of services: “With even smaller groups and more specific emphasis...we’re working on the ones trying to get them over the hump, so we divide [students] by need base.”

Elementary schools also appeared to reduce disruptive “switch-outs,” enabling frequent and sustained intervention for at least nine weeks. One tutor supervisor noted: “I haven’t had teachers this year ask to switch anyone out yet. So, I think that this understanding is like, this is the group, and this is how it’s going to be, until we decide that if they’ve met mastery then we can consider switching someone out.” Another supervisor shared that they closely track the frequency of tutoring sessions on the district’s data dashboard to ensure tutors are meeting expected targets. Tutors likewise described working with the same students daily for an extended time: “Every day I see the same group, so every day looks the same...Last year...it was the same grade level, but not the same students.”

This increased stability was closely tied to a MTSS-aligned, data-driven approach to assigning students to tutoring. As one tutor said:

last year there was not as much control over the sticking to the data driven information...teachers were asking, ‘Oh, I need to include this student—do you have room?’...and so it wasn’t supervised as well previously. But this year...I think that

they're being very data-driven about it and really trying to pinpoint the students that need the specific instruction.

This tutor went on to describe this shift as transformative, noting that smaller student caseloads were a key factor in improving tutoring quality: “it's actual high dosage and it's, it's more we're just doing a better job this year. Last year, I just had so many [kids]. This year, we can really focus on these groups of kids that need that [support].” As these reflections suggest, district initial efforts to clarify HDT activities improved understanding, but implementation did not align with those expectations until caseloads were low enough—and assignment and monitoring processes structured enough—to make consistent, high-dosage work feasible.

Persistent challenges with low-achieving and middle schools. Despite progress in school-level fidelity with HDT, schools with large numbers of below-grade-level students continued to struggle to adhere to HDT requirements. One elementary case school assigned six students to tutoring groups (instead of four) and asked the tutor to support each half of the group for 15 minutes (instead of the required 30 minutes) to make tutoring support more accessible to students. The school also used the tutor to support core or Tier 1 interventions, such as small-group instruction, delivered to all students. These adjustments meant that the school did not meet the district's HDT dosage targets tracked in the data dashboard.

Similarly, a tutor supervisor at another high-poverty elementary school noted that HDT was not feasible in a school where every student needed targeted one-on-one support. In these situations, it was better to use the tutors to support core instruction “before trying to pull 50 groups of students a day.” Similarly, a district leader argued that the standard treatment protocol cannot work in “reverse triangle” schools “where 80% or more of students are not [at grade-level]”. They went on to explain “it doesn't make sense to try to pull more than 20% of students

and put them into intervention groups. That's when we say, okay, yes, you're going to get an intervention plan, but it's going to be implemented through core [Tier 1 instruction].” Together, these reflections underscore that in schools where large proportions of students require intensive support, the scale of need makes true HDT operationally impractical because schools cannot realistically meet the required dosage for so many students.

Middle schools also faced persistent challenges that worked against the consistency of HDT achieved in elementary schools. Time was the biggest barrier. Providing 30-minute tutoring sessions, three times a week, meant students missed core instruction. Pull-outs disrupted teachers’ instructional plans, and teacher and student absenteeism complicated scheduling. In older grades, curriculum units required more time than the allotted blocks allowed, meaning that teachers were already struggling to cover grade-level content before sacrificing time to tutoring. Tutor supervisors tried to minimize disruption by pulling students from non-tested subjects or electives, but this approach had trade-offs—elective teachers did not appreciate losing instructional time, and students viewed tutoring as an added burden for not being on grade level. One supervisor summarized the challenge: “there is not much we can do...we just don’t have dedicated time for Tier 2 to happen...three 30-minute sessions has really been [hard]...when you have 200 plus students that would really need this.” Even as the pressure from the ESSER-period faded, these schools were under additional pressures that continued to frustrate implementation.

7. Discussion

Delivering tutoring to large numbers of students is inherently complex and introduces serious implementation challenges, particularly around hiring, placement, supervision, and fidelity to high-dosage practices. In GCS, these pressures were acute during the pandemic and ESSER-funded HDT period. At the same time, GCS’s experience with tutoring over time suggests that districts can respond to implementation barriers and build systems to support the

delivery of tutoring by adapting to realities on the ground and right-sizing programmatic demands and organizational capacity. By the time ESSER-funding had ended, GCS knew it had to invest in a more robust personnel infrastructure, building from its central tutoring department and streamlined HR processes and rapid mass hiring to slow down and spend more time on selective retention and strategic placement of tutors in schools. Downsizing the program to a more manageable (but still relatively large) scale in 2024–25 was pivotal: it redirected attention from quantity to implementation processes associated with quality, and created the bandwidth for more comprehensive tutor supervision and training.

Beyond right sizing the program, the district’s tighter integration of tutoring with MTSS and its investment in real-time data systems helped ensure depth and consistency of tutoring delivery. Clarifying how tutoring fits within MTSS (e.g., Tier 2/3 eligibility and standard treatment protocols) provided an anchor for schools to standardize tutoring practices—group size, data-driven assignment, dosage, duration expectations—and to leverage evidence-based curriculum resources for training tutors and guiding content choices for tutoring sessions, again especially so in early literacy. In parallel, the district established systems of pressure and support through ongoing guidance and education of schools on HDT practices, and data dashboards that enabled district leaders and schools to monitor tutoring dosage and respond when tutoring activity diverged from expectations.

By the 2024-25 school year, GCS had changed the systems supporting school-level implementation of HDT. Meanwhile, schools appeared to demonstrate greater ownership of tutoring activities and shared norms around the value of HDT. They also appeared to adopt HDT practices with greater depth and fidelity. Schools also spent more time focused on retaining tutors and integrating tutoring into core instructional processes, signaling sustainability of

tutoring in schools. These shifts were again most pronounced in tutoring for early-grade literacy in elementary schools.

The case of GCS points to several organizational conditions that may help district efforts to build systems supports for tutoring. First, leaders used downsizing as an opportunity to refine systems and programs rather than retreat from them. During the initial scaling-up phase of tutoring during ESSER, the district spent lots of energy on getting the basics in place. But once the district downsized post-ESSER, it had more capacity and knowledge to provide system-level supports for tutor hiring, training, supervision, and fidelity to core HDT activities. Second, district and school leader experience over time, and data on tutor quality and placement inequities, helped inform subsequent actions to sustain hiring pathways (e.g., creating school-based part-time roles) and rebalance staffing toward higher-need schools and subjects. Third, cross-department collaboration (tutoring, HR, finance, data, curriculum, MTSS), and responsive communication between the tutoring department and schools, supported continuous problem-solving and systems refinement.

The case of GCS also highlights how a district's progress can be shaped by structural factors that are both within and outside of its direct control. On the former, the district's intentional shift in tutor funding toward school improvement dollars prompted greater policy coherence between tutoring and MTSS, enabling leaders to draw on existing MTSS-aligned protocols to standardize group size, session frequency and duration, data-driven assignment, and the use of curriculum materials. Likewise, the decision to partner with an external vendor to build its data infrastructure supported visibility and accountability for tutoring dosage.

Some factors beyond the district's control also supported HDT implementation. For example, state legislation mandating phonics-based instruction, combined with a rich ecosystem

of early literacy interventions acquired by the district, provided helpful content for tutor training and session guidance in early literacy. At the same time, there were many external factors that constrained implementation, including rapid policy deadlines and pressures under ESSER to scale tutoring. The additional pressures and resource problems faced by middle, remote, and hard-to-staff schools also created persistent implementation barriers. In these schools, the district struggled to achieve consistent HDT activities around data use and curriculum planning, and fidelity to HDT practices such as small group size, dosage, frequency, and instructional delivery. Such challenges suggest there may be natural limits to how far districts can leverage system-level improvements to support HDT. Some structural barriers, like staffing shortages, rigid schedules, and the sheer volume of students needing intervention, may go beyond what districts can resolve on their own during the time frame of this study.

7.1 Contributions to Literature, Future Research, and Study Limitations

Our study advances the literature on post-COVID tutoring programs by going beyond cataloging implementation barriers to show how districts respond to barriers and build systems to try and address them. We also identify the evolving organizational and structural conditions that shape district action. We show how ESSER functioned as a catalyst for action and learning, creating a window for districts to engage in rapid scaling and subsequent refinement.

Future research should examine whether systems that grew out of the ESSER-funding period persist or erode without targeted funding and policy or political support. Comparative studies could identify which organizational conditions most reliably produce fidelity to HDT activities, especially in secondary math or in schools where resource gaps and scheduling constraints are acute. More research is needed to understand the effects of downsizing on instructional quality and tutor retention, the equity implications of tutor placement and staffing

decisions, and the role of policy coherence between tutoring and MTSS in sustaining tutoring across diverse school contexts.

Several limitations should be considered when interpreting these findings. First, this single-district case study limits generalizability to other contexts. GCS is a large district with an unusually extensive home-grown tutoring program, even after post-ESSER downsizing, and districts with different governance structures, staffing capacity, or policy environments may face distinct implementation dynamics. Second, although the study draws on rich, longitudinal qualitative data and triangulates across interviews, focus groups, documents, and observations, findings rely primarily on participant self-reports; we were not able to directly observe tutoring instruction at scale or systematically assess instructional quality across sites. Third, our purposive and relatively small school sample enabled in-depth analysis but may not capture the full range of implementation experiences across GCS, particularly schools implementing HDT in more minimal ways. Finally, this paper focuses on implementation processes rather than causal impacts on student achievement. While dosage and participation trends suggest improvements in fidelity over time—especially in early literacy—examining effects on student outcomes is beyond the scope of this study and remains an important direction for future research.

7.2 *Implications for Policy and Practice*

GCS’ experience suggests that policymakers (and the public) should recognize that districts need time to iterate and scale instructional interventions, especially when building home-grown programs. Pushing immediate large-scale implementation can crowd out reflection, supervision, and fidelity—all key components that emerged only *after* GCS had concluded with ESSER funding and district leaders had time and space to refine tutoring. The recurring problems in middle schools and elsewhere are a good reminder that districts may not be able to

build systems for scale that work everywhere by themselves. Federal and state guidance, and access to external vendors and partners, could help equip districts with intervention frameworks (e.g., MTSS), evidence-based curriculum resources, and data tools to standardize tutoring, strengthen tutor training, and ensure the quality of instructional lessons. GCS's progress in early literacy was aided by such structures; parallel guidance for math or other challenges at the middle school level might help districts close persistent gaps.

Moreover, our results suggest that looking only at short-term student test score changes risks missing district-level changes that unfolded during the ESSER period. While immediate outcome gains are a reasonable focus given the magnitude of ESSER funding, focusing on test score impacts alone may obscure how districts used ESSER funding to initiate or adapt systems with longer time horizons. We find that GCS directed ESSER resources toward building infrastructure—such as tutor pipelines, policy alignment, and data systems—that may continue to shape educator practice and students' academic performance.

For district and school leaders, our findings suggest the possibility of reframing implementation barriers as opportunities to refine systems to support implementation, from personnel infrastructure to policy coherence to pressure-and-support mechanisms and cross-department collaboration. A systems perspective can help districts work towards continuous improvement of programs. But doing so requires strong communication and collaboration between districts and schools, within schools, and for both districts and schools to have time to learn and improve.

Finally, our findings highlight an underlying tradeoff between flexibility and local adaptation, on the one hand, and implementation fidelity and district-wide coherence, on the other. While the scale and heterogeneity of student need can make highly prescriptive tutoring

models difficult to implement, districts can, over time, work toward greater fidelity and coherence by clarifying expectations for HDT practice, educating schools about core design elements, and building responsive relationships with school leaders. In GCS, we observed an initial pattern of leaning heavily toward flexibility, which allowed schools to adapt tutoring to local contexts but also limited the development of shared structures and evidence-aligned HDT practices. District-level efforts such as managing tutor caseloads, providing MTSS-aligned protocols and resources, tracking tutoring delivery in real time, and protecting intervention time during the school day represent ways districts might rebalance toward greater fidelity and coherence without fully sacrificing local responsiveness. We find that this balance was more readily achievable in early literacy, while middle school math and schools serving large numbers of students with intensive needs required additional resources and scheduling solutions.

References

- Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., & Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, 31(4), 73–102.
- Bhatt, M. P., Chau, T., Condliffe, B., Davis, R., Grossman, J., Guryan, J., Ludwig, J., Magnaricotte, M., Mattera, S., Momeni, F., Oreopoulos, P., & Stoddard, G. (2025). *Personalized Learning Initiative Interim Report: Findings from 2023-24*. The University of Chicago: Education Lab and MDRC.
<https://www.google.com/url?q=https://www.mdrc.org/sites/default/files/UChicago-Education-Lab-PLI-Interim-Report-06.2025.pdf&sa=D&source=docs&ust=1768411132418419&usg=AOvVaw0nwrY2ZsVmo2CJJ-sjHYVZ>
- Boughton, H., de Barros, J., Goldhaber, D., Payne, S., & Schwartz, N. (2021). *The once-in-a-generation opportunity: What states and districts can do now to learn from American Rescue Plan ESSER interventions* (CALDER Policy Brief No. 27-092). National Center for Analysis of Longitudinal Data in Education Research (CALDER).
- Carbonari, M. V., Davison, M., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Hashim, A. K., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2024). The Challenges of implementing academic interventions during COVID: Evidence from the Road to Recovery project. *AERA Open*, 10.
<https://doi.org/10.1177/23328584241281286>
- Carbonari, M. V., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Kane, T. J., McDonald, A., McEachin, A., Morton, E., & Muroga, A. (2024). *Impacts of Academic Recovery Interventions on Student Achievement in 2022-23*.
- Cobb, P. A., Jackson, K., Henrick, E., Smtih, T. M., & the MIST Team (Eds.). (2018). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Harvard Education Press.
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3–12.
- Cortes, K., Kortecamp, K., Loeb, S., & Robinson, C. (2024). *A Scalable Approach to High-Impact Tutoring for Young Readers: Results of a Randomized Controlled Trial*. National Bureau of Economic Research.
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
- DiMarco, B. (2025, May). Legislative Tracker: 2025 State Tutoring Bills. *FutureEd*.
<https://www.future-ed.org/legislative-tracker-2025-state-tutoring-bills/>

- Dusseault, B., & Pillow, T. (2021). *First look at ESSER priorities: Districts are placing their bets on what they know*. Center on Reinventing Public Education.
- Fryer, R. G., & Howard-Noveck, M. (2020). High-dosage tutoring and reading achievement: Evidence from New York City. *Journal of Labor Economics*, 38(2), 421–452.
- Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L., & Schatschneider, C. (2008). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology*, 100(3), 491.
- Fullan, M. (2000). The return of large-scale reform. *Journal of Educational Change*, 1(1), 5–27.
- Gallagher, H. A., Faw, L., & Cottingham, B. (2026). *How districts scale instructional improvement that lasts*. Policy Analysis for California Education at Stanford University. https://edpolicyinca.org/sites/default/files/2026-02/pb_gallagher_feb_2026.pdf
- Giddens, A. (1979). *Central Problems in Social Theory: Action, Structure, and Contradiction in Social Analysis*. University of California Press.
- Hallgren, K., Gonzalez, N., Choi, J., Kelly, K., Li, A., Ochoa, L., & Gill, B. (2017). *The Atlanta public schools turnaround strategy after one year: High impact tutoring and the purpose built schools partnership*. Mathematica Policy Research. <https://www.atlantapublicschools.us/cms/lib/GA01000924/Centricity/Domain/11375/Year%201-APS%20Turnaround%20Report%202017.pdf>
- Hashim, A. K., Davison, M. B., Morton, E., Leak, J., Wright, C., Dizon-Ross, E., Stephens, S., & Hamilton, K. (2024). Partnering through pandemic recovery: An examination of a research-practice partnership in Guilford County Schools. *Journal of Educational Administration*.
- Heinrich, C. J., Burch, P., Good, A., Acosta, R., Cheng, H., Dillender, M., Kirshbaum, C., Nisar, H., & Stewart, M. (2014). Improving the implementation and effectiveness of out-of-school-time tutoring. *Journal of Policy Analysis and Management*, 33(2), 471–494.
- Huffaker, E., Robinson, C. D., Bardelli, E., White, S., & Loeb, S. (2025). *When interventions don't move the needle: Insights from null results in education research* [EdWorkingPaper: 25-1295]. Annenberg Institute at Brown University. <https://doi.org/10.26300/58dd-6d02>
- Jacobson, L. (2023, January 24). Cardona's tutoring charge, 1 year later: Some progress, but obstacles remain. *The 74 Million*. <https://www.the74million.org/article/cardonas-tutoring-charge-1-year-later-some-progress-but-obstacles-remain/>
- Kraft, M. A. (2015). How to make additional time matter: Integrating individualized tutorials into an extended day. *Education Finance and Policy*, 10(1), 81–116.

- Kraft, M. A., Edwards, D. S., & Cannata, M. (2024). *The scaling dynamics and causal effects of a district-operated tutoring program* (EdWorking Paper 24-1030). Annenberg Institute at Brown University. <https://edworkingpapers.com/ai24-1030>
- Kraft, M. A., & Falken, G. T. (2021). A blueprint for scaling tutoring and mentoring across public schools. *AERA Open*, 7. <https://doi.org/23328584211042858>
- Kraft, M. A., List, J. A., Livingston, J. A., & Sadoff, S. (2022). *Online tutoring by college volunteers: Experimental evidence from a pilot program*. 112, 614–618.
- Kraft, M. A., & Lovison, V. S. (2024). The Effect of Student-Tutor Ratios: Experimental Evidence from a Pilot Online Math Tutoring Program. EdWorkingPaper No. 24-976. *Annenberg Institute for School Reform at Brown University*.
- Kraft, M. A., Schueler, B. E., & Falken, G. T. (2024). *What impacts should we expect from tutoring at scale? Exploring meta-analytic generalizability* (EdWorking Paper 24-1031). <https://edworkingpapers.com/ai24-1031>
- Makori, A., Burch, P., & Loeb, S. (2024). *Scaling high-impact tutoring: School level perspectives on implementation challenges and strategies* (EdWorking Paper No. 24-923). Annenberg Institute at Brown University. <https://doi.org/10.26300/h8z5-t461>
- March, J. G., & Olsen, J. P. (1975). The uncertainty of the past: Organizational learning under ambiguity. *European Journal of Political Research*, 3, 147–171.
- McDonnell, L. M., & Elmore, R. F. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis*, 9(2), 133–152.
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9, 171–178.
- Morton, E., & Hashim, A. K. (2023). *Accelerating student academic recovery* [Brief No. 26]. EdResearch for Recovery, Annenberg University. <https://www.edresearchforaction.org/research-briefs/accelerating-student-academic-recovery/>
- Nickow, A. J., Oreopoulos, P., & Quan, V. (2020). *The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence*. (No. 27476). NBER. https://www.nber.org/system/files/working_papers/w27476/w27476.pdf
- Robinson, C. D., Bisht, B., & Loeb, S. (2022). *The inequity of opt-in educational resources and an intervention to increase equitable access* (EdWorking Paper No. 22-654). Annenberg Institute at Brown University. <https://doi.org/10.26300/ja2n-ys82>
- Robinson, C. D., Kraft, M. A., Loeb, S., & Schueler, B. E. (2021). *Accelerating student learning with high-dosage tutoring* (EdResearch for Recovery Design Principles). EdResearch for Recovery Project. <https://eric.ed.gov/?id=ED613847>

- Robinson, C. D., Pollard, C., Novicoff, S., White, S., & Loeb, S. (2024). The Effects of Virtual Tutoring on Young Readers: Results from a Randomized Controlled Trial. *EdWorkingPapers. Com*.
- Schueler, B. E., & Rodriguez-Segura, D. (2021). *A cautionary tale of tutoring hard-to-reach students in Kenya* (EdWorking Paper No. 21-432). Annenberg Institute at Brown University. <https://doi.org/10.26300/43qs-cg37>
- Stanford, L. (2022, July 8). Biden's tutoring initiative: What will it mean for learning recovery? *EdWeek*. <https://www.edweek.org/policy-politics/bidens-tutoring-initiative-what-will-it-mean-for-learning-recovery/2022/07>
- White, S., Groom-Thomas, L., & Loeb, S. (2023). *A systematic review of research on tutoring implementation: Considerations when undertaking complex instructional supports for students* (EdWorkingPaper No. 22-652). Annenberg Institute at Brown University. <https://doi.org/10.26300/wztf-wj14>
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Sage Publications, Inc.