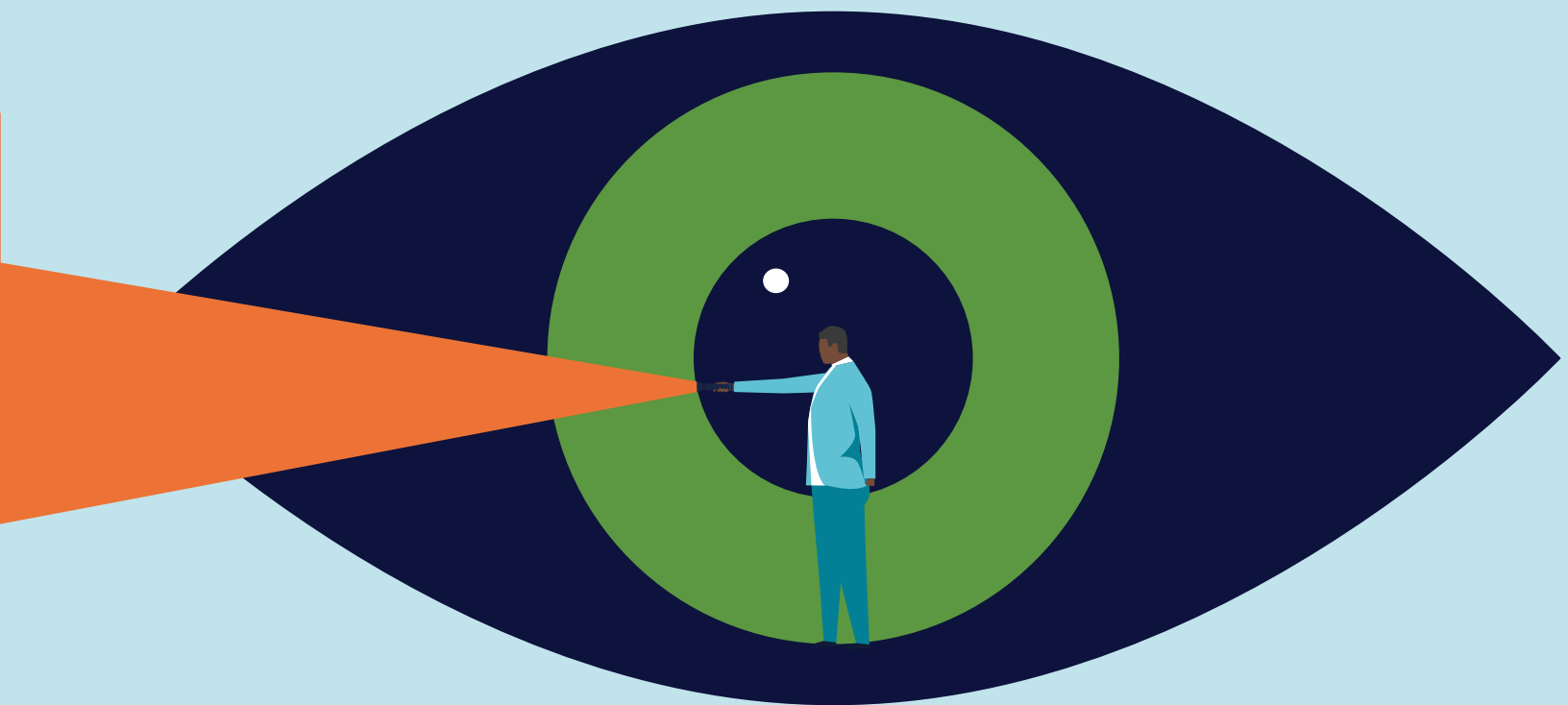


The State of Tutoring 2026:

What Leaders Can See and What They Can't

By Evo Popoff, Daimen Sagastume, & Matt Tower



Produced by

Commissioned by

W/A Whiteboard
Advisors

PEARL
EDUCATION

Foreword

By Amanda Neitzel (Inns), PhD

Associate Research Professor

Deputy Director of Evidence Research, Center for Research and Reform in Education

I recently was running validation checks on session data for a state tutoring program, and the system kicked back a tutor whose sessions overlapped. One session ended at 2:00 p.m. The next one had already started at 1:58 p.m.

I wanted to know what happened in those two minutes and in the other overlapping entries I was seeing. With one vendor, it turned out to be nothing. Its system was tracking to the second and rounded up, creating a less-than-one-minute gap. I could live with that.



But some of these overlaps were five to eight minutes, and the sessions were 20 minutes long. Those gaps couldn't be explained away as rounding errors, and nobody could explain what was happening in those five to eight minutes. Was the tutor running both groups at once? Was one group just sitting there? No one had an answer — not because anyone was hiding anything; it just hadn't occurred to anyone to look.

Those missing minutes matter more than most people realize. **One of our recent large-scale randomized controlled trials found that completing full-length sessions was even more predictive of student outcomes than attendance alone.** This makes intuitive sense, but it puts those gaps in sharper focus. Two minutes lost per session doesn't sound like much. But across a school year of 20-minute sessions three times a week, it adds up to over 200 missed minutes — nearly four weeks of tutoring lost. Every minute counts. Knowing whether students are receiving those minutes, in real time, counts too.

I run third-party evaluations for state and district tutoring programs — work I've been doing since before the pandemic and that our center now conducts at scale across the country. **The pattern I see most often isn't that programs are failing. It's that the people responsible can't see whether the programs are failing or not.**

This is obviously a problem for the students in these programs. Consider the unseen student who stops turning up to tutoring without any adult asking why, or the students who spend eight minutes of a 20-minute session waiting for an instructor to appear.

But it also creates challenges for states and districts that have invested significant time and resources into their tutoring programs with little to no visibility into what's actually happening. Sometimes program managers operate from after-the-fact summary reports that capture program averages. **These reports often conceal more than they reveal — like**

the 3 in 10 students who aren't receiving high-impact tutoring buried in an apparently successful program based on average data — and provide no opportunity to improve programs in real time.

A state we work in requires a minimum threshold of high-impact tutoring for students scoring below proficient on their state test. But like many states, the state wasn't yet collecting real-time, systematic data on what tutoring students were receiving. Some of the vendors providing reading support do not appear in [Evidence for ESSA](#) or [ProvenTutoring](#), the clearinghouses my center maintains, which means some share of the students covered by that mandate were getting something else beyond an evidence-proven tutoring model. The issue isn't unique to this state; it's common. Mandates get written, but without mechanisms to track implementation, requirements on the books can go unmet year after year — not because anyone intends to fall short but because **no one has visibility into what's actually happening. You can't fix what you can't see.**

Fortunately, there are examples of programs where leaders can not only see what's happening but use those insights to improve their programs. In those programs, the data comes in on a regular basis; it's clean, nobody's hiding anything; and program managers receive actionable insights in real time.

In Massachusetts, state leaders receive uploaded data on their early literacy tutoring program on a close-to-biweekly cadence. **The difference between seeing 65% attendance in February versus July is the difference between having an opportunity to improve implementation and simply reporting what happened after the fact.**

These insights can also sharpen the research itself. The evidence base for high-impact tutoring is strong, but it rests largely on averages, and averages can obscure what any individual child actually needs. Consider the standard benchmark of three sessions per week. Should that apply equally to a first grader who struggles to sit still and a high schooler accustomed to sustained instruction? We do not know yet, and we should. **Broadening the evidence base is how the field learns what works best for first graders specifically or for multilingual learners or for students in rural districts.** And when that data is available in real time, educators do not have to wait for the next study to act on it. They can adapt to what a student needs now.

Most of what the field knows about tutoring reflects what programs were designed to deliver: the dosage a contract specifies, the schedule a plan lays out, the model a vendor describes. What follows is different. Across 883,882 sessions and 64,044 students, this report describes what students actually received. It is a snapshot rather than a complete picture, but it is far more than the field could say a year ago, and it is a floor to build on.

In the meantime, the question I'd ask of any tutoring program isn't whether it meets the standard. It's whether anyone knows what is actually being delivered to students. Somebody should know what's happening in those tutoring sessions — before missing minutes become missing weeks.

Executive Summary

This is the second annual report on the state of K-12 tutoring. This year, the report examines what tutoring actually looks like and how close programs come to achieving research-backed ideals relating to session length and frequency as well as tutoring duration and group size. These insights are based on an analysis of 883,882 tutoring sessions delivered to 64,044 students across 28 programs in 2025-2026. We've also talked to state and district leaders to get their perspectives. This report likewise attempts to address one of the bigger challenges facing the field: Today, most tutoring data is reported as averages.

But a lot can get lost in an average: Three-quarters of all students receiving the recommended weekly dose sit inside programs whose averages fall short, while 29% of students in the programs that hit the mark are not actually getting there. Looking student by student, our analysis of that data sought to answer three questions:

Are students receiving enough tutoring each week?



- **Close but not quite.** The research-backed ideal is about 90 minutes a week, but the typical student received 72.5 minutes, and just over one-quarter reached 90 or more. The gap for the typical student is under 20 minutes a week.
- **Two students in the same program can have very different experiences.** In one program serving just over 6,000 students, some students received 23 minutes of tutoring a week, while others received 96. Most of the variation in this data happens inside programs rather than between them.

Are students participating in enough weeks of tutoring and enough tutoring sessions per week?



- **Tutoring often stops before the school year does.** The typical student was tutored for 12 weeks. Nearly 1 in 4 received fewer than eight weeks of tutoring. Just over half, 53%, reached 12 weeks or more. But that doesn't necessarily mean those students received the ideal number of sessions per week. Not all of these shortfalls reflect a program breaking down: students transfer, move out of district, graduate, or enroll mid-cycle, and some programs are designed as short-term interventions from the start.
- **Tutoring frequency varies.** Duration (the number of weeks students participate in a tutoring program) and frequency (the number of sessions students have per week) are rarely linked. So, fixing one does not fix the other. Only 10% of students received three or more sessions a week sustained across 12 weeks or more, and more than half of frequently tutored students were enrolled for fewer than 12 weeks.
- **When tutoring is available, students show up.** Students attended 83% of sessions, though attendance rates range from 79% in state-led programs to 86% in vendor-led programs. Students want help, but they may not have access to the right amount of tutoring for as long as they need.

How does the student experience differ across programs?



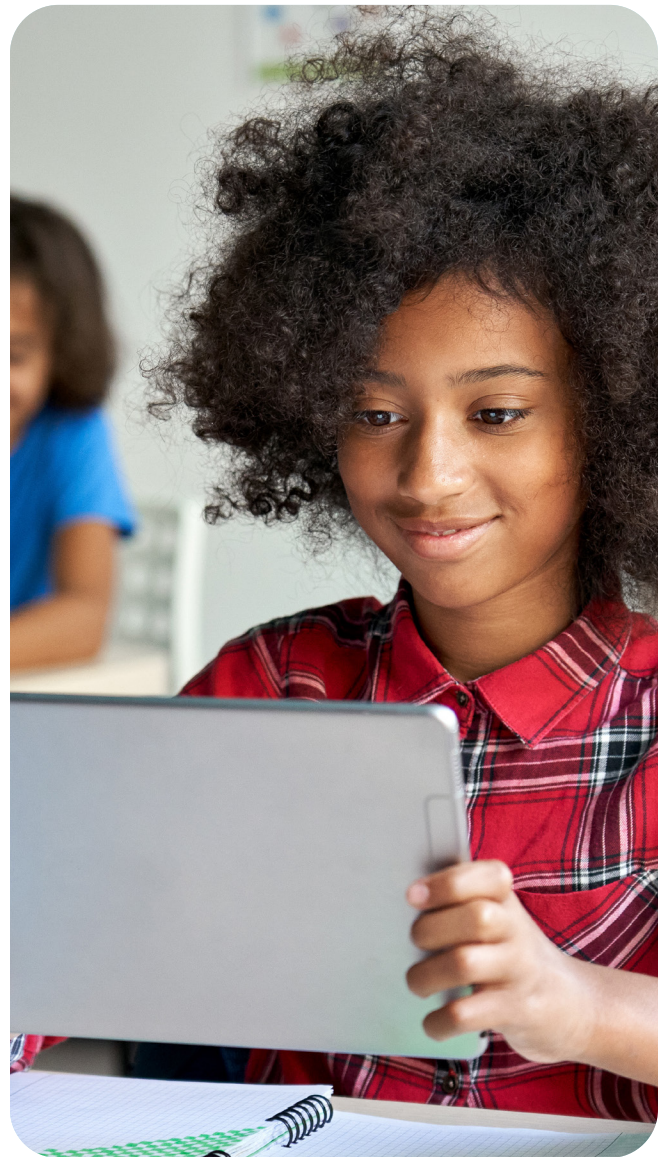
- **"Tutoring" program offerings differ.** Session length across programs runs from 20 to 105 minutes, and average group size ranges from one student to 15. Two programs, for example, deliver about one hour a week: one in 20-minute sessions in groups of three online and the other in half-hour sessions in groups of six in person.
- **Who runs the program shapes what a student receives.** In district-led programs, 84% of students kept the same tutor all year compared with roughly 4 in 10 in state-led and vendor-led programs. Vendors offer the most frequent sessions; state programs last the longest. Yet, all three land within about six minutes of each other on weekly tutoring time.

Taken together, these findings offer an encouraging foundation for the first year of post-ESSER implementation and point to opportunities to build on that progress. Students are participating in tutoring, most are receiving meaningful amounts of support, and the typical student is less than 20 minutes per week from the research-backed weekly benchmark. The challenge is not that tutoring isn't happening. It is that delivery varies considerably from student to student — and until recently, leaders had limited visibility into those differences.

That creates an opportunity. **Frequency, dosage and persistence are implementation conditions leaders can see and adjust during the school year. Better visibility makes it possible to identify students who are receiving less support than intended and intervene before those gaps compound.**

Until now, the field could describe what it spent on tutoring and how many students it reached. This year, it can describe what students actually received. While this is a notable shift, it's still only half the picture. Implementation data shows that students received very different amounts of tutoring, for different lengths of time, in very different arrangements, but it cannot yet say which of those differences changed anything for a student. Connecting the data with outcomes is what turns a record of activity into evidence about what works and for whom.

This is one of the largest collections of tutoring implementation data assembled to date, and it is still a partial view. It covers programs that use a single platform, in a specific set of states and districts, and it is explicitly not a nationally representative sample. This analysis also draws on data from Pearl, which works with states, school districts, and tutoring providers to design, manage, and sustain student support programs at scale, including tutoring, intervention, and supplemental instruction. Pearl commissioned this report — readers should weigh that accordingly.



Contents

Foreword	2
Executive Summary	4
Contents	7
About the Authors	8
Acknowledgments	9
Introduction	10
1. The Persistence of High-Dosage Tutoring: The Field's Strongest Evidence Base	13
2. The Visibility Gap: What the Data Shows	16
Finding 1: Are Students Receiving Enough Tutoring Each Week?	16
The Average is Not the Student	17
Finding 2: Is That Support Sustained Long Enough?	20
Finding 3: How Does the Student Experience Differ Across Program Models and Modalities?	21
3. From Visibility Toward Accountability: How Tutoring Systems Are Evolving	25
State Spotlight: Arkansas	27
Funding: Structure Determines Staying Power	28
Data Collection: Seeing Early Enough to Act	29
State Spotlight: Massachusetts	30
District Spotlight: Metro Nashville Public Schools	31
4. Why This Matters Now: Tutoring is Becoming Permanent	33
Conclusion: A Baseline Year	35

About the Authors

Evo Popoff is a senior vice president at Whiteboard Advisors. Named State Policy Maker of the Year by the State Education Technology Directors Association, he previously served as chief innovation and intervention officer and assistant commissioner for the New Jersey Department of Education, where he oversaw the state's edtech and school system improvement efforts. Prior to joining the department, he led the development of edtech products and school improvement solutions in collaboration with school system and state leaders and educators. Before beginning his career in education, Evo practiced law at McDermott, Will & Emery, where he worked on labor and employment, antitrust and general corporate issues. He holds a Bachelor of Arts in political science from the University of Chicago and a Juris Doctor from The George Washington University Law School.

Daimen Sagastume is a senior director at Whiteboard Advisors (W/A), where he specializes in advocacy and growth enablement across K-12 education. Prior to joining W/A, he spent five years at Emerson Collective, managing its education philanthropy investment portfolio with a focus on scaling innovative solutions for educational equity. During his tenure, he played a pivotal role in incubating Uppercase, a seed-stage edtech venture dedicated to democratizing access to world-class teaching expertise. A Stanford University graduate with a Bachelor of Science in biology, Daimen discovered his passion for education while pursuing premedical studies. At the intersection of edtech, philanthropy and strategic advisory, he works to ensure that innovative edtech solutions reach the students and educators who need them most.

Matt Tower spends his time studying the intersection of business and education. He is the editor of The EdSheet, a Whiteboard Media newsletter covering the same topic, and supports Whiteboard's Strategy and Research practices. Matt joined W/A via W/A's acquisition of ETCH, a company he founded to provide news and information to education industry investors and company leaders. Prior to working on ETCH, Matt spent 10 years in the education industry as a product manager and venture capitalist. Matt received a Bachelor of Arts in history from Amherst College, with a concentration on the geopolitics of natural resources. While at Amherst, Matt owned and operated a consignment bookstore, painted houses, and taught skiing and snowboarding. He now lives in San Francisco with his wife and children. If not reading at his desk, he can often be found running on a soccer field or floating on a surfboard.

Acknowledgments

We collaborated with administrators and field leaders to develop this report, incorporating their qualitative insights that influenced our key findings. We would like to express our gratitude to these forward-thinking Pearl partners:

- **Amy Counts**, *Director of Curriculum Projects*, Arkansas Department of Education
- **Kathy Moffitt**, *Education Administrator*, Alaska Department of Education & Early Development
- **Amanda Neitzel**, *Associate Research Professor*, Center for Research and Reform in Education
- **Dianne Orr**, *Education Administrator*, Alaska Department of Education & Early Development
- **Catherine Peretti**, *Executive Director*, CitySchools Collaborative
- **Justin Robicheaux**, *Director of Curriculum & Instruction*, East Baton Rouge Parish Public Schools
- **Linda Sewnarine**, *Assistant Director of Literacy*, Massachusetts Department of Elementary and Secondary Education
- **Rachel Tompkins**, *Executive Director of the Scholars Portfolio*, Metro Nashville Public Schools
- **Craig Waterman**, *Assistant Director of Strategy and Integration*, Massachusetts Department of Elementary and Secondary Education



Introduction

Over the past two decades, tutoring has become a core strategy for how districts approach academic recovery and acceleration. During and immediately after the COVID-19 pandemic, federal Elementary and Secondary School Emergency Relief (ESSER) funds covered the costs of tutoring programs. But even as ESSER relief funding has dissipated, a growing number of states and districts are still treating tutoring, specifically high-dosage tutoring, as a permanent part of their academic strategy.

As spotlighted in last year's report, high-dosage tutoring represents the strongest existing evidence base for what effective tutoring looks like: roughly three sessions a week, in small groups, sustained over time. It is not the only configuration that can work, and the evidence for others is considerably less developed. But it is the clearest reference point the field has. The question this report asks is whether leaders can actually see how close individual students come to the standard high-dosage tutoring sets in practice.

To answer this question, this report draws on tutoring data covering 883,882 sessions from 64,044 students that the Pearl Student Support Platform collected from 28 programs, across three program models: 16 vendor-led, six state-led and six district-led.

This dataset represents one of the largest cross-state collections of tutoring implementation data assembled to date, allowing a deep exploration of what tutoring actually looks like in practice, measured student by student rather than program by program.

When leaders judge a tutoring program, they usually look at its averages: sessions per student, hours delivered, students served. But an average describes a program — not a student — and in this data, the two often point in opposite directions.



The average is not the student

Judging a tutoring program by its average misses students in both directions.

Where are the students getting the recommended weekly dose?

17,176

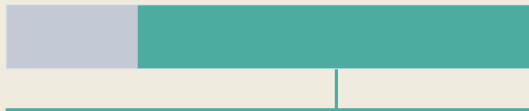
students received 90+ minutes of tutoring per week

4,273 (25%)

in programs whose average meets the standard

12,903 (75%)

in programs whose average falls below it



Three quarters of the students receiving the recommended weekly dose are inside programs a leader would likely write off.

What's actually inside the programs that look like success stories?

5,990

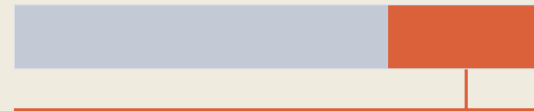
students are enrolled in programs whose average meets the standard

4,273 (71%)

are getting 90+ minutes a week

1,717 (29%)

are not



Nearly three in ten students inside the highest-performing programs aren't getting what the average suggests.

The session-by-session analysis reflected in this report provides a unique look at the state of tutoring that typical analyses miss, which, as Amanda Neitzel, associate research professor at the Center for Research and Reform in Education noted, are often based on program averages. **Those program averages hide a lot of information**, particularly from the leaders who fund and manage these programs. When data from individual tutoring sessions are compressed into a single number — average number of sessions per student, for instance — important details about the services students are or aren't receiving are buried. An analysis of this report's dataset illustrates the problem with program averages. Based on program averages, a typical student in **19 of the 28 programs in this dataset receives fewer than 90 minutes of tutoring a week, yet those programs contain 12,903 students who are receiving 90 minutes or more tutoring, three-quarters of all such students in the dataset.** In the nine programs whose averages do meet the mark, 29% of students are not actually receiving tutoring that meets evidence-backed standards. This disconnect between program averages and student experience creates a **visibility gap**, which prevents effective program implementation, continuous improvement and overall program impact. The only way to bridge this visibility gap is by looking at individual session and student data.

Research-backed evidence tells us whether tutoring works and under which conditions it works best. Implementation data tells us something different: whether students actually received tutoring in the ways that research suggests matter most. This year's dataset gives us an unusually strong view into that implementation layer, and it reveals meaningful differences in how much tutoring students receive each week, how long that support is sustained, and how it is delivered across program models and modalities.

Seeing those differences clearly and early enough to act on them is critical for successful program implementation and continuous improvement efforts. But implementation data alone doesn't tell a story about outcomes or which programs are delivering results for students under specific conditions. While these are important issues, they are not ones that this year's report set out to address, because doing so would require additional data that this year's dataset was not designed to capture.

Program visibility is the important first step; it's what makes current practice legible and improvement possible. The interviews and analyses presented in this report show where there are visibility gaps, why these gaps are important to close, how tutoring can become an even more sustainable practice, and why the next phase in tutoring's evolution will center on accountability.

A note on methodology: All figures were computed from Pearl's raw data files. The dataset was completely anonymized to remove any personally identifiable information about students. The analytic sample is 64,044 students across 28 programs and 883,882 completed sessions. Medians are student-weighted. Dosage is measured across the weeks each student was actually enrolled, excluding school breaks. This data tracks implementation through session frequency, duration, group size, tutor consistency and attendance, but it does not track student outcomes or involve any individual, personal student data. While this is a large dataset, it is drawn from a specific set of states and districts and is explicitly not a nationally representative sample.

The Persistence of High-Dosage Tutoring: The Field's Strongest Evidence Base

High-dosage tutoring remains the most rigorously studied approach to K-12 tutoring. The National Student Support Accelerator (NSSA) summarizes the evidence as showing **tutoring can produce roughly three to 15 months of additional learning**, depending on program, grade and context, and Stanford currently describes high-dosage tutoring as producing an average of more than four months of additional learning in elementary literacy and almost 10 months of additional learning in high school math. But what exactly does “high-dosage” mean?

According to the research base, the definition of high-dosage tutoring has gotten fairly specific: no more than four students meeting with a single tutor for at least 30 minutes, three times a week, sustained over time. And achieving the standard does matter. A recent study found that **students were 7% less likely to be absent on days when tutoring was scheduled**, but that the effects were even stronger when tutoring incorporated key high-impact design features, including greater frequency, smaller tutor-to-student ratios, and in-school delivery. The study found no significant attendance effect when programs had none of these features or only one of the three features, while effects increased when programs combined two or three.

But, as researchers themselves have noted, faithful implementation of this standard is difficult, and meeting an implementation benchmark isn't the same as demonstrating effectiveness. While the research on high-dosage tutoring is thorough, there is more learning to be done about how these standards apply across contexts. According to Amanda Neitzel, an associate research professor at the Center for Research and Reform in Education, “Context is incredibly important,” and she notes some considerations:

Definition of High-Dosage Tutoring

no more than
four students
meeting with
a **single tutor**

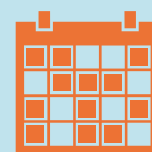


for at least

30 minutes



3x



**a week, sustained
over time**



On group size, there is flexibility. “We say high-impact tutoring is small groups of up to four kids, and people take that as: We can’t put five kids in a group. The research doesn’t say that four is amazing and five is garbage. It can work really effectively with up to six kids. It’s more about averages: an average of four kids per session. So maybe one kid needs their own session, and it evens out.” District leaders may read “four students” as a hard cap; “an average of four” is the more faithful reading of the research. Similarly, a group of four won’t be beneficial for a student who needs one-to-one attention.



On frequency, the returns are not linear. “Dosage isn’t linear. It’s not that every session gets you the same amount of learning. There’s a point below which it doesn’t help, and a point where you’ve maxed out and more sessions don’t do anything.”



On session length, developmental reality is relevant. The 45-minute daily session that anchors most high-dosage tutoring guidance was never meant to apply uniformly across grade bands. What works for a high school student is not developmentally appropriate, or even realistic, for a six-year-old. Understanding students is critical when it comes to program design. “Can you imagine first graders doing 45 minutes of tutoring a day? That’s not developmentally appropriate. We set these guidelines as the gold standard, but some of it is very context-dependent. On average, it holds, but there are places where adaptations would actually work.” In some cases, roughly 15-minute sessions are more effective than longer ones, especially for younger students.

These group-size, frequency and session-length ideals are evidence-backed conditions associated with effective tutoring, but they aren’t absolutes. A program shouldn’t be seen as a failure just because it doesn’t achieve every ideal. Craig Waterman, assistant director of strategy and integration at the Massachusetts Department of Elementary and Secondary Education, describes the value of holding these ideals loosely: “As much as I’m aware of the research, what we’ve found is unsurprising: more dosage, more learning; smaller groups, better; consistent multiyear better than single year ... The good news, and it’s what



we've found, is that tutoring most helps the lowest-performing kids and kids identified as at-risk."

These ideals provide a useful lens for understanding why better implementation visibility is needed and why connecting that visibility with outcomes is necessary to understand effectiveness. The goal of analyzing program data should be to gain visibility into how tutoring is actually being delivered and how closely it aligns with evidence-backed conditions, not to push conformity to unrealistic or inappropriate standards. When schools strive toward the high-dosage ideal, students benefit. But students are unique, and circumstances vary, so it's also important to understand where there is room for adaptation.

In practice, however, the high-dosage tutoring ideal can be difficult to achieve within the constraints of a typical school or classroom setting. Rachel Tompkins, executive director of the Scholars Portfolio, Metro Nashville Public Schools, describes the problem directly: **"We have a lot of research that talks about what high-impact tutoring is and how it should happen in theory. There has been less information on how you implement this, given the constraints in an actual school building."** Those constraints do not necessarily prevent schools from implementing effective tutoring, but they may require different approaches. Overlapping intervention blocks can shape staffing models; limited space may make virtual tutoring more practical than in-person tutoring; and transition



Other, non-high-dosage models of tutoring require further scrutiny and analysis. While they may still be able to deliver meaningful outcomes in certain circumstances, the evidence on their effectiveness is limited.

times may affect the ideal length and frequency of sessions. This is a gap in the existing research on tutoring. Districts would benefit from a range of evidence-based implementation options and clearer guidance about which approaches work best under which conditions.

Other, non-high-dosage models of tutoring require further scrutiny and analysis. While they may still be able to deliver meaningful outcomes in certain circumstances, the evidence on their effectiveness is limited.

The Visibility Gap: What the Data Shows

Waterman may have said it best: “If I were advising another state, [I’d say that in] your first year, there’s really only one question worth worrying about: Is tutoring happening the way we think it’s happening?”

Student-level analysis provides the only window into answering that question. That level of visibility can also provide an important first step to filling the research gap in the previous section by identifying what students are actually receiving. States and districts often use program averages to measure how well programs are being implemented, which don’t tell the full story of how, or how much, tutoring is actually taking place.

Finding 1: Are students receiving enough tutoring each week?

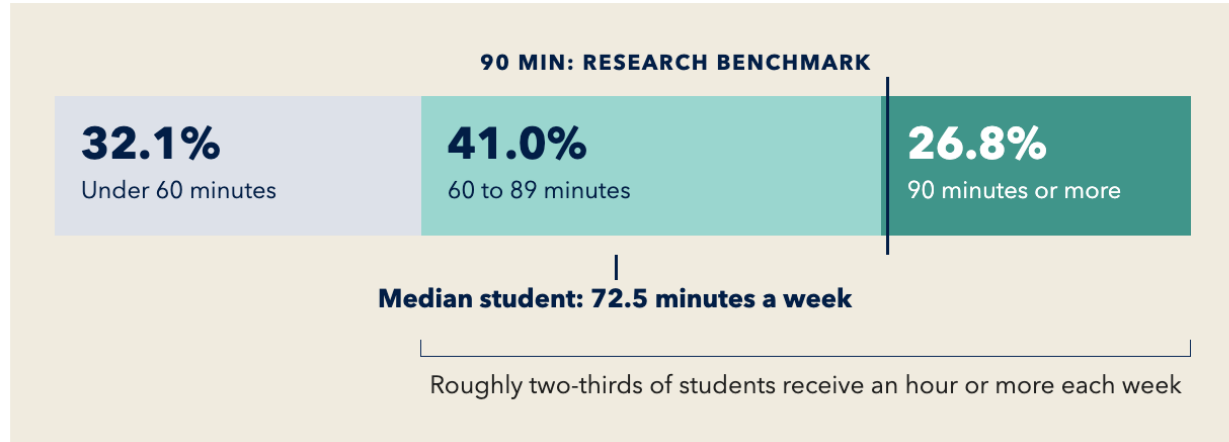
The research benchmark for high-impact tutoring works out to roughly 90 minutes of tutoring a week: about three sessions of 30 minutes each. That weekly rate is what determines whether a student accumulates enough support over a year.

The median student in this dataset received 72.5 minutes a week. Just over one-quarter, 26.8%, reached 90 minutes or more.



Most students land just below the intended weekly dose

Minutes of tutoring received per week, across all 64,044 students. The research benchmark works out to roughly 90 minutes a week.



Note: benchmark of roughly 90 minutes a week reflects three 30-minute sessions, the research-backed frequency and session length for high-impact tutoring.

Most students are not far off — another 41% received between 60 and 89 minutes a week. Together, roughly two-thirds of students received an hour or more of tutoring each week, but the distribution clusters just below the benchmark. **The gap for the typical student is not the difference between tutoring and no tutoring. It is under 20 minutes a week.**

Frequency follows the same pattern. Across all students, 20% received three or more sessions a week; 41% received two to three; 31% received one to two; and 8% received less than one.

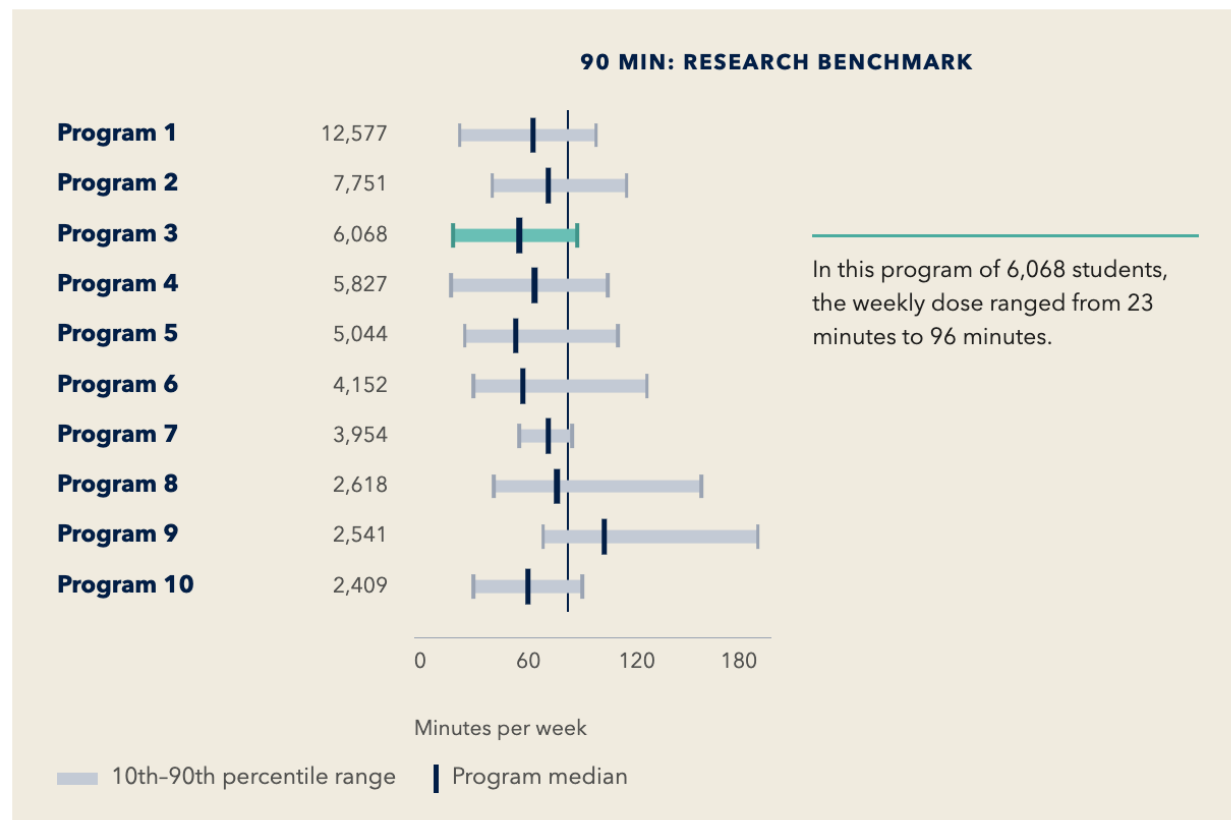
The average is not the student

If weekly minutes is the measure leaders can act on, the next question is where the variation lives.

Across this dataset, 91% of the variation in how many minutes of tutoring a student receives each week occurs among students within the same program. Only 9% is explained by which program a student is enrolled in. **A program's weekly average, in other words, describes almost none of what any individual student actually receives.**

The average describes almost no one

Weekly minutes of tutoring received, showing the range from the 10th to the 90th percentile student inside each of the ten largest programs.



In one district serving just over 6,000 students, the student at the 10th percentile received 23 minutes of tutoring a week. The student at the 90th percentile received 96 minutes — representing a fourfold difference.

The pattern holds everywhere. Every large program in this dataset contains students receiving 90 minutes a week or more and students receiving fewer than 60. There is no program where the average describes most of the students inside it.

This, again, is the visibility gap. A leader reviewing a program that averages 70 minutes a week has no way to know, from that figure, whether students are clustered tightly around it or split between one group receiving nearly two hours and another receiving barely 20 minutes.

Catherine Peretti, executive director of CitySchools Collaborative, emphasizes that dosage “is the MVP of the standards. All of the standards matter ... but dosage, those actual minutes per kid, is the support we focus on most heavily.”

It’s also worth noting that more is not automatically better. In one recent eight-district study, programs delivering substantially more instruction than the benchmark showed no detectable effects, while three programs with positive results averaged more moderate dosage. The goal is not maximum minutes, but rather to better understand which students are receiving the intended weekly dose.

Why it matters

The important questions are: which students are tracking below the intended weekly dose and how early can anyone tell? A weekly rate is a leading indicator: It can be checked in September, compared against what the program intended, and corrected while the year is still underway. Classifying a program as meeting or missing a standard at the end of the year answers a different question and answers it too late to help the students in it. Visibility into the weekly dose at the student level is what turns an implementation report into something a leader can act on.

Program averages cannot surface within-program variation. That has a practical consequence: The students furthest from the intended weekly dose are the hardest to see, because they are averaged together with the students closest to it. Identifying them requires implementation data at the student level, examined during the year rather than summarized after it. With that visibility, a leader can find the students receiving 20 minutes a week inside an otherwise healthy program and change it while the year is still underway.

Finding 2: Is that support sustained long enough?

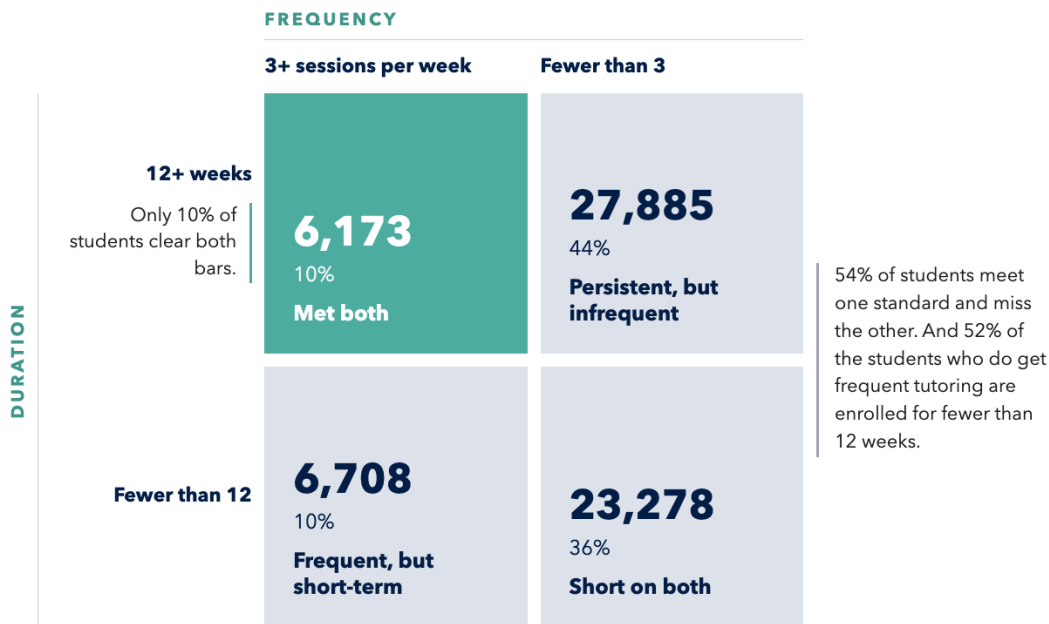
The frequency gap and the persistence gap are usually discussed as one problem. In this data, they are two, and they barely overlap. **The median student received 12 weeks of tutoring, and just over half, 53%, reached 12 weeks or more. Nearly 1 in 4 received fewer than eight weeks, and roughly 1 in 9 received fewer than four weeks.** Not all of these shortfalls reflect a program breaking down: students transfer, move out of district, graduate, or enroll mid-cycle, and some programs are designed as short-term interventions from the start.

Knowing how often a student was tutored tells you almost nothing about how long they stayed. The correlation between sessions per week and weeks enrolled is 0.03 — effectively none. A student tutored three times a week is no more likely to still be receiving tutoring in March than a student receiving tutoring once a week.

That means students fall short in different ways, and a fix aimed at one does nothing for the other. **Only 10% of students clear both bars — three or more sessions a week sustained across 12 weeks or more. Most striking, 52% of the students who do receive frequent tutoring are enrolled for fewer than 12 weeks. Frequency, on its own, buys very little.**

Two gaps, not one

Every student in the dataset, sorted by whether they met the frequency standard, the full-semester standard, both, or neither.



Note: percentages are rounded and sum to 100. Frequency standard is 3+ sessions per week; duration standard is 12 or more weeks. Correlation between the two measures across all 64,044 students is 0.03.

One number cuts against the assumption that this is a student engagement problem. Attendance across the dataset is 83% — in other words, students attended 83% of sessions. When a tutoring session is on the calendar, students overwhelmingly show up.

Attendance does, however, vary more than the headline figure suggests. By program model, it runs from **79% in state-led programs to 86% in vendor-led, with district-led at 85%**. The sharper pattern is group size: Attendance is 89% in one-to-one programs, 83% in groups of two to four, and 79% in groups of five to six. Students show up most reliably for the tutoring that most closely resembles the research ideal.

Why it matters

Programs are typically designed around frequency and group size — the conditions set at the start of the year. Persistence is not designed; it accumulates, one missed week at a time, and it is invisible until someone looks. A student who stops receiving tutoring in Week 6 still appears in the participation count, and still appears in the program's average.

Because attendance is high overall, many of the most actionable levers may be on the program-delivery side, and they are concrete: how sessions are scheduled and whether the schedule holds; whether enough tutors are staffed against the students identified; how quickly tutoring launches after students are flagged; and whether enrollment continues once it starts. None of those show up in a participation count. All of them show up in student-level weekly minutes and weeks enrolled, tracked during the year rather than summarized after it.

Finding 3: How does the student experience differ across program models and modalities?

A student's experience of tutoring depends heavily on who runs the program and how it is delivered, in ways that a single summary measure often conceals.

The sharpest difference is in who does the tutoring. **In district-led programs, 83.9% of students worked with a single tutor across the year. In state-led and vendor-led programs, 46% and 41%, respectively, did.** A student in a district program is far more likely to build a sustained relationship with one adult; a student in a vendor program is far more likely to encounter several.

The pattern reverses on frequency. Vendor-led programs put 28.1% of their students at three or more sessions a week compared with 18.8% in state-led programs and 12.2% in district-

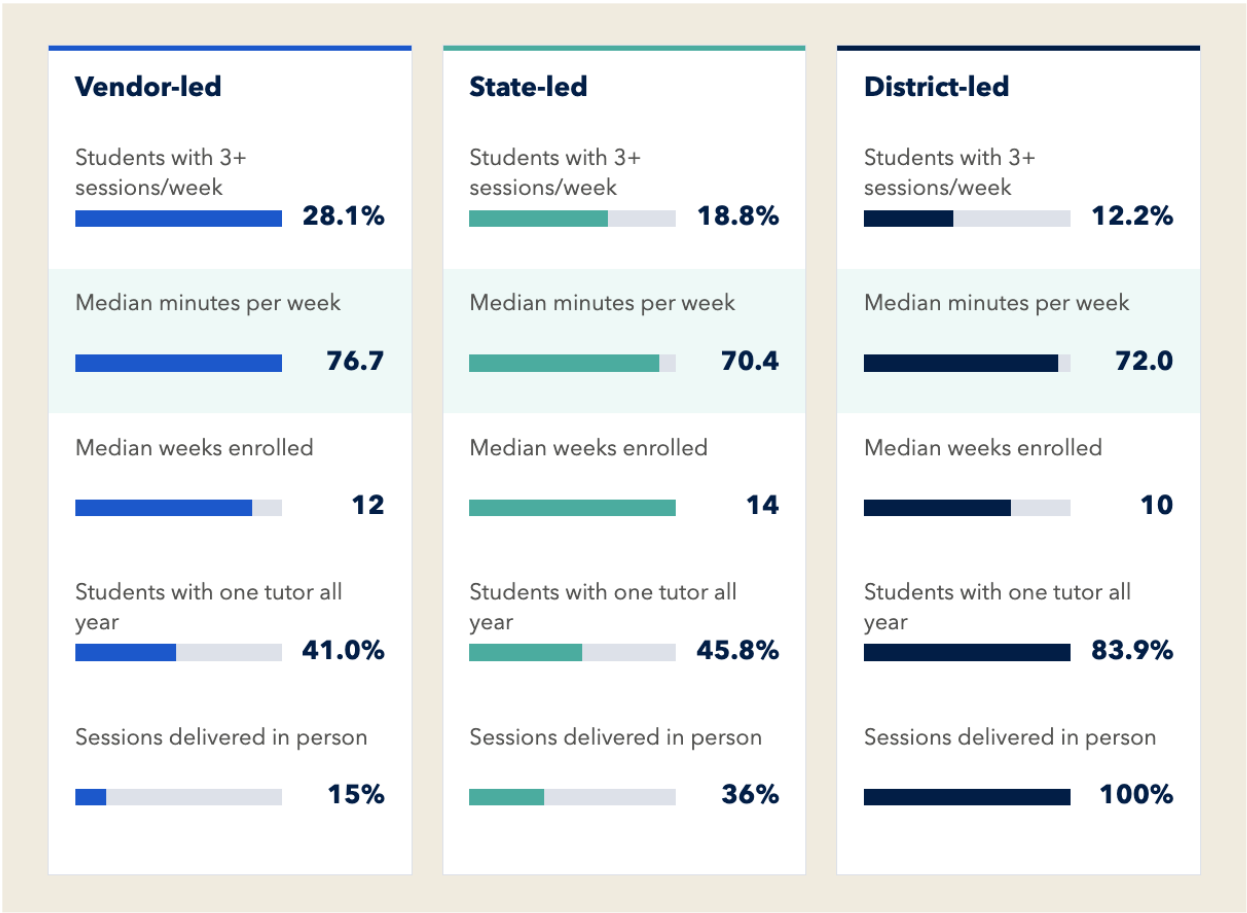
led. State-led programs run the longest, with a median of 14 weeks enrolled, compared with 12 for vendors and 10 for districts.

Yet the weekly dose is far closer than any of that suggests. **Median minutes per week were 76.7 in vendor-led programs, 72.0 in district-led and 70.4 in state-led**, a spread of about six minutes. The share of students reaching **90 minutes a week tells a similar story: 34% in vendor-led programs, 24% in district-led and 23% in state-led**. The models arrive at nearly the same weekly volume of instruction through very different arrangements of sessions, tutors and weeks.

Modality tracks the same divide. **All district-led tutoring in this dataset was delivered in person. In vendor-led programs, 15% of sessions were in person and 36% in state-led programs, with the remainder virtual.**

Three models, three different student experiences

How tutoring was actually delivered across vendor-led, state-led, and district-led programs.



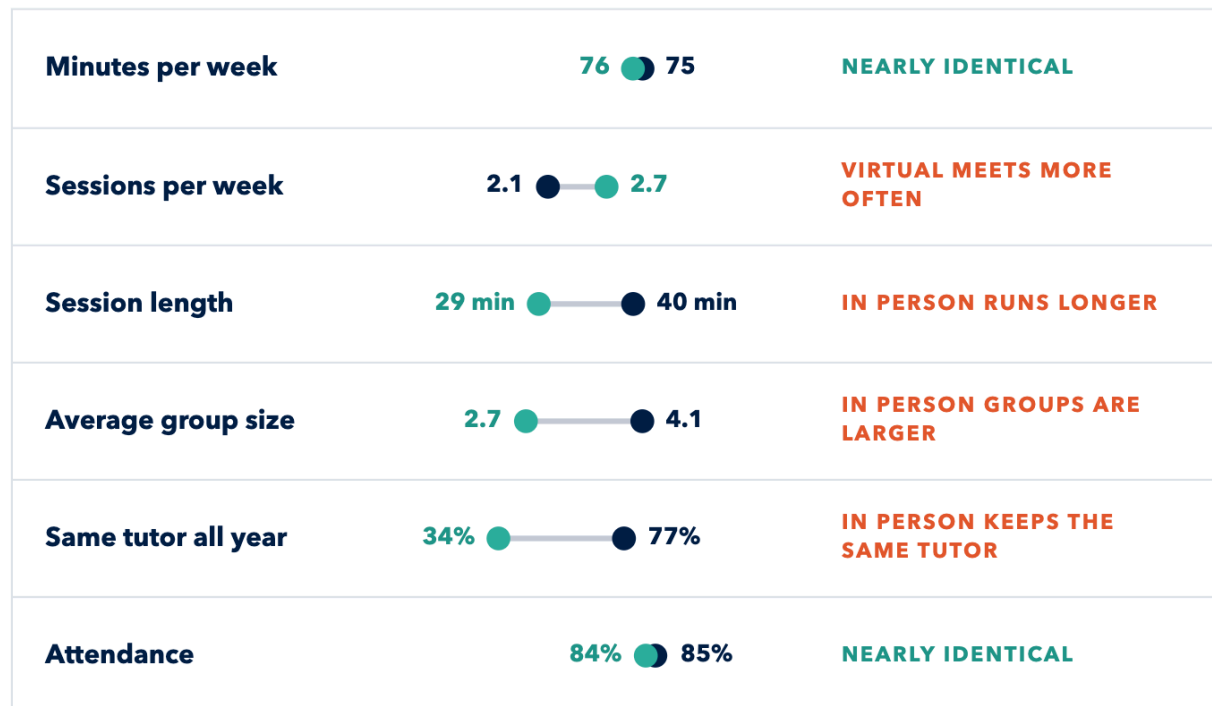
The same weekly dose, assembled differently

Overall, 61% of sessions were delivered virtually. But virtual and in-person programs deliver nearly the same weekly dose through opposite structures.

Same weekly dose, opposite structures

Programs that are almost entirely virtual, compared with programs that are almost entirely in person.

● Virtual ● In person



In this dataset, virtual programs tended to pair greater frequency and smaller groups with shorter sessions and lower tutor continuity, while in-person programs tended toward longer sessions and greater tutor continuity but lower frequency and larger groups. Consider two programs in this dataset:

- A vendor program serving 4,152 students delivers 3.1 online sessions a week, 20 minutes each, in groups of about three for about 64 minutes a week.
- A district serving 5,044 students delivers two in-person sessions a week, 30 minutes each, in groups of about six for about 60 minutes a week.

On the measure most systems collect, these two programs are identical.

Across the dataset, **the typical session ranges from 20 to 105 minutes, and the average group size ranges from 1.0 to roughly 15.** About 4,150 students are in programs where the typical session runs 20 minutes or fewer. About 11,000 are in programs averaging four or more students per group. A leader who knows that a student is receiving tutoring frequently does not know for how long or in how large of a group.

Why it matters

This is the clearest illustration of both what implementation data can do and where its limits are. It can show, for the first time at this scale, which combinations of frequency, continuity, duration and modality students are actually receiving and how sharply those combinations differ across programs that all describe themselves as offering “high-impact tutoring.” What it cannot do is say which combination works, for whom and under what conditions. That requires connecting these implementation profiles to student outcomes. Without that implementation layer, outcome studies compare programs whose actual delivery is unknown. With it, the differences visible here become testable.

Visibility into implementation is the prerequisite. Building the evidence base is the opportunity.



**Visibility into
implementation is the
prerequisite. Building
the evidence base is
the opportunity.**



From Visibility Toward Accountability: How Tutoring Systems are Evolving

From tutoring activity to tutoring effectiveness

These findings reveal why visibility into implementation matters. Understanding frequency, dosage, program model and other implementation conditions at a granular level provides critical context for understanding the student experience and which changes might improve that experience and outcomes.

As noted above, this report's analysis doesn't explore outcomes, but implementation visibility is an important first step for program accountability and improvement. Connecting implementation and outcome data can ultimately give leaders the evidence to decide which approaches to continue, adjust, expand or reconsider. Outcomes-based contracting is one application of that foundation, allowing tutoring providers to be judged in part on how much students improve. But if districts don't have the ability to collect individual students' data, then they can't use that data to make decisions about whether a given vendor has actually improved student outcomes or provided effective tutoring. Building that data infrastructure is critical.

Districts need a consistent data foundation that can continuously bring together implementation and participation information across programs and providers, giving leaders visibility into what students are actually receiving. Without that foundation, leaders may have each piece of the accountability picture somewhere in their systems but lack the connected view needed to understand what is actually working.

That progression would allow leaders to move beyond questions like “How much tutoring did we provide?” toward more consequential questions: What did students actually receive? Why are some providers more successful than others? What worked? For which students? Under what conditions? And at what cost?

What durable, well-monitored programs have in common

The visibility gaps revealed in this analysis make it difficult to determine whether students are receiving the support they need and to make the case for continued investment in these programs. They point to a broader infrastructure question: What steps can states and

districts take to better understand what students are receiving across programs and providers — and connect that implementation to outcomes over time?

Well-implemented programs take different shapes, but they share two traits: **They have a stable funding structure that allows them to operate and iterate over multiple years. And they collect data early enough to act on it rather than waiting for a data summary at the end of the year.** Putting these structures in place also makes it easier for states and districts to implement outcomes-based contracting.

As tutoring becomes a more permanent part of K-12 student support, some states and districts are beginning to build the systems needed to guide data collection, as well as outcomes-based contracting. Common data expectations, more continuous implementation monitoring, and stronger connections between service delivery and student outcomes are creating the foundation for a more accountable approach to tutoring. The examples that follow illustrate different parts of that evolution.





Arkansas



Summary: Arkansas spent three years building the data systems needed to see what tutoring students actually receive. Early on, the state could only verify part of what districts reported and had no reliable way to track how much tutoring students got. It now collects data on every session from both vendors and districts, and it gives priority for grant funding to districts that tie vendor contracts to student outcomes.

Arkansas now has one of the most advanced tutoring programs in the U.S., but it took time, iteration and experience for the state to build its data infrastructure.

In 2023, the LEARNS Act named two programs:

High Impact Tutoring, for districts and open-enrollment charters delivering math or literacy support during the school day, and the Arkansas Literacy Tutoring Grant Program, which routes funds to purchase tutoring from an approved provider list to families of K-3 students who were identified through the state screener or assessment. This has amounted to roughly \$20 million granted to about 90 districts in 2025-26.

The visibility came three years later, in stages.

- **Year 1:** Districts were required to submit student-level rosters from the start. With a two-person office, Amy Counts, who oversees tutoring for the Arkansas Department of Education, could verify only about 70% of what came in, which arrived as hundreds of individual spreadsheets with district-reported attendance rates. District after district reported perfect attendance, which seemed suspect in hindsight. The state could describe which students were being targeted but not how much tutoring the students received.
- **Year 2:** By 2024-25, the state could verify 100% of tutored students and match them to Arkansas's K-10 statewide assessment system. Dosage and frequency data remained out of reach, because districts had altered the reporting template, and there was no way to combine the files. Arkansas knew exactly who was being tutored but almost nothing about how.

- **Year 3:** Counts used leftover grant funds to begin working with Pearl to collect data from tutoring vendors in January, onboarding all but one of roughly 21 approved vendors. At the same time, the state rewrote its district reporting template to require a line of data per session attended rather than an aggregate summary, putting district-run programs on the same footing as vendor-run ones. Paired with an outcomes-based contracting cohort run with Forward Arkansas and the Center for Outcomes-Based Contracting, the state could finally compare program models, providers and modalities at its July summit. Districts on outcomes-based contracts now receive priority in the grant competition.



Counts advises other states that: “You have to invest in the data infrastructure. You can’t leave that behind ... you have to anticipate the questions leaders are going to ask you. They’re not planning ahead for what data they want, so you have to.”

Arkansas is now moving district-run programs onto the Pearl Student Support Platform, which will replace twice-yearly submissions in January and June with continuous monitoring.

Funding: Structure determines staying power

Program stability makes it possible for states and districts to invest in the infrastructure necessary to make visibility possible in the first place.

Funding is necessary but not sufficient. Stability can be challenging when tutoring is funded through annual appropriations. In those situations, district leaders can get caught in a recurring scramble: The money must be out the door within months; there is no room for long-term planning; and there is no administrative capacity to build data infrastructure. When programs are grounded in mechanisms that create durability, like multiyear authorization, predictable funding, dedicated administrative capacity, reporting requirements and consistent contracts, states and districts can build sustainable structures around the work.

When those structures are in place, program leaders are able to connect the dots between implementation and accountability, as in Massachusetts where, as Neitzel notes, “*Paying invoices for tutoring providers is conditioned on meeting the state’s data collection and sharing requirements.*”

Data collection: Seeing early enough to act

Most tutoring programs collect data, but how data is collected and used really matters. Data that arrives after the fact — a common data collection model for tutoring programs — can explain what went wrong, but little can be done with the information to help students who participated in the program; only data collected and seen in real time (or close-to-real-time) during the year can help identify what is going wrong and provide an opportunity to fix it in ways that benefit the students in the program.

As Kathy Moffitt, education administrator, Alaska Department of Education & Early Development, said, “The thing we do wrong with data is that we start down the journey planning to measure at the end of it. What we’re trying to do differently with high-dosage tutoring is monitor along the way, so you can change decision points during the year instead of analyzing at the end. That’s where we create gaps for kids.”

There is a second, underappreciated payoff. When a state requires structured implementation data as a condition of participation, it does not just enable in-year course correction — it makes program evaluation dramatically cheaper. Researchers evaluating state tutoring programs typically spend the bulk of their effort operationalizing data collection. Where the state has already required regular data collection in a consistent format throughout the school year, that work is largely done before the researchers arrive. States that build the data layer are, without intending to, subsidizing their own evidence base.

They’re also making it possible for districts to engage vendors in outcomes-based contracting, which means that tutoring providers won’t get paid if they don’t meet their own stated goals for improving student outcomes. The design challenge is threading the needle between the data a state needs and the reporting burden it places on districts. The infrastructure has to make it easy.



Where the state has already required regular data collection in a consistent format throughout the school year, that work is largely done before the researchers arrive. States that build the data layer are, without intending to, subsidizing their own evidence base.



Massachusetts



Summary: Massachusetts started its tutoring program with federal pandemic funds and now pays for it directly through a \$25 million state appropriation. Early on, the state couldn't tell which students tutoring had reached or what they experienced. It fixed that by contracting directly with tutoring providers and only paying them when they submit their data, and it is now studying tutoring's long-term impact and cost-effectiveness.

Massachusetts' tutoring program grew out of ESSER funding, but the state now explicitly funds it, with a \$25 million appropriation in the 2026 session. Early on, the state made grants to districts, and districts hired vendors from an approved list.

But the state lacked visibility into what tutoring programs were

accomplishing. In the first years, implementation was uneven, and the variations weren't visible in the outcome data. As Craig Waterman, assistant director of strategy and integration at the Massachusetts Department of Elementary and Secondary Education, describes it, early efforts to measure impact were limited by inconsistent student-level participation and implementation data. Without reliable rosters, dosage information and outcome measures, the state could not confidently determine whom tutoring had reached or what students had experienced.

Massachusetts then shifted to contracting directly with providers, and data submission became part of the vendor invoicing process. Although the expectations required some providers to adjust their data systems and practices, the state made consistent data submission a routine part of program implementation and accountability.

Massachusetts has not done rigorous cost-benefit analysis, and Waterman is direct about the fact that, per hour, tutoring is expensive relative to system-level fixes. A federal Education Innovation and Research (EIR) grant will allow Massachusetts to study tutoring's short- and longer-term impacts and cost-effectiveness and the conditions needed to implement the model successfully at scale. But at this point, Waterman can say that more dosage produces more learning; smaller groups are better; and consistent multiyear tutoring beats single-year tutoring. And the effects are largest for the lowest-performing students and those identified as at-risk.



Metro Nashville Public Schools



Summary: Metro Nashville's tutoring program serves about 8,500 students in more than 105 schools, using a mix of virtual providers, district tutors, school staff, and retired educators. Each school chooses its own approach, but the district reviews every plan to make sure it meets research-based standards for dosage, frequency, and group size. Session data helps the team spot where implementation is strong or struggling, and school visits help them understand why.

Metro Nashville's tutoring program, Accelerating Scholars, sits alongside summer school and Saturday school in what the district calls its Scholars Portfolio. Launched in 2021 as a volunteer-based, virtual high-dosage tutoring initiative, Accelerating Scholars now serves roughly 8,500 students in grades K-11 across more than 105 schools in English language arts, math and ACT preparation.

Accelerating Scholars uses a range of tutoring providers, including virtual providers, district-employed tutors, school-based staff receiving supplemental pay and a local nonprofit that engages retired educators. Schools design schedules and structures around their students' needs, so the conditions for implementing tutoring vary across buildings. Rather than requiring every school to use the same delivery model, the district works with schools to identify an approach that fits their context while maintaining the research-based elements of high-impact tutoring.

The program was initially supported through a combination of state and federal funding and is now sustained primarily through local investment.

For Accelerating Scholars, the focus is not simply on offering tutoring but on ensuring that implementation aligns with research-backed best practices. The district's program support team provides schools with guidance on building tutoring into their schedules and reviews school plans to ensure that dosage, frequency and group size reflect the evidence on effective tutoring. Tutor management systems then provide session-level data on attendance and implementation across tutoring providers and schools.

That quantitative data serves as a starting point for continuous improvement. It helps the team identify schools where implementation is particularly strong as well as

places where schools may be encountering challenges. The next step is understanding why. As Rachel Tompkins, executive director of the Scholars Portfolio, Metro Nashville Public Schools, explains, “We have the quantitative data. We now need to figure out qualitatively what is actually happening in the schools, and we need to do it systematically, because it’s 105 schools.”

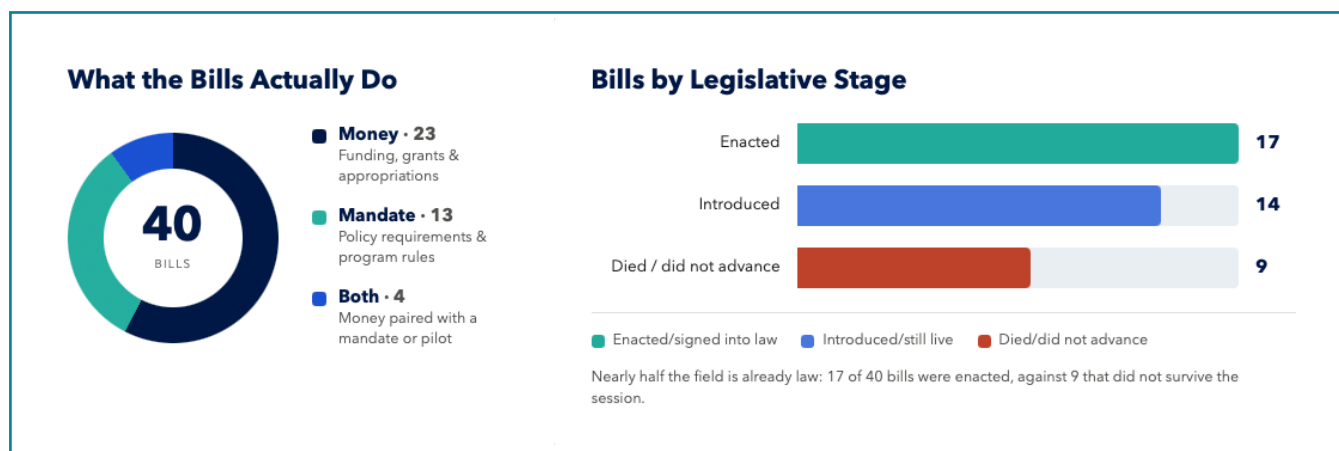
The team uses school visits, conversations with school-based leaders and implementation data together to understand what is driving those patterns. Lessons from that work can then inform training, tools and support for school-based leaders — and help the district distinguish between solutions that are specific to one school’s context and practices that it can successfully replicate across many schools.



Why This Matters Now: Tutoring is Becoming Permanent

In last year's report, we explored how tutoring evolved from an afterschool enrichment program to a core pandemic intervention and asked what the future held. While still not ubiquitous, tutoring appears to have established a foothold in K-12 education. Both parents' and districts' interest in tutoring has held steady, and policymakers in statehouses around the country have taken notice, regardless of their party affiliation. Tutoring legislation in 2026 drew 19 Republican, 10 Democratic and nine nondesignated sponsors. Across 18 states and Washington, D.C., 40 bills (including companion bills) mentioning tutoring were introduced during the 2026 legislative session, reflecting a growing effort to sustain tutoring as part of long-term state education systems. Where tutoring was once siloed, it is now gradually being folded under the broader umbrella of "intervention."

Seventeen bills have been enacted, and about 60% of them are broad appropriations vehicles in which tutoring is one funding line among many.



Policy bills, as opposed to budget lines, are rarer, and most set expectations rather than provide money. Ohio's S.B. 19 requires free intervention for students scoring at the limited level in math or English language arts and names high-dosage tutoring (three-plus times a week) as an approved option. Georgia's Math Matters Act requires structured supports, including tutoring, for advanced-math students. Louisiana's H.B. 316 extends literacy mandates through grades 4-8 with tutoring among the required interventions. Connecticut's S.B. 220 requires individualized reading plans for grades 4-9 and establishes a high-dosage literacy tutoring grant program.

Methodology: This analysis covers state-level legislation introduced or signed during the 2026 regular sessions (through July 27) that funds, mandates, defines or shapes tutoring delivery. Research was conducted through two avenues: FutureEd's 2026 state tutoring legislative tracker served as a starting point, and we then did a deep dive on each state, performing news article skimming and searching each state's legislative database for terms such as "tutoring," "high-dosage," "high-impact," "intervention," and "accelerated learning." The team then assessed notable legislation against its text to confirm tutoring was meaningfully affected. Bills moving identical language through both chambers, or clarifying a main bill's appropriations, were classified as companions and not treated as separate policy to avoid double-counting funding, though they are included in volume counts to reflect a state's overall focus on tutoring.

Read together, these bills point to a definitional problem. Each one mandates extra, personalized academic instruction, but only some call it tutoring. The rest use terms like "individualized reading plans," "structured supports" or "small-group intervention." That inconsistency has consequences. Tutoring has an established research base and field-tested standards for what quality looks like; these adjacent categories do not. When states fund personalized instruction under a different name, they lose access to that evidence base and the accountability that comes with it. You cannot measure, fund or improve a thing the field cannot consistently define or name.



Conclusion: A Baseline Year

Last year, in the wake of the ESSER funding cliff, it was unclear whether tutoring would survive. Very few districts were tracking tutoring program data. Now, as policy solidifies permanence, **tutoring has grown into more than just a short-term intervention.** In 2025-26, there was more data tracking in place, and we unearthed what we consider the field's current implementation baseline: tutoring that is close to the research ideal but delivered somewhat less often and for somewhat fewer weeks. The gaps are not one gap. **Frequency and persistence move independently, and the same weekly dose arrives in shapes so different that a leader looking only at minutes cannot tell them apart.**

This year's analysis establishes the implementation side of the picture. The next frontier is connecting these patterns to student outcomes to understand what works, for whom, under what conditions and eventually at what cost.

The opportunity now is to turn visibility into evidence. Connecting what students actually receive with what happens as a result can build a stronger evidence base for tutoring and give leaders the information they need to make better decisions about how programs are designed, delivered, improved and sustained.

