Accelerate's 2024-25 Call to Effective Action

A Synthesis of Lessons Learned

REPORT PREPARED BY

Sophie Bright, Gregory Chojnacki, Kamillah Smith, Mathematica

Jason Godfrey, Accelerate — The National Collaborative for Accelerated Learning

Acknowledgments

Accelerate would like to thank our partners at Mathematica for contributing their expertise and insights in the creation of this synthesis. We are grateful to work alongside Mathematica to share these lessons learned and appreciate their broader efforts to collect, analyze, and apply data to solve urgent challenges.

Thank you also to our grantees doing the important work of proliferating effective academic interventions and bringing evidence-based practices to scale. Special thanks to the Accelerate grantees featured in this report:

Cover photo: North Carolina Education Corps (Getty Images)

About Accelerate

The National Collaborative for Accelerated Learning

Accelerate is a national nonprofit that helps states turn strong evidence into real results for students. By aligning research, policy, and practice, Accelerate helps states scale proven strategies in public schools. Through grantmaking, research partnerships, and state implementation support, Accelerate ensures that what works in studies translates into measurable gains in classrooms nationwide.

Accelerate is supported by Arnold Ventures; Arrow Impact; Gates Foundation; Griffin Catalyst; Bill & Crissy Haslam Foundation; The Leon Levine Foundation; Overdeck Family Foundation; and Walton Family Foundation.

For more information, visit www.accelerate.us.

Executive Summary

This report synthesizes lessons learned from Accelerate's 2024–25 Call to Effective Action (CEA) cohort, with a focus on eight grantees whose evaluations and implementation experiences offer insights on high-impact, scalable tutoring and personalized learning.

Overview of the 2024–25 CEA cohort and focal grantees

In 2024–25, Accelerate funded program implementation and program evaluation research for 16 grantees administering a range of tutoring and personalized learning models in grades K–12. Across the cohort, programs varied by subject area (literacy/English language arts, math, or both), mode (in-person and virtual), staffing approach (certified teachers, paraprofessionals, trained community members, and technology-enabled tools without a human tutor), and student-tutor ratios (from 1:1 to 9:1). Evaluation designs included randomized controlled trials, quasi-experimental designs, pre-post studies, mixed methods designs, and observational studies.

This synthesis elevates findings from eight focal grantees that address critical evidence gaps (for example, curriculum alignment and understudied learner profiles) as well as emerging approaches in personalized learning, including artificial intelligence (AI)-enabled tools.

Key findings on student learning outcomes

Five of the focal grantees with rigorous[1] comparison-group research designs reported positive, statistically significant impacts of the tutoring and personalized learning models on student achievement. These effects were substantial, ranging from about 1.5 to 15 months of additional student learning. A sixth grantee reported marginally statistically significant impacts equivalent to about 1.3 months of additional learning. The remaining two focal grantees did not find evidence of impact; in both cases, low dosage—meaning the amount of tutoring students received—likely limited the effects. Among the eight grantees, five reported effects that differed based on students' gender, baseline achievement, grade level, or tutoring group size.

For focal grantees with statistically significant impacts, tutoring efficiency (or the hours of tutoring associated with one additional month of student learning) ranged from 0.7 to 14 hours—highlighting the importance of considering learning gains and dosage when comparing programs and making investment decisions.

Findings on implementation, scaling, and research

Across the focal grantees, several implementation and measurement lessons emerged:

- **To support stronger instruction,** align tutoring content to core instructional materials and pacing to improve teacher buy-in, reinforce classroom learning, and create a more coherent instructional experience for students.
- **To achieve adequate dosage (especially for tech-enabled models),** set clear student usage expectations and pair tutoring materials with routines and tools that help strengthen implementation, enable providers and school leaders to quickly monitor and communicate about student progress, and provide teachers the ability to adjust use of edtech tools to match instructional pacing.
- **To enable continual improvement of program design and implementation,** use real-time monitoring of tutoring data (e.g., attendance, formative performance measures, platform data, and observation data) to inform timely instructional adjustments, understand school-level buy-in, and address logistical hurdles.
- **To scale while maintaining quality,** invest in high-quality reusable content and curricula, improve operational and process efficiency, and prioritize strong district partnerships and integration into systems.
- **To capture data that improves evidence-based decision making,** measure cost-effectiveness of the tutoring programs and their comparison condition[2] and standardize how student dosage is reported.

[1] In this report, "rigorous research designs" are defined as randomized controlled trials (RCTs) and propensity score matched designs that demonstrated equivalence between tutored and non-tutored groups at baseline.

[2] The "comparison condition" includes any instruction, supports, and activities nonparticipating students received while the participating group of students received tutoring.

Looking Ahead

Insights from the 2024–25 CEA cohort inform Accelerate's ongoing research and field-building priorities. These priorities include addressing evidence gaps on understudied student groups, grade levels, and subjects. Accelerate is also elevating cost measurement to support education leaders' decision-making, while continuing to generate high-quality evidence—particularly on virtual tutoring at scale and expanding the evidence base on AI solutions. Providing education decision makers with the highest-quality evidence possible empowers them to identify which tools stand to benefit students most. As a result, clearer evidence on what works will create powerful incentives for service providers to continue innovating and improving the effectiveness of tutoring and other personalized learning supports.

Photo: North Carolina Education Corps

Table of Contents

I

The Call to Effective Action

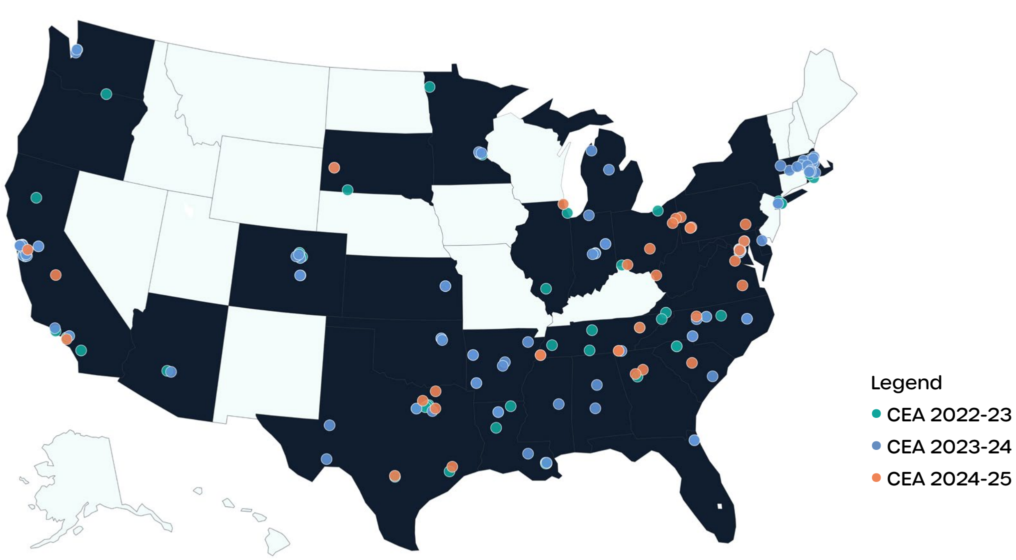
As part of its third Call to Effective Action (CEA), Accelerate awarded more than $3.5 million in grants to 16 partners to develop, scale, and evaluate sustainable, cost-effective models of tutoring and personalized learning for K-12 students across the country in the 2024-25 academic year. This synthesis focuses on findings from eight of the 16 2024-25 CEA cohort grantees, with the goal of elevating grant findings that offer novel evidence on effective approaches to high-impact, scalable personalized learning.

A) The CEA program

Since 2022, Accelerate has awarded funding to over 80 grantees through its CEA program, totaling $10 million across three cohorts with funds allocated to both program implementation and the direct costs of research and evaluation. These grants have led to nearly 50 rigorous studies, meeting ESSA Tier 2 evidence standards—including 28 randomized controlled trials (RCTs)—that have contributed much needed evidence on the implementation and effectiveness of high-dosage tutoring models.

Moreover, these grants have delivered tutoring to more than 300,000 students in 227 geographically diverse districts across the United States (Exhibit 1).

Exhibit 1. States where CEA grantees operated



Through the CEA program, Accelerate awarded different types of grants, which included Promise and Innovation grants. Promise grants were awarded to support scalable, established tutoring models that had prior evidence suggesting their effectiveness in improving student learning. The purpose of these grants was to support high-fidelity program implementation and to further develop each model's evidence base with rigorous, causal impact evaluations, including quasi-experimental designs (QEDs) and randomized controlled trials (RCTs). Innovation grants were awarded to support scalable tutoring models that had promising designs but did not yet have early-stage evidence of their impact on student outcomes. The purpose of these grants was to help high-potential tutoring models undergo further development, support high-quality implementation, and produce early evidence of their impact.

B) Overview of Accelerate's three-year CEA strategy

Although this report focuses on findings from the third and final cohort of the CEA program,[3] Accelerate's approach for the third CEA reflects the evolution of its strategy based on findings from the first two cohorts. Insights from those cohorts informed the design and investment decisions that Accelerate made in the 2024–25 CEA portfolio, as well as the development of subsequent grant programs, including the Evidence for Impact (EFI) and Call for Effective Technology (CET), which build on CEA insights to support more rigorous evaluation and scalable implementation. Specifically, in 2024, Accelerate decided to (1) fund research using a "funnel approach," that began with broad exploration across a range of tutoring designs and delivery models in the first two CEA cohorts and then used findings to narrow the focus toward specific gaps in the evidence base, (2) increase the rigor of research for both early-stage and established providers, (3) emphasize the comparability and usability of results for decision-making, and (4) pilot new data collection tools.

1. **Fund research using a funnel approach to identify and focus on gaps in the evidence base.** Starting broadly, Accelerate used the first CEA grantee cohort during the 2022–23 school year to vary its investments across multiple dimensions, including student populations (such as age groups), academic subjects, and tutoring modalities (virtual, in-person, or hybrid). In each subsequent cohort, Accelerate drew on evidence from the preceding cohort to focus on grantees that had promising designs and to answer new questions that emerged in response to new findings. By the third CEA cycle in 2024–25, Accelerate had narrowed its focus to areas that lacked a strong research consensus, maximizing the lessons learned from new grants. For example, the 2024–25 CEA did not include grants to providers focused exclusively on one-to-one, in-person, high-dosage tutoring because evidence already shows that these programs are effective—and potentially difficult to scale due to their high cost. Since previous syntheses showed that meeting dosage expectations definitely matters, the 2024-25 portfolio emphasized selection of providers with a history of tracking and meeting dosage targets.[4]

   In the 2024–25 CEA grants, Accelerate targeted specific learning gaps in the evidence base on tutoring. Accelerate's research priorities included (1) building evidence for understudied groups, (2) testing the effects of using tutoring curricula aligned with core classroom curricula, (3) exploring the role of community-based organizations (CBOs) as tutoring providers, and (4) examining education technology products such as AI-enabled platforms and other learning technology. The following grants provide examples of how this funnel approach translated into more targeted investments:

   - **Building evidence for understudied groups.** Off2Class used its grant to examine how tutoring might improve English proficiency for newcomer English learners who have low literacy skills in their home language. This effort aligned with Accelerate's key research priority of building evidence for understudied student groups who might need tutoring the most.

   - **Testing curriculum alignment as a core design feature.** Tennessee SCORE used its grant to explore the importance of alignment between tutoring curricula and core (i.e., Tier 1) curricula, finding positive and marginally significant effects of the instructionally aligned tutoring on early elementary literacy.[5]

[3] For more information about the 2024–25 CEA grantees, including program descriptions and study designs, see: https://accelerate.us/2024-cea-grantees/

[4] See Accelerate's 2022–23 and 2023–24 CEA Syntheses of Lessons Learned.

[5] The term marginally significant refers to estimates with a p-value less than 0.10 but greater than or equal to 0.05."

- **Exploring the role of CBOs in delivering tutoring.** Literacy Mid-South received a grant to investigate whether its model, as a community-based organization, could drive meaningful academic outcomes. Accelerate was interested in exploring whether CBOs might be best positioned as tutoring providers when education is part of their core identity, as it is for Literacy Mid-South, rather than a new service offering.[6] This grant was informed by prior evidence suggesting that CBO-led tutoring was less effective when delivered by nonprofessional tutors compared with trained teachers or paraprofessionals (Nickow et al. 2024).

- **Prospecting AI-enabled and computer-based programs as a scalable alternative to tutoring.** Accelerate's grants to OKO Labs and Coursemojo focused on how AI-driven personalized learning tools (which did not involve human tutors) might provide readily scalable benefits to students.

2. **Increase the rigor of research for both early-stage and established providers.** Accelerate supports a continuum of evidence across which certain methodological designs make sense for different programs, depending on their stage of development. While rigorous evaluation design has been a priority for Accelerate across all grant years, over the course of the CEA portfolios, more recent grants have increasingly focused on leveraging the most rigorous type of evaluation design—RCTs—to test program impact, resulting in a larger share of RCTs in each cohort over time. The share of grants that were RCTs rose from 11 percent in the first CEA (2022-23 school year) to 34 percent in the second (2023-24 school year) and 56 percent in the third cohort (2024-25 school year). RCTs provide the most reliable information for education leaders and tutoring providers to use when making decisions about which tutoring to utilize. Two grantees in the 2024-25 CEA cohort also randomly selected students to receive different versions of their programs, which allowed them to build high-quality evidence on how to refine program designs.

   Accelerate's Innovation grants—which focus on assessing implementation and building early evidence on student outcomes—also incorporated more rigorous designs in 2024–25 than in previous cohorts. For example, rather than simply measuring the change in student test scores between the beginning and end of the implementation period, many of these evaluations measured outcomes for a comparison group that did not receive tutoring. These comparison-group designs provide a more accurate measure of the effect of the tutoring. The number of Innovation grants with this type of comparison increased from 18 of 27 Innovation grants (67%) in the 2022-23 CEA to 8 of 10 Innovation grants (80%) in the 2024-25 CEA.

3. **Emphasize the comparability and usability of results for decision making.** As the CEA initiative progressed, Accelerate sought to make grantees' evaluation outcomes more comparable by:

   - Ensuring outcomes were reported as effect sizes, a standardized metric.
   - Synthesizing findings using a tutoring efficiency metric introduced by Accelerate.[7]
   - Providing a common approach to defining tutoring dosage.[8]
   - Piloting and refining new measurement tools (see next point).

[6] For more insights into how CBOs can best support tutoring in their communities, see Schools Need Tutoring Help from Their Communities—But Doing It Well Isn't Easy, an opinion piece written by Accelerate's Jennifer Bronson and published by The 74.
[7] See Accelerate's 2023-24 Call to Effective Action synthesis for additional information on recommendations for calculating tutoring efficiency.
[8] For additional information on definitional alignment around tutoring dosage, see Accelerate's December 2025 Quarterly Research Note (QRN).

4. **Create and pilot new measurement tools.** When a tool did not exist, Accelerate developed new tools to address persistent gaps in tutoring research—especially the lack of standardized measures of dosage, cost, and comparison conditions—to make evaluation findings more comparable and useful for decision makers. Accelerate then used its grantmaking as a test bed to pilot, refine, and disseminate these new tools, including its cost tool, DATA standards, and a teacher survey, which are highlighted in more detail throughout this report.

   - **Accelerate's cost tool.** In February 2025, Accelerate introduced a cost tool that standardizes the process of calculating program costs for any tutoring program model and was used by three 2024–25 CEA grantees.

   - **Increased consistency in reported data, leading to Accelerate's DATA standards.** To improve consistency and comparability in tutoring measurement, Accelerate developed the Data Alignment and Tutoring Assessment Standards (DATAS), an open, user-friendly reporting framework and toolkit that defines a practical common language for tutoring records at both the session and student level.

   - **Accelerate's teacher survey, used to gather information on the comparison condition.** Accelerate partnered with Drs. Robin Jacob and Catherine Asher at the University of Michigan's Youth Policy Lab to refine a teacher survey to collect information on a tutoring evaluation's comparison condition—the types of instruction, supports, or activities that non-tutored students received while tutored students received tutoring. One grantee in the 2024–25 CEA cohort and all four grantees in Accelerate's 2025-26 EFI program have used or will use this survey.

   Beyond tools, Accelerate also drew on the expertise of partners to support early-stage grantees in its 2024–25 Innovation portfolio, contracting with Mathematica to provide technical assistance to these grantees when planning and conducting their evaluations. This assistance led Innovation grantees to use more rigorous methods (three RCTs and five comparison-group designs) and to describe their methods more clearly, ultimately helping them produce more meaningful evidence, even in early-stage evaluations. Mathematica's technical assistance also promoted uniform measurement of the amount of tutoring students received during the evaluations, a key goal as Accelerate seeks to normalize the measurement of tutoring efficiency and cost-effectiveness to inform decision makers.

By the 2024-25 CEA, Accelerate refined and increased the transparency of their application process, including clearer guidance on reporting requirements and grant expectations. For select Promise and Adoption applicants, Accelerate added an interview step. They also strengthened their approach to grantee selection by examining applicants' existing conditions for successful implementation and evaluation such as committed district partnerships, evidence of strong dosage delivery, and a track record that suggested readiness to execute at scale.

Building on lessons from the first two cycles, Accelerate also shifted how evaluation and funding were structured. They developed and funded external research partnerships to lead rigorous third-party impact evaluations of Promise and Adoption[9] grantees, and they made an intentional decision to coordinate the research function in-house so the evaluations were truly independent. Finally, grantee payments were aligned to program implementation dates and key deliverables. Accelerate's goal through this process was to select grantees that would deliver cost-effective services at sizable scale and improve outcomes for students, particularly students in historically underserved communities.

[9] Accelerate's Adoption Grant targets scalable, established tutoring models with prior evidence of causal impact on student outcomes, with the primary aim of testing program implementation fidelity and scalability through a Randomized Control Trial (RCT) in new districts, i.e. replicating results in a different context. These grants were intended to support a sizable expansion of program delivery — whether into new school settings, to new student populations, or through other well-considered adaptations not previously covered by the grantees' evidence of program impact.

Overall, throughout the progression of the CEA program, Accelerate sought to generate practical, action-ready findings for the education field on the implementation of high-quality tutoring models and the interpretation of their effects on student achievement. As this synthesis—and previous CEA syntheses—have outlined, this effort has resulted in a better understanding of high-leverage tutoring approaches that can move the needle on outcomes for students. These lessons learned come at a critical moment in the landscape of tutoring and tutoring research, where investment is strong but scale-up has remained challenging due to inconsistencies in implementation. The remainder of this report helps to address that challenge by highlighting key takeaways in the context of findings from the 2024–25 CEA cohort and providing lessons learned from implementation of tutoring across a range of tutoring types and contexts.

Photo: Tennessee SCORE

C) Report focus and structure

This synthesis primarily focuses on findings from eight of the 16 2024–25 CEA cohort grantees. We selected these eight grantees based on multiple criteria (described in Section III), with the goal of building novel evidence on effective approaches to high-impact, scalable personalized learning.

Before we present findings from the eight focal grantees, this report describes our approach to developing this synthesis (Section II) and summarizes the key program features of the 16 grantees in the 2024–25 CEA Innovation and Promise cohort (Section III). It then discusses notable findings on program effects, efficiency, and design features (Section IV); themes on scaling; and remaining research gaps (Section V). The report concludes with a summary of key takeaways—including policy-relevant implications for the field—and Accelerate's planned next steps (Section VI).

Throughout the report, readers will find callout boxes from Accelerate-sponsored research studies and strategic insights across its three CEA cohorts. These two types of highlights might be particularly useful for other intermediary organizations, tutoring providers, researchers, or funders of tutoring initiatives.

Analytic or strategic insights

These callout boxes offer helpful insights on analytic approaches, the interpretation of impact or efficiency findings, and implications for strategic decision making.

Accelerate's strategic evolution

These callout boxes tell the story of Accelerate's strategic evolution across three CEA cohorts and note implications for future programming.

Photo: Air Reading

Data sources and methods

To inform our synthesis of the eight focal grantees, we used the following data sources and methods.

Data sources. This report presents findings from three sources: (1) evaluation reports from the evaluators of the 16 2024–25 CEA grantees, describing the impact of grantees' programs on student learning; (2) responses from the 16 grantees to a survey administered by Accelerate in June 2025; and (3) interviews with the eight focal grantees to capture more information to contextualize notable program features and findings.

To guide our interview questions with the eight grantees, Mathematica researchers co-developed a list of domains with Accelerate about what we hoped to learn from the grantees as a whole. We then tailored questions to the domains relevant for specific grantees and the grantees' specific findings to maximize the additional information each interview would yield.

Selection of the eight focal grantees. Mathematica worked with Accelerate to select the eight focal grantees based on multiple criteria, with the ultimate goal of elevating grant-related findings that offer novel evidence on effective approaches to high-impact, scalable personalized learning. Specifically, we selected grantees based on a combination of the following criteria: research design, findings on specific student populations, notable program design features, and AI applications. All focal grantees except OKO Labs used a matched comparison or RCT design. We chose several grantees because they examined which students might benefit most from specific models, including models that had outsized impacts on English learners, students with lower baseline literacy skills, and boys. We selected two of the grantees because they provide evidence on AI-driven personalized learning tools.

Analysis methods. To conduct this analysis, the research team used an AI-supported approach to extracting and synthesizing findings and key takeaways from grantees' evaluation reports and the grantee interviews. Specifically, the team first used ChatGPT 5.1 to extract key program design information (such as grade level, modality, and dosage) and findings on student improvement (including differences for student groups) from each evaluation report, and the team reviewed responses from grantee surveys. The Mathematica team then worked with Accelerate to select the eight grantees for interviews and develop interview protocols focused on key topics. After conducting each interview, Mathematica researchers used ChatGPT 5.1 Pro to identify key findings from interview transcripts related to the topics of interest, reviewed these findings, and synthesized them into this report.[10]

[10] Both uses of AI (extracting data from grantee reports and summarizing interview transcripts) drew on applications of large language models (LLMs) that Mathematica has validated internally, including one through an internal RCT. The research team also regularly cross-checked LLM extracts with source materials (such as grantee reports and interview transcripts) to confirm accuracy and spot-check for errors. In the case of interview summaries, the AI-supported analysis was conducted by one of the two lead interviewers who had conducted the interview as an additional layer of validation.

II
Overview of 2024–25 CEA Grantees

Accelerate selected 16 grantees from a pool of nearly 200 applicants to participate in the third CEA cohort and implement their tutoring programs in the 2024–25 academic year. Seven of these grantees received Promise grants, and nine received Innovation grants. The grantees included high-dosage tutoring programs implementing models that served 20,000 students from kindergarten through 12th grade across 39 districts in 13 states and Washington, DC. Grantees delivered tutoring in literacy or English language arts (ELA), math, or both subjects.

Grantees varied widely in organizational size and reach, but the CEA-funded programs they implemented for their evaluations often served a narrower subset of students. A typical grantee's program had about 124 tutors (from 0 to 755 tutors, with 0 being for tech-enabled providers) and served about 719 students on average (from 18 to 5,000 students). Compared with all districts nationally, the districts served by these grantees had larger proportions of students who qualified for a free or reduced-price lunch, had an Individualized Education Program (IEP), or identified as multilingual learners (Exhibit 2). Students served by grantees were also more likely to be from underrepresented racial and ethnic groups (Exhibit 3).

Exhibit 2. Demographics of students in 2024–25 CEA districts versus the average U.S. district

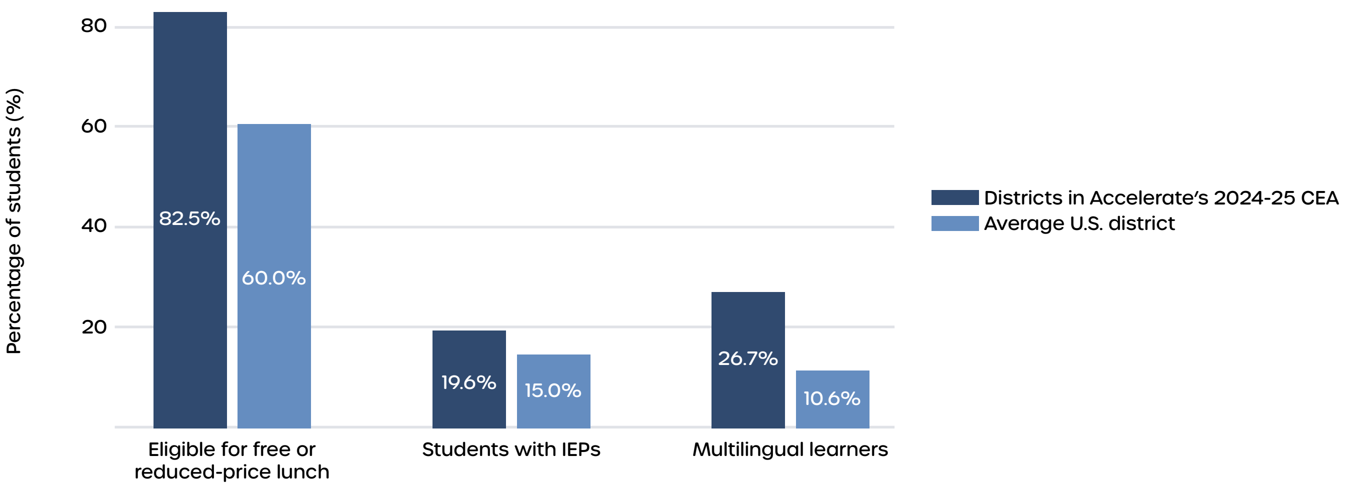
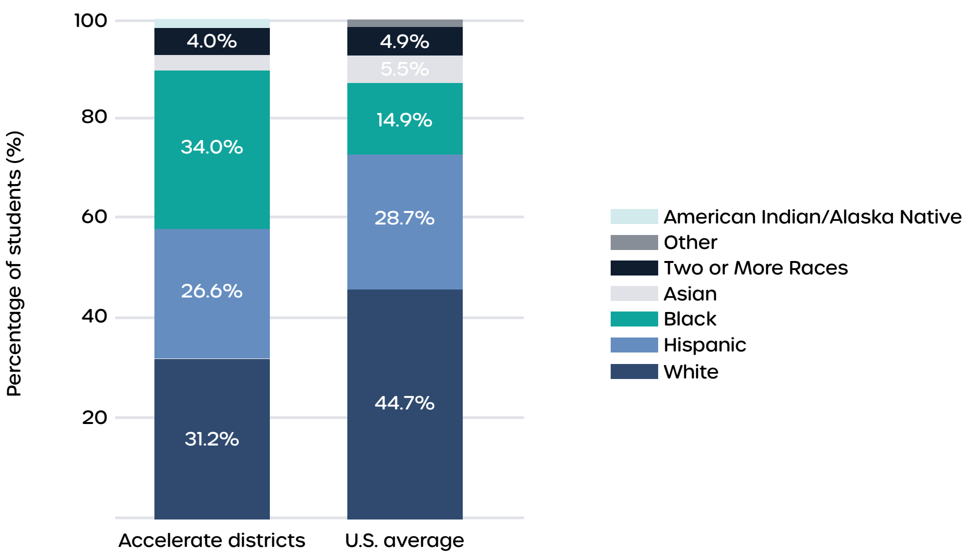


Exhibit 3. Race and ethnicity of students in 2024–25 CEA districts versus the average U.S. district



Tutoring program features. Accelerate selected a diverse set of grantees in terms of their subject-area focus, grades served, mode of delivery, and tutor type for the 2024–25 CEA cohort (Exhibit 4). This diversity is evident across the full portfolio of 16 grantees and among the eight grantees selected for deeper analysis. Collectively, these programs span literacy and math, elementary through high school grades, in-person and virtual delivery models, a range of tutor types, and multiple rigorous evaluation approaches.

Photo: Literacy Mid-South

Exhibit 4. Key program features of grantees

Grantee	Subject focus	Grades served	Mode of delivery	Tutor types	Study design	Student-tutor ratio	Scheduled dosage	Actual dosage
Air Reading	Literacy	K-6	Virtual	Paraprofessionals with a bachelor's degree or above and at least one year of teaching experience	RCT	1:1 to 4:1	14 weeks and 27 hrs. (Cohort 1) 19 weeks and 37 hrs. (Cohort 2)	13 weeks and 19.9 hrs. (Cohort 1) 17 weeks and 27.1 hrs. (Cohort 2)
Carnegie Learning	Math	6–8	Virtual	State-certified math teachers	Matched comparison	3:1 to 4:1	12 weeks 14 hrs.	12 weeks 11.53 hrs. (School A) 7.97 hrs. (School B)
Carnegie Mellon University (PLUS Tutoring)	Math	6-8	Virtual	Trained undergraduate and graduate college students	Regression discontinuity	2:1 to 10:1	20 weeks 15 hrs.	15.9 weeks 19.2 hrs.
Cignition	Literacy and math	K–12	Virtual	Certified teachers	RCT (includes cost analysis)	4:1	30 weeks 60 hrs.	26 weeks 35.9 hrs.
CitySchools Collaborative	Literacy and math	3–8	In-person	Paraprofessionals, college students, and other nonteacher tutors	Weighted comparison[11]	4:1	10 weeks 15 hrs.	14.8 weeks 27.3 hrs.
Coursemojo	ELA	6	In-person	AI-powered assistant teacher (nonhuman), with support from classroom teachers (no human tutor)	Matched comparison	1:1	28–32 weeks 10-30 hrs.	28–32 weeks 4.5–13.5 hrs.
ExpandEd Schools	Literacy and math	Grades K–2 (literacy) Grades 6–8 (math)	In-person	Trained tutors	Pre-post regression analyses	1:1 to 4:1	10 weeks 15 hrs.	10 weeks 14.6 hrs.
Future Forward	Literacy	K–2	In-person	Professional tutors	RCT	5:2	13 weeks 19.5 hrs.	13-14 weeks 16.15 hrs.
Literacy Mid-South	Literacy	2–5	In-person	College students and community members with at least a high school diploma	RCT	3:1	18 weeks 40.5 hrs.	14–19 weeks 32.25 hrs.
Magpie (Kinder Beta)	Literacy	K	Virtual	Teachers oversee use of digital literacy platform	Matched comparison	1:1	30 weeks 20 hrs.	22 weeks 7.97 hrs.
Math Corps (Ampact)	Math	K-3	In-person	Trained AmeriCorps members	Pre-post	1:1 or 3:1	7-25 weeks[12] 6-25 hrs.	7-25 weeks 6-25 hrs.
North Carolina Education Corps	Literacy	K-3	In-person	Retirees (including former teachers), parents and caregivers, college students, and community members	RCT (with cost analysis)	3:1	28–30 weeks 50 hrs.	14.1 weeks 21.3 hrs.
Off2Class	Literacy	6–12	Virtual	English language development teachers or a professional online tutor trained to deliver the curriculum	RCT (with cost analysis)	1:1 to 9:1	17 weeks 19 hrs.	12 weeks 10.3 hrs.
OKO Labs	Math	4–5	Virtual	AI-facilitated intelligent tutor with oversight by 4th- and 5th-grade teachers (no human tutor)	Pre-post	3:1 to 5:1	9 weeks 9 hrs.	9 weeks 2.52 hrs.
Teachley	Math	2–4	In-person with a mix of physical and digital tools	Paraprofessionals	Matched comparison with a small RCT substudy	4:1	12 Weeks 15 hrs.	19 Weeks 10.8 hrs.
Tennessee SCORE	Literacy	K-3	In-person	Teachers and paraprofessionals	RCT	3:1 to 4:1	24 weeks 83.7 hrs. (grades 1–3) 14 weeks 49.1 hrs. (Kindergarten)[13]	18.4 weeks 63.4 hrs.

[11] Inverse Probability of Treatment Weighting (IPTW) is a method used to control for confounding in observational data by assigning weights to treatment and comparison observations that result in the weighted samples having similar observed baseline characteristics. This method was implemented using grade band-specific analyses that the grantee aggregated using weighted averages by n-size to produce overall estimates.
[12] The wide range in Math Corps's reported scheduled and actual dosage reflects a key design feature of the tutoring model, which is the constant progress monitoring of participating students. Students are cycled in or out of tutoring depending on their need, resulting in a flexible dosage range rather than one that is prescribed based on program design.
[13] While Tennessee SCORE's tutoring program began in October 2024 for students in grades 1-3, tutoring for kindergarten students began in January 2025 to allow them time to acclimate to school.

Comparing eight focal grantees with the 2024–25 CEA cohort. The program features of the eight focal grantees are similar to those of the 2024–25 CEA cohort, with two exceptions: the eight grantees had more rigorous study designs and more of an ELA or literacy focus. Exhibit 5 shows the similarities and differences of the eight focal grantees versus all 2024–25 CEA grantees.

Exhibit 5. Comparison of all 2024—25 CEA grantees and eight focal grantees selected for interviews

Program feature	All 2024-25 CEA grantees	Eight selected grantees
Study design	44% RCT, 25% matched comparison; 13% pre-post, 13% regression or weighting analyses; 6% regression discontinuity	63% RCTs; 25% matched comparison; 13% pre-post
Subject	50% literacy or ELA; 31% math; 19% mixed subject	75% literacy or ELA; 25% math
Grade range	31% early elementary (K–3); 19% elementary (1–5); 19% middle school (6–8); and 31% spanning across grade bands, including students in high school (9–12)	25% early elementary (K–3); 25% elementary (1–5); 25% middle school (6–8); and 25% spanning across grand bands, including students in high school (9–12)
Tutor type	13% certified teachers only; 19% paraprofessionals or teachers only; 19% digital tutors (AI-based or digital platform); 38% noncertified trained tutors	25% certified teachers only; 13% paraprofessionals or teachers only; 25% digital tutors (AI-based or digital platform); 38% noncertified trained tutors
Modality	50% in-person; 44% virtual; 6% mix of in-person and virtual	50% in-person; 50% virtual
Student-tutor ratio	1:1 to 10:1	1:1 to 9:1

NOTE: Percentages may not sum to 100% due to rounding


Photo: Carnegie Learning

III

Synthesis of Findings for Eight Selected Grantees

This section synthesizes the impact and implementation findings from the eight focal grantees from Accelerate's 2024-25 CEA. Here we describe their programs' effects on student achievement, the amount of tutoring needed to achieve those effects (tutoring efficiency), and lessons learned from grantees' program design.

A) Student outcome findings across grantees

Key takeaways

- Five of the seven grantees with rigorous comparison-group research designs had positive effects on student learning. These positive effects were substantive, from 1.5 to 15 months of additional student learning. The sixth grantee had marginally significant impacts on student literacy equivalent to 1.3 months of extra learning. The seventh grantee did not show an effect on student learning, likely due to low dosage.
- The eighth grantee did not use a rigorous design but instead focused on developing early-stage evidence as a foundation for conducting an RCT in 2026. This grantee did not show an effect on student learning.

This synthesis examines how grantees' effects on student learning varied by program features (such as virtual versus in-person tutoring). However, these comparisons are descriptive and do not establish whether any specific program feature caused larger impacts compared to others because programs differed in multiple ways and program features were not experimentally tested. The comparison between virtual and in-person tutoring illustrates this: for example, nearly all virtual tutoring grantees focused on literacy, whereas in-person grantees spanned both literacy and math. As a result, any observed differences between virtual and in-person tutoring might reflect the delivery mode, subject focus, or other program characteristics.[14]

Student learning improvement.
Among the seven grantees that had more rigorous research designs (RCTs or matched comparison designs), five had positive, statistically significant impacts on student learning (Exhibit 6).[15] Two of these grantees had large effects (greater than 0.2 standard deviations [SDs]), and three had moderately sized effects (0.1 to 0.2 SDs). The other two grantees did not have a statistically significant effect on student achievement.[16] Likewise, the additional grantee that did not use a rigorous design did not have a statistically significant effect on student achievement.

[14] A focus of the 2025-2026 EFI portfolio includes multi-arm RCTs to test the direct effect of program design features, a development in the research scope of Accelerate's portfolio that demonstrates a continued focus on addressing known unknowns in the tutoring literature.
[15] We included Literacy Mid-South in this count because it had a statistically significant impact (0.20 SDs) on lower-achieving students in the study sample (students in the bottom half of the achievement distribution at baseline). Literacy Mid-South typically serves lower-achieving students but had to expand eligibility for this study due to recruitment challenges. As a result, the sample of students in the evaluation included high-performing students with an average baseline i-Ready score at the 80th percentile of the national distribution. Because Literacy Mid-South's typical students are lower performing, we focused on the impact estimate for these students in the sample (students with an average i-Ready score at the 49th percentile nationwide).
[16] One of these two grantees, Tennessee SCORE, measured an impact of 0.12 SD that was marginally significant with a p-value less than 0.10. That evaluation compared the outcomes of students who received similar amounts of tutoring but in which one group's tutoring used curriculum aligned with core classroom instruction while the other group's tutoring used a different curriculum to test the value of curricular alignment.

Exhibit 6. Program design features and dosage for eight interviewed grantees

Grantee/tutoring provider	Grade	Modality	Subject	Learning improvement (SDs)	Scheduled dosage	Actual dosage	Study design and strength
Air Reading	K-6	Virtual	Literacy	0.29*	14 weeks and 27 hrs. (Cohort 1) 19 weeks and 37 hrs. (Cohort 2)	13 weeks; 19.9 hrs. (Cohort 1) 17 weeks, 27.1 hrs. (Cohort 2)	● RCT
Carnegie Learning	6–8	Virtual	Math	0.14**	12 weeks 14 hrs.	12 weeks 11.53 hrs. (School A) 7.97 hrs. (School B)	◐ Matched comparison
Coursemojo	6	In-person	ELA	-0.01	28–32 weeks 10-30 hrs.	28–32 weeks 4.5 hrs.- 13.5 hrs.	◐ Matched comparison
Literacy Mid-South	2–5	In-person	Literacy	0.20*b	18 weeks 40.5 hrs.	14–19 weeks 32.25 hrs.	● RCT
North Carolina Education Corps	K–3	In-person	Literacy	0.18*	28–30 weeks 50 hrs.	14.1 weeks 21.3 hrs.	● RCT
Off2Class	6–12	Virtual	Literacy	0.35*	17 weeks 19 hrs.	12 weeks 10.3 hrs.	● RCT
OKO Labs	4–5	Virtual	Math	0.02c	9 weeks 9 hrs.	9 weeks 2.52 hrs.	○ Pre-post assessment
Tennessee SCORE	K–3	In-person	Literacy	0.12†	24 weeks, 83.7 hrs. (grades 1–3) 14 weeks, 49.1 hrs. (Kindergarten)	18.4 weeks 63.4 hrs.	● RCT

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

[a] ● High ◐ Moderate ○ Emergent

We define strong studies as RCTs with a low level of attrition, as defined by the What Works Clearinghouse. We define moderate-strength studies as matched comparison designs or RCTs with a combination of deviation from random assignment procedures, high attrition, and significant differences in baseline characteristics. We categorize one study as emergent to reflect its pre-post design.

[b] Students near the national median i-Ready score.

[c] The evaluation compared test score growth among OKO Labs users to their projected growth, based on typical growth among students with similar baseline scores, and showed that the difference was not statistically significant (p = 0.79). In other words, students who used OKO Labs in this early implementation study had learning growth that was similar to growth among their peers nationally.

Average student learning improvement by modality, grade level, and subject.

Prior research suggests smaller impacts for virtual tutoring programs versus in-person tutoring, and for tutoring for middle and high school students versus elementary school students (see Nickow et al. 2024). However, our evaluations of the eight focal grantees suggest that these patterns might not always hold true. Although the number of studies in this analysis is too small to rigorously compare the effect of different program features, we averaged the learning improvement by modality, grade level, and subject areas to explore whether patterns in our studies differ from the patterns in prior evidence (Exhibit 7). Grantees that delivered virtual tutoring and those that served middle and high school students both had large effects on student achievement, contrary to prior evidence. This suggests it might be valuable to continue exploring the use of tutoring in areas where previous research showed smaller impacts. Although the two math-focused programs in our sample exhibited smaller impacts than the literacy-focused programs, it is important to note that students were engaged with each program for less than 15 hours, on average. Moreover, one of the programs, Carnegie Learning, exhibited a high degree of efficiency in terms of how hours of tutoring translated to gains in student learning, as discussed in the next section.

These findings suggest that tutoring models have evolved over time to address challenges associated with tutoring virtually, with older students, and in math. For example, previous evidence reviews had few virtual studies to draw on, and investments in remote learning and educational technology (edtech) by many districts during the pandemic might have improved some conditions that enable successful virtual tutoring. At the same time, virtual providers might have enhanced their implementation quality and student interfaces by incorporating lessons learned in early implementation (Conroy et al. 2023).

Exhibit 7. Average learning improvement across eight grantees by modality, grade level, and subject

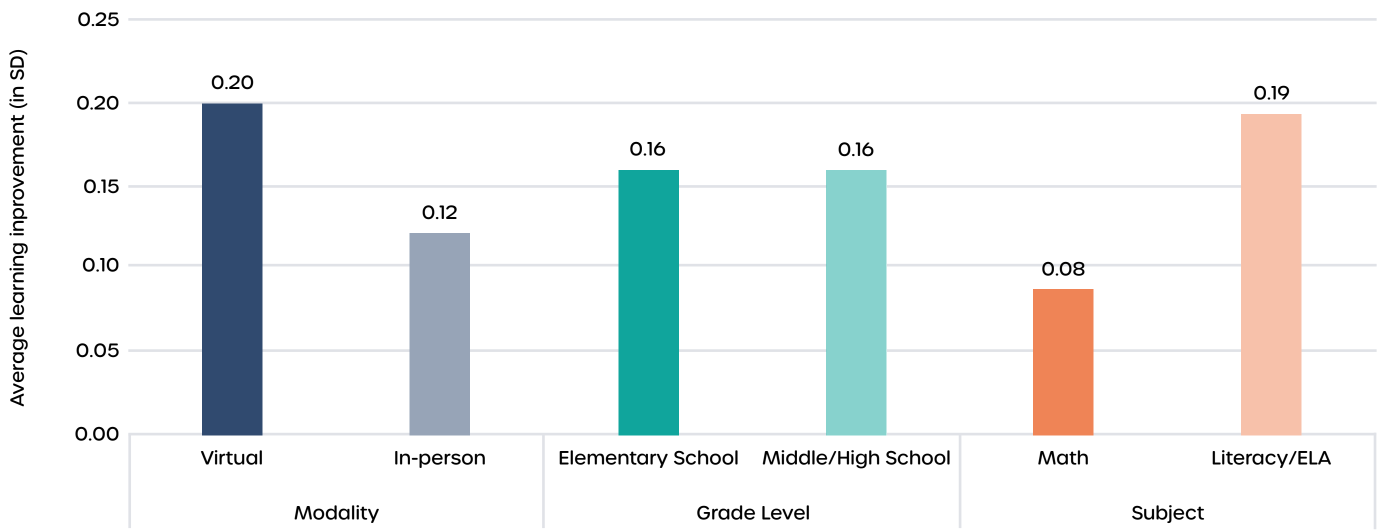


Photo: Literacy Mid-South

Effects of tutoring programs for student groups and group size

Among the eight grantees, five reported effects that differed based on students' gender, baseline achievement, grade level, or tutoring group size.

Gender: Three grantees, Air Reading, Tennessee SCORE, and North Carolina Education Corps, reported that their tutoring had larger effects on boys than on girls. All grantees said more research was needed to confirm these differential effects and to better understand them, consistent with recent evidence that gender-based differences in learning trajectories require nuanced examination (Kuhfeld and Burchinal 2025). As initial hypotheses, North Carolina Education Corps suggested the structured, small-group format and strong tutor-student relationships might particularly benefit boys' engagement and focus. The grantee also said observed gender differences could partly be a reflection of other factors, specifically citing the need to investigate whether boys were more likely to be assigned to tutoring groups and whether girls in some sites were more likely to be English learners. Another grantee—Tennessee SCORE—offered similar hypotheses: (1) higher incidence of low baseline performance among boys in early grades (see the next section on "Baseline achievement" for evidence demonstrating larger effects for lower-performing students) and (2) the potential for tutoring to support boys' learning styles through immediate corrective feedback and small-group settings.

Analytic insight: Findings on differential effects by gender demonstrate the importance of reporting effect sizes by student groups

The differential effects by gender—specifically larger effects for boys compared to girls—found by three of the 2024-25 CEA grantees demonstrate the broader importance of reporting effect sizes by various student groups (including student characteristics, baseline achievement, and tutoring program design). In turn, the field should use these data to strengthen meta-analyses and investigate potential links between specific program characteristics—such as in-person delivery—and larger effect sizes among specific student groups. This type of analysis can sharpen our collective understanding of what tutoring models work for whom and under what conditions. Equipped with this information, decision makers can better select models that might work better for certain students or in certain settings.

Baseline achievement: Two grantees, Literacy Mid-South and Tennessee SCORE, found that their tutoring had larger effects on lower-performing students than on higher performers. Literacy Mid-South hypothesized that these larger effects might be a result of its program's focus on below-grade-level or foundational literacy skills.. Meanwhile, Tennessee SCORE said this greater effect on lower-performing students might be a result of its program's efforts to align tutoring with Tier 1 instruction, potentially reducing the cognitive burden in accessing instruction, which is especially important for students with lower beginning skill levels.

Grade level: Two grantees, North Carolina Education Corps and Off2Class, said the effects of their tutoring varied by grade level. North Carolina Education Corps had the largest effect on kindergarteners' achievement (about 0.4 SDs on both the DIBELS composite and a study-created index that focused on the foundational literacy skills most targeted by tutoring). Interestingly, among 3rd graders, the effect on students' DIBELS composite score was lower than that on kindergarteners', but the index of targeted skills showed a large impact (0.3 SDs). Program leaders said this might reflect that the DIBELS composite score for higher grades places less weight on foundational literacy skills—the skills that lower-achieving students focus on during tutoring. This suggests that tutoring can yield large impacts in the early elementary grades, but impacts in later elementary grades might be concentrated on foundational literacy skills that require careful measurement to detect.

Off2Class had larger effects on achievement for high school students than for middle school students. The grantee attributed this to larger effects of the tutoring on long-term English learners, who are more likely to be in high school than middle school. Off2Class found these results surprising given that many long-term English learners lose motivation after years of participation in English language services and repeated testing with minimal improvement.

Tutoring group size: Tennessee SCORE reported differential effects by tutoring group size, finding higher impacts on groups of four students than on groups of three. This finding contradicts the expectation that smaller student-tutor ratios are beneficial, but the grantee suggested two possible reasons for it: (1) the larger group creates more opportunities for students to talk with one another during peer learning (they can talk in pairs or with partners in a group of four) and (2) pacing discipline (larger groups reduce the tendency to linger on a single student's difficulty).

B) Tutoring efficiency of grantees

Key takeaways

- 2024–25 CEA grantee programs were much more efficient than programs examined in prior research and prior CEA cohorts, suggesting that tutoring programs funded in this cohort identified time-saving ways to support student learning.
- Consistent with the prior research, the 2024-25 CEA grantees found that math tutoring tends to require fewer hours to generate a month of learning than literacy programs. However, one literacy grantee who served English learners demonstrated efficiency on par with math programs.

Evidence on the amount of tutoring (or the dosage) needed to yield additional learning—also known as tutoring efficiency—is critical for decision makers considering whether and how to deploy tutoring. Programs that yield the same amount of additional learning with fewer hours of tutoring could maximize the benefits of tutoring within the constraints of district budgets and school hours. We used a measure developed by Accelerate (Kohlmoos and Steinberg 2024) to explore the tutoring efficiency of Accelerate's 2024–25 CEA grantees. Consistent with previous studies, we found that math tutoring might be more efficient than literacy tutoring overall, but not for English learners.

We calculated the tutoring efficiency index for the five grantees with a positive, statistically significant effect on achievement.[17] The index is based on the hours of tutoring associated with one additional month of learning (Kohlmoos and Steinberg 2024). Grantees with a lower index value are more efficient because they need fewer hours to yield one month of additional learning. Three of the five grantees—Air Reading, Literacy Mid-South, and North Carolina Education Corps—had efficiency index values of around 10 hours per month of additional learning, but two grantees—Carnegie Learning and Off2Class—had values lower than three hours per month of additional learning (Exhibit 8). This range of values (0.7 to 13) was smaller than the range reported among four 2023–24 CEA grantees, which required 2 to 34 hours per month of additional learning (Accelerate 2025).

Analytic insight: The importance of the comparison condition

The impact of tutoring programs is not the impact of tutoring in isolation—but the impact of tutoring compared with what non-tutored students receive.

A sixth grantee included in Exhibit 8, Tennessee SCORE, had a marginally significant effect on achievement. At first glance, the number of tutoring hours associated with Tennessee SCORE's impact estimate might seem notably high (that is, the program may seem inefficient). But the impact being measured for Tennessee SCORE's efficiency index is not the effect of the amount of tutoring, because both treatment and comparison students received equal hours of intervention support using the same tutoring approach but different curriculum materials. In other words, this impact estimate demonstrates the effect of a type of tutoring curriculum rather than the impact of tutoring itself. As a result, dosage does not serve a particularly meaningful explanatory purpose, and the efficiency index is not a helpful measure. This example shows the importance of the comparison condition when interpreting the impacts and efficiency measures of tutoring programs. As noted in Section I, Accelerate has worked with the Youth Policy Lab at the University of Michigan to refine a helpful teacher survey as a tool for collecting information on the comparison condition. This survey is in use in four ongoing evaluations under the 2025 Evidence for Impact (EFI) grantee cohort.

[17] We considered impact estimates with a p-value less than 0.05 to be statistically significant. The term marginally significant refers to estimates with a p-value less than 0.10 but greater than or equal to 0.05.

Exhibit 8. Efficiency of tutoring programs with positive, statistically significant impacts

Tutoring provider	Subject	Grades served	Mode	Learning improvement (SDs)	Learning improvement (months of extra learning)[a]	Average tutoring dosage (hours)	Hours of tutoring required for one month of extra learning (efficiency index)	Study design	Instruction or supports received by students in comparison group
Air Reading	Literacy	K-6	Virtual	0.29*	3.0	27	9.0	RCT	Students received business-as-usual literacy instruction, with the possibility of receiving other supports such as small-group instruction, after-school programs, or interventions provided by classroom teachers.
Carnegie Learning	Math	6–8	Virtual	0.14**	3.8	10	2.7	Matched comparison	Students received standard classroom math instruction supported by MATHia, but without the supplemental tutoring.
Literacy Mid-South[b]	Literacy	2–5	In-person	0.20*	3.1	32	10.4	RCT	Students received business-as-usual instruction during the intervention period, consisting primarily of teacher-facilitated, computer-assisted instruction. Notably, these students were not in the same physical space as those participating in the tutoring intervention.
North Carolina Education Corps	Literacy	K-3	In-person	0.18*	1.5	21	13.8	RCT	Students received business-as-usual literacy instruction and supports provided by the school.
Off2Class	Literacy	6–12	Virtual	0.35*	14.7[c]	10	0.7	RCT	Students used the Off2Class literacy platform for English learners, with engagement facilitated by their classroom teacher rather than a tutor.
Tennessee SCORE	Literacy	K-3	In-person	0.12†	1.3	64	N/A[d]	RCT	Students received intervention with an alternative curriculum that was not aligned with core classroom instruction.

[a] This measure was based on a simple grade-level average of annual expected growth (see Hill et al. 2008) since we wanted to standardize the approach to calculating tutoring efficiency across all studies, some of which did not report grade-level sample sizes.

[b] The learning improvement of 0.20 SDs for Literacy Mid-South is based on low-achieving students in the evaluation sample. The tutoring program typically focuses on these students but, due to recruitment challenges, expanded the pool of students for the study to include higher performers. We used 0.20 SDs as the value for learning improvement because the low-achieving students in the sample aligned more closely with the program's typical eligibility requirements for students.

[c] This evaluation used the WIDA ACCESS, an assessment of English language proficiency; although the exam measures different skills from state ELA assessments, the expected growth in scores from year to year, in standard deviation terms, is similar to expected growth on the standardized assessments documented in Hill et al. (2008).

[d] The efficiency index for Tennessee SCORE is not presented because the grantee didn't measure the effect of a contrast in amount of tutoring as the other grantees did. For more information, see the Analytic Insight box on pg. 20.

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

The 2024–25 CEA grantees had much higher efficiency than a larger set of programs examined by Kohlmoos and Steinberg (2024). Kohlmoos and Steinberg found an average tutoring efficiency index value of 25.9 hours per month of additional learning, whereas all of the efficiency index values in Exhibit 8 are below 15. This suggests that the tutoring programs funded through the 2024–25 CEA found more efficient ways to support student learning.[18] This finding also underscores the value in accurately measuring the hours of tutoring students receive (the actual dosage) and the impacts of tutoring on their learning.[19]

Kohlmoos and Steinberg (2024) had previously found that math programs tend to require fewer hours to generate a month of learning than literacy programs, and the very low index value for Carnegie Learning (which focuses on math) follows the same pattern. However, the literacy-focused Off2Class intervention had a similar-sized value, driven by its combination of a large effect size and a low average dosage. This suggests that English language acquisition among older English learners might be a meaningfully different learning process than building foundational literacy skills among younger learners.

Analytic insight: Placing learning improvements in context

Given the importance of context when interpreting the learning improvements and tutoring efficiency associated with different programs, this section provides examples of conclusions that help place evaluations' key findings in context. The following simple, brief statements describe what the evaluations revealed and highlight key factors that a decision maker should remember when using the findings:

- Students in 1st to 4th grade who received 27 hours of Air Reading virtual tutoring gained an extra 3 months of learning, on average, compared with peers who received typical literacy supports during their intervention block.
- Among 2nd to 5th graders near the national average in literacy skills, receiving 32 hours of tutoring from Literacy Mid-South led to 3.1 months of additional learning, on average, relative to peers who received business-as-usual learning supports during their literacy block.
- English learners in 6th to 12th grade who received the Off2Class literacy curriculum during 10 hours of virtual tutoring in groups of three achieved an estimated 14.7 additional months of learning, relative to peers whose teachers delivered the Off2Class curriculum without the help of tutors.
- Middle-schoolers in grades 6 to 8 who received 10 hours of virtual math tutoring from Carnegie Learning that was tailored to student progress on the MATHia learning platform achieved nearly four months more learning, on average, than similar students who used the MATHia platform but did not receive tutoring.
- Students in Kindergarten through 3rd grade who received Tennessee Score's literacy tutoring using the same curriculum as their core classroom teacher gained an additional 1.3 months of learning, relative to their peers who received a similar amount of tutoring (64 hours) but with a different curriculum. However, this measurement of additional learning did not have enough precision to rule out differences between the two groups being due to chance.

[18] This difference could also be driven in part by changes in how tutoring dosage is measured over time. Because the reports that Kohlmoos and Steinberg (2024) drew on did not uniformly report actual dosage received, the authors calculated efficiency index values using intended dosage, which tends to yield larger efficiency index values. (The average dosage in Exhibit 8 is 20 hours (excluding Tennessee SCORE), but the average dosage in Kohlmoos and Steinberg was 56 hours.)

[19] Measuring impacts precisely may be especially important for programs tutoring older students; as the amount of learning expected in a year (in standard deviation terms) is smaller for older students than younger ones, imprecise impact measures can lead to very imprecise measures of tutoring efficiency.

Accelerate's strategic evolution: Empowering decision makers through standard measures of tutoring efficiency

Accelerate's efforts to gather uniform, accurate information on the amount of tutoring students received led to standardized measures of tutoring efficiency across the 2024–25 CEA cohort. These standardized measures provide valuable information for decision makers seeking to maximize their investments in tutoring, and this practice highlights the value of funders requiring grantees to collect standardized data on tutoring implementation. To continue obtaining this information, Accelerate has required all tutoring grantees for the 2025–26 school year to provide a uniform data standard on the amount of tutoring students received. At the same time, Accelerate's more recent investments in AI-driven personalized learning tools as part of the CET grant highlight the need for dosage measures that can be adapted for that context, as tutor hours are not the primary driver of cost for those types of tools.

C) Program design features of grantees

Key takeaways

- Findings suggest that alignment of tutoring curricula with core classroom instructional materials can lead to greater improvement in student learning.
- Tutoring programs continue to use a variety of tutor types, with one grantee finding that students tutored by paraprofessionals without master's degrees showed larger learning gains than those working with more credentialed tutors, and two edtech grantees not using human tutors at all.
- Grantees collect many types of data for real-time monitoring and share these data with different stakeholders—including district or school leaders, educators, or tutors—to achieve a variety of goals, such as improving tutor instruction, ensuring students receive adequate dosage, supporting midcourse adjustments to instructional content and pacing, understanding school-level buy-in, and addressing implementation barriers.
- Continued research is needed on the efficacy of edtech and AI-enabled tutoring tools. Findings from three edtech-enabled grantees show variability in their effects on student learning, with one grantee finding a large improvement in learning, another whose effect was moderately sized, and a final grantee that found no effect.

Alignment with core instructional materials

Most of the eight grantees sought to align their materials with school or district practices or materials, with some explicitly aligning their tutoring curriculum with core classroom instructional materials. One grantee, Tennessee SCORE, conducted an RCT that demonstrated greater improvements in student learning when the curriculum was consistent with core instructional materials (0.12 SD, $p < 0.10$). Four grantees said aligning tutoring with core instructional materials is critical, for several reasons:

- **Alignment improves teacher buy-in.** Without alignment with the core curriculum, teachers might struggle to see how tutoring supports their classroom instruction, which can create friction between the tutoring program and teachers' core instructional priorities, resulting in lower dosage (Coursemojo).
- **Alignment supports more consistent practice with classroom content.** One grantee said tutoring aligned with the core curriculum provides students additional practice on the skills and concepts they are learning in their classrooms (Air Reading).
- **Alignment creates a more coherent learning experience.** Two grantees said instructional alignment ensures that tutoring provides students with learning experiences connected to their learning in the classroom. When tutoring reinforces the same terminology and instructional approaches used in the classroom, students experience tutoring as an extension of core instruction rather than as a disconnected intervention (Tennessee SCORE and Off2Class).

Grantees aligned their tutoring with core instructional materials in a variety of ways. For example, Carnegie Learning built a standard syllabus that the tutoring provider adapted to state or district contexts, and used weekly tutoring exit-ticket data to make midcourse adjustments to the tutoring. Coursemojo explicitly tied tutoring to teachers' lesson objectives and HQIM. Literacy Mid-South adapted tutoring content to district-selected literacy materials and used tutor support managers to coordinate tutoring content with district pacing guides, ensuring weekly tutoring lessons reinforced classroom instruction.

Furthermore, two grantees, Literacy Mid-South and Off2Class, noted different levels of alignment with core instructional materials, suggesting that a superficial level or static track might involve simply using the same curricular materials or aligning with state assessments, whereas a deeper level or more dynamic track might involve alignment with day-to-day needs, such as teachers' adaptations to materials and real-time pacing.

Strategic insight: Multiple approaches exist to align tutoring with core classroom curricula, with distinct paths to efficiency and scaling

Different approaches to curriculum alignment require different levels of effort—and thus have different implications for (and associated approaches to) scaling. For example, Tennessee SCORE uses supplemental materials from the Tier I core curriculum as its tutoring curriculum, which does not require customization or development of new materials. In contrast, Coursemojo creates new modules aligned to each standard in a school's curriculum, resulting in a vast library of new curricular resources and benefiting from the use of AI to reduce cost and turnaround time.

From a strategic perspective, it's critical to note that these differing approaches to alignment have logistical implications for scaling. For example, a grantee such as Tennessee SCORE might need to invest in training tutors on a new curriculum when starting up in a new district, but a grantee such as Coursemojo might need to invest in developing new materials when it partners with a district that uses curricula Coursemojo is not yet aligned with.

Tutor type

As reported in Accelerate's 2022–23 CEA synthesis, tutoring providers continue to use various types of tutors, including those without a formal background in education, to effectively meet tutor pipeline demands.

One grantee, Tennessee SCORE, found that students tutored by paraprofessionals without master's degrees had larger learning gains than those working with more credentialed tutors. According to the grantee, this difference might have been the result of paraprofessionals' greater fidelity to scripted, aligned materials. The grantee noted that classroom teachers might choose to deviate from scripts based on professional judgment, whereas paraprofessionals tend to "stick to the script," preserving the rigor and alignment offered by the tutoring materials. Finally, the grantee highlighted workload differences (for example, teachers having to juggle multiple prep periods and broader responsibilities versus paraprofessionals' narrower, repeated small-group delivery), which might support more consistent implementation by paraprofessionals.

Across grantees, findings and reflections were mixed on the benefits and drawbacks of certain tutor types, particularly the benefits of using experienced or certified educators with extensive subject matter expertise compared with nonteacher tutors, who might not have competing professional priorities or a tendency to deviate from scripted materials. One grantee, Carnegie Learning, attributed its strong tutor retention and low substitution to the recruitment of certified, mission-aligned teachers; professional development to provide ongoing training and support; and predictable 10-week tutoring cycles with clear expectations for tutors. Another grantee, OKO Labs, said educators who already had strong math discourse and collaboration practices tended to be most effective, but they also offered a hypothesis that deeper domain knowledge might sometimes lead educators to jump in and correct too quickly, whereas less 'math-heavy' educators might more readily facilitate peer-to-peer discourse.

Two grantees used learning tools that did not involve tutors at all: OKO Labs and Coursemojo. Their findings also speak to the range of potential staffing approaches available to deliver personalized learning, including AI tools used by teachers. We discuss these grantees further in "Use of edtech and AI in tutoring programs".

Real-time monitoring

Grantees conducted real-time monitoring of their tutoring to (1) improve tutors' instruction, (2) monitor students' attendance and progress, and (3) support continual improvement of the program. Depending on the goal, grantees used different types of data for monitoring, including the results of tutoring exit tickets, students' tutoring attendance, changes in student learning, tutoring platform data (e.g., system-generated usage data such as time spent on a task, activity completion, and session participation), conversations with school leaders, and observations of tutoring sessions. These types of data might be shared with stakeholders, including district or school leaders, educators, or tutors. Exhibit 9 provides examples of the types of data used for monitoring, their use cases, and their aligned goals. Tutoring providers, tutors, and school leaders can benefit from clear communication about the types of data to be shared, how frequently, and for what purposes to support transparency and collaborative decision making.

Exhibit 9. Types of data used for monitoring, their use cases, and their aligned goals

Data type	Use cases	Aligned goal
Student attendance	Regularly share and communicate with school leaders who can take action on attendance pain points	Ensure students receive adequate dosage
Formative student performance data (e.g., exit-ticket results)	Regularly share with tutors and school or district leaders to monitor program's effect on student learning	Support real-time, midcourse adjustments to instructional content and pacing or reteach content, if necessary
Platform data	Regularly share with tutors and school or district leaders to monitor program usage	Depends on data type—might ensure students receive adequate dosage or support real-time decision making to improve student outcomes
Conversations with principals and educators	Use to support strategic decision making, buy-in, and relationship building	Understand school-level buy-in and address logistical hurdles with implementation
Tutoring observation data	Use to inform tutor training and professional development to support improvement	Improve tutors' instruction

Use of edtech and AI in tutoring programs

As schools and districts consider various edtech and AI-enabled tutoring tools, high-quality evidence on their efficacy remains limited. Two grantees highlighted in this report—Coursemojo and OKO Labs—are AI-driven learning tools in which a platform either acts as a tutor (OKO Labs) or as an assistant teacher to provide immediate individualized feedback (Coursemojo). A third grantee, Off2Class, offers a learning platform that can be utilized by teachers as part of their classroom instruction or by tutors in online sessions. Findings from these grantees show variability in their effects on student learning, with one grantee reporting a large improvement in student learning (Off2Class) and the other two grantees reporting no effect (OKO Labs and Coursemojo). Collectively, these tools provide insights on the potential benefits and challenges of integrating edtech tools into classrooms, while demonstrating the need for continued research on edtech and AI-enabled tutoring solutions.

Accelerate's strategic evolution: Continual improvement practices of edtech grantees influenced Accelerate's investment decisions

Findings from the full set of 2024–25 CEA grantees were key to shaping Accelerate's subsequent research strategy and investment decisions for its EFI and CET grant programs. For example, projects with OKO Labs and Coursemojo generated lessons about continual improvement processes, such as product market research, design and implementation sprints, and real-time product adjustments and additions in response to user feedback. Insights from these grants directly informed the design of the CET initiative, which supports the development, implementation, and evaluation of promising AI-powered and edtech tools.

OKO Labs and Coursemojo found that the teachers using their AI-driven products needed additional guidance on when and how to deploy the tools to achieve intended student dosage. Coursemojo noted that they had implemented more direct teacher training in response to its evaluation findings to support teachers in integrating its tool into their lessons. At the same time, they also adopted more explicit communication about the expected amount of time in which teachers should use Coursemojo. To support higher student usage in future implementation, OKO Labs also expected to communicate with building leaders about dosage expectations and how to schedule time to use OKO Labs. Further, they planned to recommend that teachers guide students to use OKO Labs in pairs and on familiar learning standards before moving to larger student groups and standards they had not yet taught with OKO Labs support.

Off2Class developed evidence of promise in multiple use cases for its tool. Specifically, they found that students who received tutoring through the Off2Class platform showed additional learning gains. In addition, the small sample of control-group students in middle school who received instruction from their core teacher using Off2Class also showed learning improvement. The control group gains were 1.6 times higher than gains for similar students in a recent study (Poole and Sahakyan 2024) that evaluated the progress of English learners using typical methods, suggesting Off2Class's effectiveness in improving student learning using either modality (teachers in classrooms or tutors in online sessions). The distinct experiences and findings of these three grantees underscore the value of continued investment in building evidence on how to integrate new technologies into classrooms and how different applications of the same technology can support student learning.

Accelerate's strategic evolution: Refinements to research questions about AI-enabled tutoring models

Early findings and implementation insights from these grants prompted refinements to research questions about AI-enabled tutoring models. These refinements are reflected in Accelerate's 2025 research agenda, which prioritizes generating new evidence on several little-understood topics: the impact of AI-enabled tools on tutor performance and student learning (including differential impacts on particular student groups); whether AI supports more personalized delivery of HQIM compared with non-AI-enabled provider models; whether AI can reduce implementation barriers that affect student participation and fidelity in tutoring programs; the potential of AI to improve tutor coaching and instructional feedback loops; and how the efficiency of AI-enabled tutoring compares with human-led models.

IV
Lessons Learned and Research Gaps

This section synthesizes and elevates lessons learned from the eight grantees regarding the scaling of tutoring programs, how to achieve adequate dosage, and unresolved research gaps.

A) Lessons learned on scaling up tutoring programs

The eight 2024-25 CEA grantees highlighted in this report cited similar factors that facilitated or impeded scaling, as articulated in Accelerate's 2023-24 CEA synthesis.

Facilitators of scaling. Across providers, three common types of scaling facilitators emerged: high-quality reusable content and curricula, operational and process efficiency, and strong district partnerships and integration into systems.

- **High-quality reusable content and curricula.** Grantees with structured and reusable curricula can scale more easily because instructional quality becomes less dependent on individual tutors. Examples from grantees include the following:
  - Standardized lessons, assessments, and rounds (such as 12-week periods of tutoring) enable replicable implementation across many districts (Carnegie Learning).
  - Searchable repositories of curriculum content, coherence maps (making it easier for educators to navigate and select content within the zone of proximal development for each group of students), and semantic search help teachers quickly find targeted material (OKO Labs).
  - Ready-made materials and standardized training, which is particularly important if a district itself lacks strong curricula (North Carolina Education Corps).

- **Operational and process efficiency.** High-functioning operations—especially around technology, workflows, and internal systems—are key to scaling. Examples from grantees include the following:
  - Clear implementation processes and tutoring platform improvements are essential (Air Reading).
  - Efficient workflows (which can be further powered by AI and paired with modular HQIM-aligned content) reduce the time spent developing tutoring content and curricula (Coursemojo).
  - Improvements in the stability of the technology required for tutoring (such as audio/video checks and usability upgrades) to streamline delivery (OKO Labs).
  - Straightforward teacher workflows and curriculum updates make adoption easier (Off2Class).

- **Strong district partnerships and integration into systems.** Grantees stressed that scaling is not just adding tutors—tutoring requires supporting systems. Scaling is smoother when tutoring aligns with district structures, leadership, and multi-tiered systems of support. Examples from grantees include the following:
  - Embedding a tutoring coordinator in the district who offers monthly professional learning, engaging principals as program supporters, and creating a cohesive tutoring experience that ensures multiple systems work together rather than in isolation (Tennessee SCORE).
  - Emphasizing district leadership support and integration into existing multi-tiered systems of support (Off2Class).
  - Enabling clear communication across district-school levels and explicitly defining the roles of educators, tutors, and school leaders, which can support implementation fidelity (North Carolina Education Corps).

Barriers to scaling. The interviewed grantees identified three types of barriers to scaling that persist from prior CEA cohorts: staffing constraints, district variability and communication gaps, and implementation challenges that create a cognitive burden for teachers.

- **Staffing constraints.** Whether tutoring models use only certified teachers or offer more flexibility on credentials, identifying enough tutors can make scaling difficult. One grantee said its reliance on certified educators as tutors limited its supply of tutors when recruitment pipelines weaken. Another grantee that allows a broader range of tutors said the variability in tutor expertise required strong tutoring materials and tutor training. See Accelerate's 2022–23 CEA synthesis for more information on the implications of flexible tutor types for the tutor pipeline.
- **District variability and communication gaps.** Differences in district structures, curriculum quality, and leadership capacity can slow or complicate expansion. One grantee identified communication gaps between districts and schools as a major barrier. Another grantee said English language development structures vary widely, with some schools requiring more complex tutoring models because they lacked dedicated blocks for English language development.
- **Implementation challenges affecting teacher buy-in.** Particularly relevant for edtech grantees, any challenges that create a cognitive load for educators, such as complicated workflows, technical burdens, and repeated curriculum updates, can affect teacher buy-in. Factors that affect the usability of these tools, such as technology issues or difficulty finding curricular materials that meet educators' needs, can also hinder scaling. See Accelerate's 2023–24 CEA synthesis for more information about the impact of teacher buy-in on effective tutoring implementation.

B) Lessons learned for achieving adequate dosage

The 2022–23 and 2023–24 CEA syntheses revealed a strong link between delivering an adequate amount of tutoring and achieving measurable impacts on student learning. In the 2024–25 cohort, we found that for the two programs that did not achieve statistically significant learning improvement (Coursemojo and OKO Labs), students received a far smaller dosage than they were scheduled to receive. These findings show that technology-enabled programs also face—and must rise to meet—the challenge of achieving adequate dosage to effectively improve student learning. To do so, Coursemojo and OKO Labs identified three factors that contributed to low dosage and the strategies they took (or plan to take) to address them (Exhibit 10).

Exhibit 10. Factors contributing to low dosage in edtech tutoring programs and responsive strategies to address them

Factors contributing to low dosage	Responsive strategies
Lack of teacher buy-in (e.g., perceptions that the tool was "complex" or "extra" rather than an integrated part of core instruction)	• Strengthen implementation supports (beyond basic product training) to help schools embed small-group pedagogy and routines • Redesign activities to more clearly align with daily lessons • Provide starter lesson plans to make weekly integration more routine
Unclear expectations of teachers (e.g., not aware of minimum thresholds for student usage of tutoring program)	• Clarify expectations for minimum usage of tutoring program • Develop dashboards to more systematically monitor and communicate about usage • Schedule platform usage in partnership with school leaders
Lesson-pacing constraints (particularly when students required multiple rounds of feedback to demonstrate mastery of the material)	• Add tools that give teachers greater control over tutoring session pacing (e.g., adding "skip buttons" or other navigation tools within the tutoring platform) to align activities with classroom curriculum

C) Remaining research gaps

Findings from the 2024–25 CEA evaluations highlight several important research gaps that necessitate continued data collection to inform decisions about tutoring design, implementation, and sustainability.

Cost analyses. Cost analyses remain an area of high priority for Accelerate, one that warrants continued development. Although several of the 2024–25 grantees discussed program costs in their report or interviews, only four of the 16 grantees 2024–25 CEA cohort conducted cost analyses detailing the specific inputs and associated input costs necessary to implement their programs. More research is needed to better understand the true per-student costs of different tutoring and personalized learning models, cost variances, and cost-effectiveness (or the return on investment).

Measurement of the comparison condition. Although most grantees offered a vague or general description of the comparison condition, few provided detailed information on the content, frequency, and intensity of instruction or support received by comparison-group students. More research is needed to document the type and intensity of instruction or support received by nonparticipants during the tutoring period to more precisely measure the achievement contrast between tutored and non-tutored peers.

Effect of lower-cost tutoring or personalized learning interventions, including AI-based platforms or short-burst, targeted tutoring. Findings from Tennessee SCORE suggest that alignment with existing district-purchased curricula might offer a lower-cost design pathway, reducing reliance on separate intervention materials and enabling paraprofessionals to deliver tutoring with high fidelity through structured, scripted instruction. More research is needed to assess the cost-effectiveness and scalability of low-cost, curriculum-aligned tutoring models in different district contexts. Findings from OKO Labs suggest that AI-enabled analysis of student discourse and collaboration might support highly targeted, lower-cost instructional approaches that do not rely on dedicated tutors. Again, more research is needed to understand how such models can be implemented consistently, how impacts should be measured for targeted skill development, and whether AI-based platforms can produce meaningful learning gains at scale.

Photo: OKO Labs

V
Moving the Field Forward

Accelerate has drawn on findings from its three CEA cohorts to set priorities for the 2025–26 school year and beyond. These priorities are reflected in Accelerate's 2025 Research Agenda, its request for proposals for the EFI and CET grant opportunities, and its recently awarded Education Innovation and Research grant from the U.S. Department of Education to evaluate virtual and in-person tutoring at scale in Oklahoma. These research activities complement Accelerate's tools and support for policymakers, such as its Tutoring Field Guide (released March 2024), and its technical assistance and networking resources for tutoring providers, such as its Community of Practice.

Accelerate's ongoing and future research investments are focused on the following high-priority needs identified through its past CEA cohorts:

1. **Addressing evidence gaps on understudied student groups, grade levels, and subjects.** Accelerate-funded studies have revealed impacts among tutoring programs focused on specific learner profiles, such as adolescent English learners (Off2Class) and students with dyslexia (Reading Futures). There continues to be an urgent need to learn about tutoring approaches that support specific learner profiles, including older students and multilingual learners (MLLs), since the bulk of the evidence on tutoring centers on early grades and literacy instruction among non-MLL students (Kraft et al. 2024). To address these evidence gaps, Accelerate's 2025 Research Agenda prioritized grants to providers serving students in higher grades, especially 6 to 12; those working in math as well as literacy; and models designed for diverse learners, such as students with IEPs and multilingual learners.

2. **Elevating the role of cost analysis in program evaluations.** Considering education leaders' need to maximize the impact of finite budgets, evaluations are particularly helpful when they provide information on the cost-effectiveness of tutoring programs. Accelerate developed a cost-efficiency measure and cost analysis tool to support the collection and analysis of accurate cost data, and they are partnered with multiple grantees to pilot, refine, and update this tool.

   Beginning in the 2025–26 school year, all Accelerate grantees will produce estimates using the cost analysis tool to help provide critical information on cost-effectiveness. At the same time, Accelerate is investing in studies of lower-cost tutoring options such as short-burst tutoring, in which tutors work one-on-one with students in 5- to 10-minute sessions focused on specific foundational literacy skills. Accelerate has also supported the development of statewide outcomes-based contracting (OBC), which ties tutoring funding to measurable student results. Through its States Leading Recovery program, Accelerate has partnered with 12 states to build the infrastructure for sustainable, high-dosage tutoring at scale—with several states, including Arkansas and Texas, now piloting OBC frameworks that reward providers for achieving student outcomes rather than simply delivering services. Investing in lower-cost models, rigorous cost measurement, and outcomes-based accountability can help identify and scale the most cost-efficient programs.

3. **Building evidence on virtual and in-person tutoring at scale.** Accelerate continues to seek strategies for boosting the amount of tutoring students receive at scale. The cost per student represents an important constraint on efforts to deliver tutoring at scale, and two key drivers of cost are the number of students in each group a tutor serves and the number of tutoring sessions per week.

In partnership with the Oklahoma State Department of Education and Mathematica, Accelerate was recently awarded an Education Innovation and Research (EIR) grant to conduct a study measuring how tutoring program design features—virtual versus in-person, group size, and dosage—affect both learning outcomes and cost efficiency. The central goal is to examine whether variations in these design features yield different returns on investment and to develop guidance on ways to improve program cost-effectiveness and long-term sustainability. The study will include approximately 6,000 students in grades K-5.

4. **Expanding the evidence base on AI solutions.** Accelerate invested in two AI-driven personalized learning tools in the 2024–25 CEA cohort, Coursemojo and OKO Labs. Evaluation findings indicated that although the tools show promise, key questions remained about the support teachers need to effectively integrate them into their classrooms. Accelerate is continuing to inform education leaders about emerging technology for personalized learning through 10 CET grants during the 2025–26 school year. These grants will fill gaps in the evidence base around the usability, usefulness, and promise of AI tools for delivering effective personalized learning support. The early-stage evaluations of 10 AI- and tech-enabled tools will also begin to produce evidence on the cost-effectiveness of using this technology.

To maximize the benefits to students from tutoring and other approaches to personalized learning, Accelerate seeks to use the collective efforts above to encourage a culture of generating and using evidence to drive student learning.

Because the tutoring efficiency index and related cost-effectiveness measure are most informative when based on precise, accurate impact estimates, Accelerate continues to fund rigorous studies that provide these estimates. Accelerate's 2025 research grant programs—EFI and CET—prioritize studies that produce reliable impact estimates of innovative approaches to tutoring and personalized learning. For example, the EFI grants aimed to support multi-arm randomized designs and large samples (study samples of at least 200 treated/tutored students per treatment arm, so as to meet ESSA Tier 1 evidence standards) to produce precise impact estimates. Furthermore, all 2025-26 EFI grantees are required to use Accelerate's teacher survey to gather accurate information on the comparison condition, which is a critical input for producing useful measures of tutoring efficiency. Where possible, the EFI grants will use multiarmed randomized trials to directly measure the difference in impacts from raising the number of tutoring hours or modifying other key program features. The CET grants offer providers an incentive to use a study design that meets the U.S. Department of Education's standards for high-quality ESSA Tier 2 evidence.

Ultimately, it is critical to provide education decision makers with the highest-quality evidence possible about which tools benefit students most. This evidence can help create powerful incentives for service providers to continue researching and improving the effectiveness of tutoring and other personalized learning supports. As this field evolves, Accelerate will continue targeting investments designed to produce the evidence the field needs most to carry on the urgent work of expanding students' learning opportunities.

Photo: Air Reading

VII
References

Accelerate. "Quarterly Research Note." Brief, vol. 1, no. 4, 2025a.
https://accelerate.us/wp-content/uploads/2025/03/1.4-QRN.pdf.

Accelerate. "Quarterly Research Note." Brief, vol. 2, no. 1, 2025b.
https://accelerate.us/wp-content/uploads/2025/07/2.1-QRN_FINAL_070325.pdf.

Conroy, K., K. Place, and G. Chojnacki. "Selecting and Supporting Math Tutoring Programs: Recommendations for School Districts." Washington, DC: Mathematica, 2023.
https://www.mathematica.org/publications/selecting-and-supporting-math-tutoring-programs-recommendations-for-school-districts.

Hill, C., H. Bloom, A. Black, and M. Lipsey. "Empirical Benchmarks for Interpreting Effect Sizes in Research." Child Development Perspectives. 2, 2008, pp. 172 - 177.
https://psycnet.apa.org/record/2009-06147-010

Kohlmoos, L., and M.P. Steinberg. "Contextualizing the Impact of Tutoring on Student Learning: Efficiency, Cost-Effectiveness, and the Known Unknowns." Accelerate, May 2024.
https://accelerate.us/wp-content/uploads/2024/05/Accelerate-Research-Report-Efficiency-and-Cost-Effectiveness-1.pdf.

Kuhfeld, M., and M. Burchinal. "Gender Gaps in the Early Grades: Questioning the Narrative That Schools Are Poorly Suited to Young Boys." EdWorkingPaper 25-1297. Annenberg Institute at Brown University, September 2025.
https://doi.org/10.26300/a14s-d296.

Nickow, A., P. Oreopoulos, and V. Quan. "The Promise of Tutoring for PreK–12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence." American Educational Research Journal, vol. 61, no. 1, 2024, pp. 74–107.

Poole, G.A. and N. Sahakyan. "Examining English learner testing, proficiency, and growth: Continued trends since the COVID-19 pandemic. Research Report No. RR-2024-1. Madison, WI: Center for Applied Linguistics.
https://wida.wisc.edu/sites/default/files/resource/Research-Report-Examining-English-Learner-Testing-Proficiency-Growth-2024.pdf

Accelerate
National Collaborative for Accelerated Learning