Accelerate

# Research Agenda

2026-27

## OVERVIEW

**Accelerate's theory of action is bold: invest catalytic funds to build evidence, align policy and practice, and scale cost-effective interventions across the education ecosystem — so that millions of students, especially those historically underserved, experience measurable gains in academic achievement and life outcomes.**

This research agenda is organized around the areas of focus reflected in Accelerate's Theory of Action. Each section identifies a specific area we are working to advance and the research required to move it:

FOCUS AREA #1
**PROVING WHAT WORKS**
Proving that core educational interventions work and reach the students who need them

FOCUS AREA #2
**SCALING WHAT WORKS**
Understanding what it takes to implement what works, and sustain it at scale

FOCUS AREA #3
**PATHWAYS THAT WORK FOR LIFE**
Reaching the life outcomes our theory of action promises, not academic achievement alone

FOCUS AREA #4
**LEARNING FROM OUR WORK**
Learning across our portfolio to strengthen how we support and carry out our work

The research questions themselves aren't organized around any single grant program — they reflect what we already know, what we've learned through our own work, and where the most important gaps remain. This agenda is intentionally ambitious. Accelerate will not answer every question here on its own; our aim is for the work we fund to move the field toward these answers, alongside other research partners, funders, and agencies working on the same questions.

That scope now extends beyond high-dosage tutoring to the full range of interventions we're working to understand. Since 2022, Accelerate has funded research on 105 tutoring providers, including 45 randomized controlled trials. That evidence base is what makes it possible to pose sharper questions now than the field could pose a few years ago, and what this agenda builds on.

## FOCUS AREA #1: PROVING WHAT WORKS

For high-dosage tutoring, the evidence is well established. Nickow and colleagues' 2024 synthesis of 89 randomized studies estimated an average effect of 0.29 standard deviations; Kraft and colleagues' synthesis of 265 studies estimated 0.42. Our own portfolio of 45 randomized trials across 105 funded providers adds evidence from studies we publish regardless of the result: positive, null, or negative. That gives a less filtered view of what tutoring achieves in practice than the published literature alone. Given that base, our focus is shifting: rather than asking whether high-dosage tutoring works, we are refining how it works best — for which students, in what form, over what time horizon, and at what cost. Where it matters most, we will prioritize fewer, better-powered studies that can answer these questions over a larger number of studies that cannot.

For high-quality instructional materials and AI- and technology-enabled tools, the evidence is far less developed. Both are still at the stage of establishing basic efficacy: whether these interventions improve outcomes at all, and under what conditions. This section holds all three together because they represent different points on the same continuum, from refining a proven intervention to establishing a new one, and the same evidentiary bar applies regardless of where an intervention sits on that continuum.

Within the next five to ten years, we aim for the field to know as much about what works in instructional materials and technology-enabled learning as we now know about high-dosage tutoring, and for our investments to have contributed meaningfully to closing today's evidence gaps.

| Pillar | Research Question | We're Focused On |
|---|---|---|
| **High-Dosage Tutoring** | Does high-dosage tutoring's impact hold across grade levels, student subgroups, and over time? | • Replication of effects in middle and high school, where the evidence base is thinnest<br>• Differential effects by subgroup, including English Language Learners, race/ethnicity, gender, and justice-involved youth; prioritizing studies designed and powered to test these differences directly rather than reporting them only when they happen to appear<br>• Durability of gains a year or more after a program ends<br>• Which program design features — tutor type, group size, delivery modality — are associated with stronger effects |
| **Out-of-School Tutoring** | Does out-of-school tutoring, such as after-school programs or home-based tutoring, achieve dosage and outcomes comparable to in-school delivery, and which specific models work best? | • Whether out-of-school tutoring can achieve adequate dosage outside the school day, and what design features make that more likely<br>• How out-of-school outcomes compare to in-school delivery, and which specific models — after-school programs, home-based or virtual tutoring — perform best, a question made more urgent by new scholarship and tax-credit funding streams that expand what families can spend tutoring dollars on<br>• Whether effects differ by student group and age, and which models best reach and retain historically underserved students |

| | | |
|---|---|---|
| **High-Quality Instructional Materials (HQIM)** | Does a high-quality instructional curriculum improve student outcomes, and what enabling and supporting conditions are necessary for that improvement to occur in practice? | • Which materials, including digital and AI-enabled tools, teaching practices, and implementation decisions strengthen classroom instruction and, in turn, improve student outcomes<br>• What enabling and supporting conditions (such as coaching and ongoing professional learning) determine whether high-fidelity adoption succeeds or stalls<br>• Whether using HQIM as the basis for intervention and/or differentiated core instruction leads to better outcomes than other materials<br>• How implementation develops over time, from initial readiness to sustained ownership, and what the trajectory on leading indicators looks like before student outcomes shift |
| **AI & Technology-Enabled Learning** | Which AI-and technology-enabled learning tools are backed by rigorous evidence, rather than by promise alone? | • Which AI-and technology-enabled tools have causal evidence of effectiveness<br>• Which design features of AI- and technology-enabled tools, such as adaptive difficulty, human-in-the-loop oversight, and personalization, show up as cross-cutting patterns linked to stronger effects<br>• How the strongest tools compare to human-delivered tutoring on both effect size and cost<br>• Where AI-and technology-enabled tools may produce unintended negative effects — over-reliance, reduced human interaction, widened equity gaps — alongside their benefits |

## FOCUS AREA #2: SCALING WHAT WORKS

Demonstrating that an intervention is effective in a controlled trial is a necessary but insufficient condition for improving outcomes at scale. Kraft, Schueler, and Falken's 2026 synthesis of 263 tutoring trials found that an average effect of 0.40 standard deviations falls to 0.22 once implementation reaches 400 to 999 students, and to 0.16 at 1,000 students or more. It establishes that impact can erode substantially as a program scales, without implying that every intervention will erode by the same amount or through the same mechanism. We interpret this finding as a motivating mechanism to learn more.

Rather than treating implementation as a secondary concern to be addressed after efficacy is established, we treat it as a distinct research problem, one with identifiable components whose relationship to fidelity, scale, and outcomes can be understood and tested empirically. Because this is newer territory for us than proving efficacy, the questions below are still more granular than elsewhere in this agenda; we expect them to consolidate as the work matures.

This section applies that lens to tutoring, high-quality instructional materials, AI-enabled tools, and any intervention we fund, with a focus on two questions: what it takes for a program to retain its impact as it moves from a single site to a full system, and whether coordinating tutoring with a school's existing system of tiered supports (MTSS) improves how well it is implemented relative to delivering it independently. It also applies to states and districts aiming to improve core instruction, where the question is not whether an external program holds its impact at scale but whether a system builds the internal capacity to deepen, spread, and sustain strong practice. Over time, we also want to understand how strong core instruction, assessment and data, and supplemental supports work together as a coherent system. Our goal over the next five to ten years is to understand precisely what it takes to implement and sustain a program well as it scales across an entire system so that impact doesn't erode as programs grow.

| Pillar | Research Question | We're Focused On |
|---|---|---|
| **Implementation & Scale** | Which implementation factors are necessary for a program, or a district's own core instruction, to retain its impact at scale? | • How staff selection, training, coaching — including which components, delivered by whom, and at what dosage — and performance monitoring need to work to preserve fidelity at scale<br>• Which organizational conditions — data systems, administrative policies, accountability structures, and dedicated time for coaching and collaborative planning — and which surrounding policy, funding, and political conditions enable a program to scale, including whether these differ between state-level (SEA) and district-level (LEA) implementation<br>• What district and school leaders do to support implementation at scale, including removing barriers, adapting as challenges arise, and building bench strength so that progress survives leadership turnover<br>• Whether dedicated implementation teams and structured improvement cycles change outcomes relative to programs without them<br>• How implementation progresses from readiness through full implementation and sustainability, how long each stage takes, whether fidelity gains persist or erode once formal support ends, and whether that support can be deliberately tapered as internal capacity grows rather than withdrawn all at once<br>• How much local adaptation to a program's design is compatible with strong outcomes, and at what point adaptation becomes erosion of the elements that actually matter<br>• Whether these implementation factors must be in place together, or whether gaps in one can be offset by strength in another, and how together they relate to fidelity, scale, and student outcomes — including whether gains are distributed evenly across schools and student groups, or concentrated among those already best served |
| **Tutoring's Role Within Multi-Tiered Systems of Support (MTSS)** | Does embedding tutoring within a school's MTSS structure improve implementation quality compared to delivering it as a standalone program? | • Whether tutoring delivered within a coordinated MTSS structure achieves higher fidelity and more consistent dosage than standalone delivery<br>• What operational conditions — shared data systems, joint scheduling, coordinated caseload management — are necessary for tutoring to be implemented well within MTSS<br>• Whether embedding tutoring within MTSS improves coherence with a student's other tiered supports, rather than duplicating or crowding them out |

## FOCUS AREA #3: PATHWAYS THAT WORK FOR LIFE

Career-connected learning has drawn growing attention from states, districts, and philanthropy as a strategy for improving students' long-term outcomes, not only their academic ones. That interest has outpaced the evidence base: rigorous, well-powered studies of career-connected learning models — career academies, work-based learning, apprenticeships, credentials of value — remain scarce, and much of what does exist predates the labor market students encounter today.

Our theory of action's own definition of impact names life outcomes alongside academic achievement, yet none of our current outcome metrics measure the former. This section begins to close that gap. It is deliberately narrower than the other two: rather than spreading thin across every possible life-outcomes question, it concentrates on career-connected learning, an area where our own SIF-funded work on credentials of value already gives us a foothold to build rigorous evidence on rather than starting from nothing. Within the next five to ten years, our goal is to build meaningfully more rigorous evidence on whether and how career-connected learning improves students' long-term life outcomes.

| Pillar | Research Question | We're Focused On |
|---|---|---|
| **Career-Connected Learning** | Do career-connected learning models improve students' long-term life outcomes? | • Among current SIF-funded grantees focused on credentials of value, which near-term implementation and take-up measures like credential completion are most related to leading indicators and most predictive of downstream employment outcomes<br>• Whether credentials of value and career and technical education (CTE) pathways translate into higher employment and earnings, not just credential completion<br>• Which specific model — career academies, apprenticeships, work-based learning, dual enrollment — produces the strongest and most durable results |

## FOCUS AREA #4: LEARNING FROM OUR WORK

Focus Areas 1 through 3 are organized around what we're trying to prove about a specific intervention or outcome. But some of what we learn doesn't come from any single study. Instead, it comes from comparing many different funded efforts to each other and noticing what holds true across them — sometimes within one kind of intervention, sometimes across entirely different kinds of interventions.

| Pillar | Research Question | We're Focused On |
|---|---|---|
| **Patterns Across Our Work** | Which patterns show up again and again across our work, and what do they tell us about what drives strong outcomes? | • Which practical features and implementation conditions — including how technology is used — are consistently linked to strong outcomes, whether we're looking within one kind of intervention or across very different ones<br>• Which student populations and school contexts see consistently strong or weak results, no matter which intervention is involved<br>• Which of those associations reflect true causal mechanisms versus features that are merely correlated with success<br>• Whether the gains achieved through initiatives that we fund last after an intervention ends |