



# High-Impact Tutoring: State of the Research and Priorities for Future Learning

Carly D. Robinson & Susanna Loeb

## EXECUTIVE SUMMARY

**Research consistently demonstrates that tutoring interventions have substantial positive effects on student learning.** As a result, tutoring has emerged as a promising strategy for addressing COVID-related learning loss and affording greater educational opportunities for students living in poverty. The effectiveness of tutoring programs, however, varies greatly, and these variations may drive differential gains in student learning. Therefore, determining the program characteristics that do and do not drive positive student outcomes will be key to providing guidance for policymakers and practitioners who want to implement high-impact tutoring at scale.

The goal of this document is to highlight the programs, characteristics, and conditions that evidence suggests make for effective tutoring and to create an evidence-based framework for delivering and evaluating tutoring interventions. In addition, this document identifies promising questions for future research.

### CONTENTS

Executive Summary	1
Introduction	7
Research Summary	10
A Research Agenda Moving Forward	25

### Research Agenda Acknowledgements

*The authors are deeply appreciative of the thoughtful feedback and support we have received from many people, including Kathy Bendheim, Christina Claiborne, Kristine Gaffaney, Nora Gordon, Jonathan Guryan, Hiep Ho, Brian Jacob, Matt Kraft, Monica Lee, Sam Madison, Paul Manna, Samantha Maroshick, Annie O'Donnell, Phil Oreopoulos, Elisabeth Ostrow, Charlotte Perez, Lavanya Sathyamurthy, Nate Schwartz, and the National Student Support Accelerator Advisory Board and Working Group members. We are also grateful for support from the Bill & Melinda Gates Foundation and the Walton Family Foundation.*

**Suggested Citation:** Robinson, Carly D., and Susanna Loeb. (2021). High-Impact Tutoring: State of the Research and Priorities for Future Learning. (EdWorkingPaper: 21-384). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/qf76-rj21>

**We begin by summarizing the existing academic research on tutoring to identify the characteristics of tutoring programs and how they can vary.** The majority of the characteristics fall under one of the following categories: program focus, safety, equity, cohesion, tutor, instruction, learning integration, and data use. Within each category, tutoring programs may differ significantly from one another, as detailed below under each category name:

#### ■ **Program Focus**

- Subject area;
- Grade level; and
- Students (which students receive the tutoring: can be needs-driven, curriculum-driven, or universal).

#### ■ **Safety, including safety protocols.**

#### ■ **Equity**

- Language (whether bilingual tutors are available);
- Diversity of tutors;
- Cultural competency focus; and
- Social-emotional focus.

#### ■ **Cohesion, including organizational culture.**

#### ■ **Tutor**

- Tutor type (e.g., certified teachers, paraprofessionals, volunteers, university students, private tutors, family members, and peers);
- Tutor recruitment;
- Tutor training; and
- Tutor support and feedback.

#### ■ **Data Use**

- Design and improvement process;
- Formative assessment;
- Student progress;
- Program evaluation; and
- Tutor effectiveness.

#### ■ **Instruction**

- Delivery (online or in-person);
- Dosage (the frequency and length of sessions, as well as the program duration);
- Student-tutor ratio;
- Student grouping (how students are grouped together if tutoring occurs in small groups);
- Tutor consistency;
- Tutor-student relationship;
- Structure;
- Quality of materials; and
- Facilitation (i.e., the quality of instructional teaching).

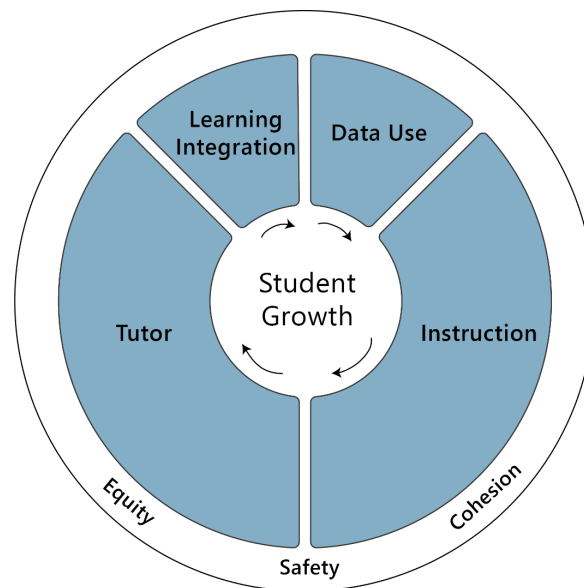
#### ■ **Learning Integration**

- During-school or out-of-school time (OST) setting;
- Displacement (i.e., the counterfactual: what the student would have experienced without the tutoring);
- Take-up (is the tutoring required or do students and/or their parents have to enroll);
- Curricular alignment (is the curriculum aligned with what the student is learning in class);
- School and teacher engagement; and
- Family engagement.

Because the vast majority of research on tutoring has evaluated the effect of individual tutoring programs, we cannot definitively say whether or not a specific program characteristic leads to increases in student learning. We can, however, look at what characteristics are present in high-impact tutoring programs and begin to draw inferences from there. Using this approach, **we recognize that high-impact tutoring programs tend to include the following characteristics:**

- **High-dosage (i.e., three or more sessions per week of required tutoring),**
- **A stated focus on cultivating tutor-student relationships,**
- **Use of formative assessments to monitor student learning,**
- **Alignment with the school curriculum, and**
- **Formalized tutor training and support.**

### Framework for High-Impact Tutoring



*This visualization conceptualizes how the components of a high-impact tutoring program relate to one another and lead to student growth.*

**In the second half of this document, having reviewed and discussed the existing literature on tutoring, we set forth an ambitious research plan in our “Research Agenda Moving Forward” section.** This section identifies two priorities to guide future research. First, we highlight promising avenues for further investigation into what features are necessary for high-impact tutoring and articulate worthwhile research questions — those with answers that could help policymakers and practitioners effectively implement tutoring initiatives.

Second, in addition to identifying the characteristics of effective tutoring, researchers will need to better understand how to scale tutoring programs in order to reach the greatest number of students in need. We have begun to articulate a second line of suggested research questions that broadly address issues such as: Where will the supply of tutors come from? What is the demand for high-impact tutoring? And how are tutoring programs most successfully implemented?

## Research Priority: Identifying the Characteristics of Effective Tutoring

Our first research priority, **Identifying the Characteristics of Effective Tutoring**, addresses several focus areas including:

- The cost-effectiveness of different tutoring models;
- Tutor skills;
- Tutor-student relationships;
- Adoption by districts, schools, families, and students;
- How programs increase engagement and implementation fidelity; and
- How best to extend the research focus.

Below, we provide a brief overview of our “Identifying the Characteristics of Effective” section, contemplating how future studies might evaluate these specific areas of interest.

First, **evaluating the characteristics that drive tutoring program costs will be critical to scaling** so that all students experiencing poverty can have access to high-impact tutoring. With a goal of determining how to deliver tutoring in the most cost-effective way, we propose conducting evaluations that intentionally vary: student-tutor ratio; time students spend with tutors each week; whether tutoring is delivered in-person, virtually, or through a combination of the two (i.e., blended); and the use of computer-assisted learning technologies. Additionally, we need to study the amount of training each tutor type requires to boost student outcomes.

Second, **understanding what skills and supports tutors will need to be successful will contribute to evaluating both the costs and quality of the tutoring**. In addition to assessing tutor type and training, we propose studying what recruitment policies are most effective, what forms of ongoing support are most beneficial for tutors and students, and how training and support can ensure quality and safety.

Third, high-impact tutoring involves human interaction between a tutor and a student. Many programs state that cultivating positive tutor-student relationships is a critical feature. Therefore, **understanding how to build strong relationships and how the relational component of tutoring impacts student learning is important for evaluating and learning to create quality**. Potential opportunities to explore the import of tutor-student relationships could involve: testing the importance of having the same tutor; intentionally embedding relationship-building content in tutoring programs; matching students with tutors based on their traits; and “looping” students and tutors such that students learn from the same tutor over the course of multiple years (not just one).

Fourth, successfully scaling high-impact tutoring requires buy-in and adoption from some combination of districts, schools, educators, parents, and students. Research efforts focused on **understanding how programs facilitate or hinder program adoption can provide necessary information for scaling**. These efforts might explore the impact of: need-driven vs. universal tutoring; opt-in vs. opt-out enrollment; different recruitment strategies; during-school vs. out-of-school time tutoring; and curricular alignment.

Fifth, once tutoring is offered, schools, teachers, parents, and students need to actively engage with the program. For tutoring to be successful, **we need to consider the types of practices that lead to increased**

**engagement among various stakeholders.** We propose testing levers for promoting engagement, including: school-based coordinators; tutor-teacher interaction; and family involvement.

Finally, **there is great opportunity to learn more about the multidimensional benefits of tutoring.** While our agenda focuses on student learning outcomes, high-impact tutoring has the potential to affect a wide range of student outcomes. For instance, in addition to tested outcomes, tutoring could also result in improvements in student grades, motivation, relationships, advanced course-taking, attendance, behaviors, and school completion. By including additional measures when assessing tutoring program characteristics, we can learn how and why tutoring leads to student success. We further need to consider the long-term outcomes associated with tutoring, and if they vary by tutoring program characteristics. Also, tutoring may have benefits for the tutors. Tutoring may lead tutors to complete more school, enter teaching, engage more positively with their community, or improve on a range of other outcomes. Tutoring may also reduce unemployment in the short run and lead to more consistent employment in the longer run.

---

## Research Priority: Implementing Tutoring at Scale

**Our second research priority, *Implementing Tutoring at Scale***, broadly address issues such as: Where will the supply of tutors come from? What is the demand for high-impact tutoring? And how are tutoring programs most successfully implemented?

The goal of “Implementing Tutoring at Scale” section is to identify questions that guide research on how practitioners and policymakers will implement high-impact tutoring.

For instance, to implement tutoring at scale, **we need to understand (1) where the supply of tutors would come from and (2) the extent to which districts want and are capable of implementing tutoring initiatives.** We need answers to questions such as:

- What skills do tutors need?
- When are tutors available?
- How can the tutoring be delivered?
- What types of supervision do tutors need?
- Who can be tutors?
- What is the best way to recruit tutors?
- How many students can a tutor handle?
- How much do tutors cost?
- Where is the money for tutoring coming from?
- What are the constituencies that might want high dosage tutoring?
- What stakeholders at different levels need to be brought in to make high-dosage tutoring happen?
- How does state or district governance shape demand for tutoring?
- Which students receive tutoring?

Simultaneously, we have developed a line of implementation-related questions that will guide researchers in building a deeper understanding of how district leaders, school leaders, teachers, parents, and students

experience a tutoring initiative. Each question addresses either the motivations, opportunities, logistical needs, or barriers for delivering tutoring to students. The Answering questions at these multiple stakeholder levels will ultimately help answer the overarching question: **What are the necessary conditions required to effectively implement a high-dosage, school-driven tutoring program?**

These two research priorities cover just a handful of the promising questions for researchers to address regarding high-impact tutoring. By no means is this list comprehensive. We hope educational researchers across disciplines will unite to determine how high-impact tutoring can lead to more equitable outcomes for the students who need it most.

## INTRODUCTION

High-impact tutoring interventions regularly produce positive learning outcomes for a wide range of students. For this reason, tutoring has emerged as a popular strategy for mitigating learning loss in the wake of COVID-19. In this document, we aim to summarize the research on high-impact tutoring, to identify and describe existing tutoring models, and to recommend promising future directions for research and practice.

First, we provide an overview of the existing academic research on tutoring and describe ways in which tutoring programs can differ. We highlight the characteristics and conditions that evidence suggests make for effective tutoring programs. Then, we identify promising questions for future research, the answers to which could benefit decision makers in choosing effective tutoring approaches.

### What is High-Impact Tutoring?

Before delving into the research, we first define what we mean by “high-impact tutoring.” Tutoring is a form of teaching, one-on-one or in a small group, towards a specific goal. High-impact tutoring leads to substantial learning gains for students by supplementing (but not replacing) students’ classroom experiences. High-impact tutoring responds to individual needs and complements students’ existing curriculum. Other “personalized learning” options exist and may promote student learning by replacing traditional class periods, but we do not include them under our umbrella definition of tutoring. For example, at this point we are not focusing on initiatives such as: pull-out services (e.g., when students receive personalized help *instead* of attending a class), in-class small group instruction by a second teacher (e.g., co-teaching), or learning pods.

We recognize high-impact tutoring programs as those that either have directly demonstrated significant gains in student learning through state-of-the-art research studies or have characteristics that have proven to accelerate student learning. These characteristics of high-impact tutoring programs currently include: substantial time each week spent in required tutoring; sustained and strong relationships between students and their tutors; close monitoring of student knowledge and skills; alignment with school curriculum; and oversight of tutors to assure quality interactions.

### The Tutoring Landscape: Past and Present

The landscape of tutoring programs is vast and varied. While the research evidence shows that tutoring interventions can have positive impacts on student learning, past efforts to scale tutoring programs have not always been successful. To put the concept of high-impact tutoring into context, we briefly discuss several of the most well-known initiatives and programs in the past few decades. In the following paragraphs, we introduce four tutoring programs, discuss how they fared, and note the characteristics of the programs that may have influenced whether the tutoring delivered was high-impact.

- 1) The ***No Child Left Behind*** legislation of 2001 allowed schools to use funding to offer tutoring through ***Supplemental Educational Services***. These Supplemental Educational Services referred to free extra academic help, such as tutoring or remedial help, that is provided to students in subjects such as reading, language arts, and math. This initiative allowed low-income families to enroll their child in Supplemental Educational Services if their child attended a Title I school. Parents had to choose the provider and enroll their child in out-of-school time (OST) tutoring. The tutoring could be provided by a large number of diverse organizations with widely varying hourly rates, service costs, tutor qualifications, tutoring session length, instructional strategies, and curricula.

Evaluations of the Supplemental Educational Services programs showed that the flexibility, lack of accountability, and minimal connections with the students' schools resulted in varied quality and mixed results (see Heinrich et al., 2014; Zimmer et al., 2010). The few instances where the Supplemental Educational Services did positively impact student learning tended to involve minimum dosage requirements, structured sessions, tutor coordination with schools, and more tutor experience (Heinrich et al., 2014).

- 2) ***America Reads*** (sometimes referred to as *America Reads\*Counts*) is a federally funded work-study program that places college and university students in the community to tutor elementary through high school students in literacy and math. College tutors can receive federal funding in the form of work-study grants for participating in an approved program. The program grew from the America Reads Challenge, a major 1997 initiative of the Clinton administration that envisioned an "army of tutors in schools teaching students to read." Approximately 1200 institutions of higher learning still participate in America Reads nationwide in 2020.

Early reports on the program suggest that it was very expensive and had mixed results (Worthy et al., 2003). Some small studies suggest there may be positive outcomes for students when tutors receive substantial training (Borges & McCarthy, 2015; Fitzgerald, 2001; Fitzgerald et al., 2002), but there have not been any large-scale rigorous evaluations of the program. Worthy et al. (2003) conclude that coordinating a national tutoring program of volunteers and work-study students requires more than just work-study funds, but also needs start-up and overhead funding; supervision; ongoing training; and motivated tutors.

- 3) ***AmeriCorps*** is an arm of the Corporation for National & Community Service that mobilizes 75,000 individuals who work in non-profits, public agencies, and religious organizations across the United States. AmeriCorps members are generally recruited by and embedded in the staff of their partner organization, although some members complete intensive, residential projects in AmeriCorps teams. Of the 75,000 members, approximately 44,000 serve in the education sector in programs embedded in public schools, tutoring non-profits, or other education-focused agencies.

While the majority of AmeriCorps and City Year school-based volunteers are not focused explicitly on tutoring, there are several successful AmeriCorps tutoring initiatives. These involve intensive training for tutors, a minimum of 1.5-3 hours of one-on-one tutoring per week, subject-focused curricula, and site coordinators at the school:



- *Readings Corps*, one of the largest of the AmeriCorps initiatives, is implemented across 12 states and the District of Columbia. Evaluations of the programs' impact in Minnesota, Wisconsin, and Florida find that it is very effective when implemented with fidelity (Markovitz et al., 2014).
  - *Reading Partners* is a uniquely structured program where AmeriCorps members serve as volunteer coordinators and site leadership at partner schools. Members are responsible for managing pull-out tutoring sessions between volunteers and students in a designated reading center at each school, and evaluations find it improves literacy for young students (Jacob et al., 2015).
  - *Math Corps* (Minnesota) is a community-based intervention that leverages AmeriCorps members as both tutors and recruiters of local community members. A teacher or school staff member is designated to manage the implementation of the program at each site, and an external coach validates the implementation fidelity on a regular basis at each site. The program meaningfully increases math achievement (Parker et al., 2019).
- 4) *Saga Education* is a national nonprofit organization that partners with public school districts to supplement teacher instruction by offering trained tutors for students who are falling behind in math. The tutorial model offered by Saga Education was founded in the Match Charter Schools in Boston, MA where daily 2:1 tutoring sessions are embedded in students' schedules. Saga Education, too, recruits AmeriCorps members as tutors for at least one full year of service.

The Saga Education model has some of the strongest evidence of success among tutoring programs, especially for students in secondary schools (Cook et al., 2015). Additional pilot and non-experimental evaluations find similar positive results, suggesting that these results are not idiosyncratic to the one randomized controlled trial (RCT) (Cook et al., 2015; Fryer Jr, 2014). Saga Education deploys many of the program characteristics that we associate with high-impact tutoring, including: intensive training for tutors; high-dosages; an aligned, subject-focused curriculum, on-going tutor support, and a focus on building tutor-student relationships.

**While there exist many more programs, these four are consistently part of conversations about tutoring and its ability to effectively impact student outcomes and scale.**

## RESEARCH SUMMARY

The research evidence is clear: high-impact tutoring positively affects student learning. Of course, not all tutoring initiatives are high-impact and some programs appear to be more effective than others. Though private instructors and in-home tutoring constitute the oldest form of schooling, there has been an increased interest in the past few decades in deploying cost-effective tutoring to bolster educational outcomes. In this section, we summarize the key takeaways from the academic literature on tutoring with an eye toward identifying the characteristics of effective tutoring interventions.

Because so many tutoring models exist, we clarify the focus of this review and how we are evaluating evidence.

- First, we adopt the criteria that tutoring is one-on-one or small group instruction that supplements but does not replace classroom-based education and that involves a human teacher, though it might also make use of computer-based materials.
- Second, we rely on research that includes a clear comparison group, usually randomly assigned, so that the estimated effects of the tutoring intervention are believably causal.
- Third, the included studies provide strong estimates of the causal impacts of programs. However, they usually compare programs with multiple features to a control of no tutoring, and, as a result, do not tell us the causal impact of individual tutoring features. We draw inferences based on what characteristics tend to be present in more and less successful tutoring interventions, but without research specifically designed to assess the effects of a given feature, we cannot definitively say whether or not a specific characteristic causes better student outcomes.

### Overall Effectiveness of Tutoring Interventions

Two recent meta-analyses found that, on average, tutoring interventions increase student learning outcomes by over one-third of a standard-deviation (Dietrichson et al., 2017; Nickow et al., 2020). The average effects of tutoring interventions are considered large for educational interventions (Kraft, 2020) and translate to between three and fifteen additional months of learning for students (Bloom et al., 2008). These analyses (as well as those of Ritter et al., 2009) focus exclusively on student learning outcomes, and do not speak to whether these tutoring interventions impact other, non-academic student outcomes.

In their 2017 meta-analysis, Dietrichson, Bog, Filges, and Jorgensen examined interventions that aimed to improve the educational achievement for low socioeconomic status (SES) students in elementary and middle school. Of all the interventions examined, tutoring was both the most common (36 of the 101 studies employed a tutoring component) and the most effective, with an average effect size of 0.36-standard deviations on standardized academic tests (95% confidence interval [CI] [0.26, 0.45]).

More recently, Nickow, Oreopoulous, and Quan (2020) surveyed 96 K-12 tutoring interventions evaluated in RCTs. They found strikingly similar results to those of Dietrichson and colleagues — the pooled effect size was 0.37-standard deviations on academic learning outcomes (95% CI [0.30, 0.43]). While the majority of these studies had relatively small sample sizes, 15 studies included over 400 participants and the magnitude of the effect attenuated only slightly. The larger-scale tutoring programs had an average effect size of 0.25-standard deviations on learning outcomes, and the effect size does not appear to get smaller as the sample size of large programs increased. These findings provide evidence that implementing tutoring programs at scale can meaningfully improve student outcomes.

## The Drivers of High-Impact Tutoring

While the reviews of research document positive effects, the effectiveness of individual tutoring interventions vary. Several potential mechanisms could explain how tutoring interventions positively impact student learning and why different programs show different effects. Here we list some of the mechanisms theorized in the academic literature and how they might manifest differently based on the characteristics of a tutoring program:

- Increased instructional time
  - programs vary in intensity, including the number of hours per week
- Increased customization/personalization of learning
  - programs vary in how much they monitor student learning and how closely they align instruction with student capabilities
- Improved pedagogy, such as new strategies for smaller instructional groups, better use of resources such as online materials, or more time on task due to fewer disruptions
  - programs vary in the number of students per tutor in a session and in their use of instructional materials as well as in their selection, oversight, and professional development of tutors
- Positive social connection between students and their tutors that promotes engagement with educational content and increased motivation
  - programs vary in their emphasis on the tutor-student relationship and whether students meet with the same tutor over time

Few studies to date have directly tested the importance of the different mechanisms or the specific characteristics of tutoring programs. Studies have, instead, assessed the effects of programs which are an amalgamation of features. Nonetheless, we can gain some understanding of the effects of program features by comparing the characteristics of more and less effective programs. In this way, the meta-analysis by Nickow and colleagues provides insights into how effectiveness varies by program characteristics.

## Creating a Framework for High-Impact Tutoring

With a goal of creating a framework for high-impact tutoring, we have identified some of the characteristics of tutoring programs worth considering. These include: Subject area; Grade level; Students (what students receive the tutoring); Safety protocols; Organizational culture; Language; Diversity of tutors; Cultural competency training; Social-emotional training; Tutor type; Tutor recruitment; Tutor training and support; Tutor oversight; Delivery (online or in-person); Dosage; Student-tutor ratio; Student grouping; Tutor consistency; Tutor-student relationship; Structure; Quality of materials; Facilitation; During-school or out-of-school time (OST) setting;

Displacement (i.e., the counterfactual; what the student would have experienced without the tutoring); Take-up (is the tutoring required or do students and/or their parents have to enroll); Curricular alignment; School and teacher engagement; Family engagement; Design and improvement process; Formative Assessment; Student progress; Program Evaluation; and Tutor effectiveness. We review the research on these characteristics below.

Figure 1 illustrates how we are conceptualizing the components of high-impact tutoring, how they relate to one another, and how they lead to student growth.

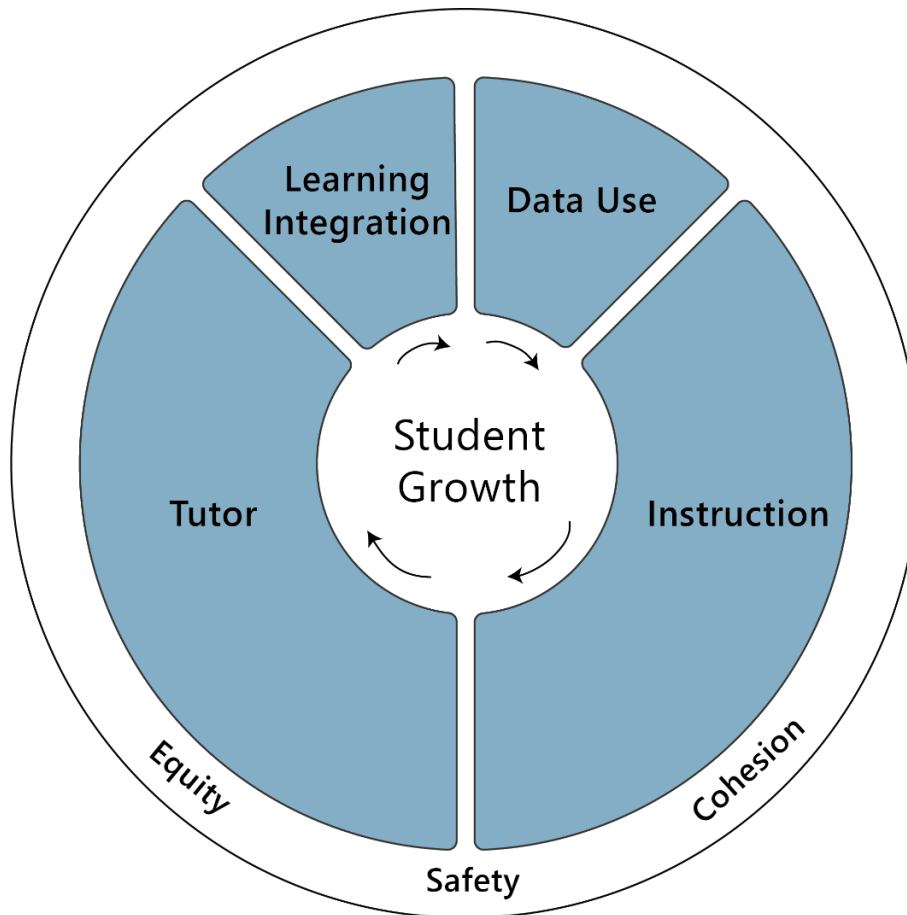


Figure 1. Framework for High-Impact Tutoring

## Tutoring Program Elements and Characteristics

So that stakeholders can better understand and apply our framework to their own needs, we have grouped the many characteristics of tutoring programs under Program Focus or one of seven broader elements: **Safety; Equity; Cohesion; Tutor; Session; Learning Integration;** and **Data Use.**

### Program Focus

#### Subject

Tutoring interventions have focused primarily on producing learning gains in literacy and math. Across all evaluated programs at all levels, the effect sizes for literacy and math are similar to one another ( $ES = 0.35$  and  $0.38$ , respectively). Some of the variation in the impact of these programs varies by grade level, which we discuss further below. We found little rigorous evidence evaluating tutoring programs that are not subject-focused (e.g., comprehensive tutoring that targets students' overall learning), or other subject areas besides math and reading (e.g., science and social studies).

Reviews have concluded that one-to-one tutoring is the most effective intervention known to improve the reading achievement of struggling students in elementary school ( $ES = 0.41$ ), but small group tutoring (led by teachers or paraprofessionals) can also improve student literacy ( $ES = 0.24$ ) (see Neitzel et al., 2020). While relatively few tutoring programs in the U.S. focus on reading for secondary students, a review of seven UK-based studies found that tutoring can improve secondary students' reading achievement (Baye et al., 2019).

Tutoring appears to have a robust impact on elementary math achievement as well. One-to-one and small group tutoring appear to be equally effective for improving math achievement for elementary school students ( $ES = 0.19$  and  $0.30$ , respectively) when utilizing teacher or paraprofessional tutors (Pellegrini et al., 2020). We discuss the small number of rigorous evaluations for math tutoring in secondary schools in the Middle and High School section below.

#### Grade Level

##### Pre-Kindergarten – Grade 1

Mirroring the effects of early investments in children's education (Heckman, 2006), tutoring interventions aimed at students in first grade and below tend to have the comparatively biggest impact on student learning ( $ES = 0.45$  and  $0.42$ , respectively). Reading-focused tutoring interventions for first graders are among the most popular of programs evaluated ( $ES = 0.43$ ), and it would appear that this age could be a "sweet spot" for targeting efforts to improve literacy. Math tutoring interventions aimed at first graders are less prevalent than reading interventions but also show positive effects ( $ES = 0.38$ ), including in promising large-scale evaluations (Fuchs et al., 2013; Gersten et al., 2015).

### Grades 2 – 5 (Elementary School)

While the effect of tutoring interventions appears to decline as students progress through elementary school ( $ES = 0.29$ ), this pattern is primarily driven by decreasingly effective literacy programs (although note that this may be an artifact of testing formats, available evidence and differential learning rates by age). Literacy tutoring programs become relatively less effective as students progress beyond kindergarten and first grade ( $ES = 0.22$ ). However, math tutoring programs actually become more effective for older elementary students ( $ES = 0.44$ ).

### Grades 6 – 12 (Middle & High School)

There are very few published randomized controlled trials evaluating tutoring initiatives in middle and high school. While the small number of studies make it impossible to make generalized claims, a large-scale high school tutoring intervention generated promising results. The program, facilitated by Saga Education, involved service fellows tutoring ninth and tenth grade male students for the school year. Math test scores of tutored students increased by an average of 0.19- to 0.30-standard deviations (Cook et al., 2015). Recent, ongoing evaluations appear to replicate these positive results on students' math achievement.

While tutoring interventions appear to get less effective as students get older, it is important to note that average year-to-year learning gains get smaller as students age and it is, in general, harder to impact student learning outcomes in older students. (Bloom et al., 2008). Thus, small effect sizes in higher grades may still be valuable, particularly given that other potential interventions tend to have much smaller effects. We believe delivering and evaluating tutoring interventions at the secondary level has great potential to bolster student learning.

## Students

Who receives tutoring? Most tutoring interventions tend to be *need-driven* and targeted to students who are struggling and perform below particular thresholds. In Chicago, incoming male 9<sup>th</sup> and 10<sup>th</sup> graders were identified as at high-risk of dropping out and assigned to receive intensive 2:1 tutoring sessions for the school year (Cook et al., 2015). Tutoring can also be *curriculum-driven* and provided at critical moments when students tend to fall behind. For instance, Reading Recovery tutoring programs occur during first grade because it is a crucial point for literacy development (Sirinides et al., 2018). Finally, tutoring can be *universal*. This model, in which all students receive tutoring, has been implemented by Match Education charter schools.

There is little research on what student target method is most effective, but problem-driven tutoring may be more efficient (from a cost and logistics standpoint) and may be able to direct resources toward those students who have been left behind. Universal tutoring, on the other hand, may make tutored students feel less stigmatized, could address the needs of the (oft-forgotten) mid-performing students, and may provide a vehicle in underserved communities for high-performing students to excel.

---

## Safety

### Safety Protocols

All tutoring programs need to consider how they ensure student safety. This may involve some combination of meeting district, state, and institutional regulations, screening tutor candidates, training, and oversight. For tutoring delivered in a virtual setting, there are additional considerations involving cyber safety and student privacy. While we do not discuss the research on safety in schools, we believe tutoring programs cannot operate without making student safety a pillar of their operation.

---

## Equity

### Language

In the majority of interventions reviewed, tutoring was conducted in English. However, tutoring programs may provide tutoring in students' native languages. Providing English-language learner (ELL) students access to tutors who speak their language may complement classroom-based learning that occurs in English. At this point, we lack specific evidence suggesting whether ELL students would benefit from having a tutor who speaks their native language. There are a number of tutoring interventions on which we currently do not focus specifically geared at teaching ELL students English.

### Diversity of Tutors

We discuss tutor types below, but tutoring programs can have a more or less diverse group of tutors to deploy. Given that the teaching force currently is overwhelmingly white, recruiting and employing a more diverse group of tutors may help to provide role models for students of color and produce greater learning gains (Gershenson et al., 2018). The section on "Tutor-Student Match" addresses this idea in greater detail.

### Cultural Competency and Social-Emotional Focus

One way to promote equity among tutoring programs is to explicitly elevate and train tutors on relevant concepts. Cultural competency training and social-emotional training, in particular, are likely to increase tutors' ability to connect with students. Tutoring programs that provide tutors with greater amounts of training and support may be successful in part because they prepare tutors to handle the social-emotional needs of students and offer a framework for reaching students of all backgrounds. Future research might consider the benefits of incorporating training in these principles in addition to focusing on content and instruction.

---

## Cohesion

### Organizational Culture

Successful organizations have clear missions and goals that direct the daily operations of their staff. Tutoring programs are no different, and those that are well-run and have a cohesive and aligned organizational culture will likely be more effective in the long run. Studying how the organizational features of tutoring organizations

might impact tutor and student outcomes could contribute to and extend the research base on educational institutions as organizations.

## Tutor

### Tutor Type

A wide range of people can serve as tutors. We group tutor types into seven broad categories: *teachers* (certified teachers), *paraprofessionals* (school staff members, service program fellows, community organization staff), *volunteers* (post-graduates, community members, seniors), *university students* (undergraduate students, graduate students), *private tutors*, *families* (parents, siblings, other family members), and *peers* (classmates, near-peers).

### Teachers

The research suggests that tutoring is most effective when classroom teachers fulfill the role of the tutor (effect size [ES] = 0.50). All of the tutoring interventions evaluated that had teachers as tutors were for grades K-5, and the majority were focused on literacy gains in first grade. The Reading Recovery program, which requires teachers to complete a year-long graduate-level course before tutoring first graders to improve their reading ability, is among the most effective tutoring interventions evaluated. Other tutoring interventions with teacher tutors, even those with substantially less training involved, consistently improve student outcomes.

### Paraprofessionals

For this review, we consider paraprofessional tutors to be either school staff members, service program fellows (e.g., AmeriCorps fellows), or community organization staff. Tutoring interventions led by paraprofessionals are only slightly less effective than those led by teachers (ES = 0.40). Given that these interventions exhibit extremely consistent effects, employing paraprofessionals as tutors may be more feasible from a cost and logistical standpoint than employing teachers.

Several paraprofessional tutoring programs have proven to effectively scale, including those that deploy service fellows as year-long tutors via AmeriCorps (Markovitz et al., 2014; Parker et al., 2019) and Saga Education (Cook et al., 2015). Employing master's- or doctoral-level students studying education or relevant subject areas part-time also appears to be effective (Fuchs et al., 2013; Gersten et al., 2015; Mattera et al., 2018). Neitzel et al. (2020) found there were no statistically significant differences between elementary reading tutoring programs that used teachers as tutors and those that used paraprofessionals (either teaching assistants or paid service program fellows).

### Volunteers & University Students

While we conceptualize volunteer and university student tutoring interventions as distinct approaches, here we combine the evidence from the two types of programs because they are often conflated in the literature. Moreover, university students are often considered “volunteers,” even if they receive stipends as part of a work-study program.

Overall, tutoring programs staffed by volunteers and university students display positive, but consistently smaller average effect sizes on student learning outcomes than those staffed by teachers and paraprofessional



( $ES = 0.21$ ). Volunteer- and university-staffed tutoring programs that have been experimentally evaluated tend to focus on literacy in elementary school, including the AARP Experience Corps program, which found positive results on student reading ( $ES = 0.10$ ). In general, paid volunteers appear to be more effective than unpaid volunteers (Neitzel et al., 2020).

### Private Tutors

Private tutors are individuals who run or are employed by for-profit or non-profit tutoring organizations. Schools may hire tutoring organizations as vendors or parents may individually hire private tutors. The present review does not include tutoring organizations that use professional tutors (i.e., non-service fellows), but these programs may provide unique benefits for students.

In the United States, parent-hired private tutors typically are limited to students from upper- middle- and upper-class families who can afford the expense. These private tutors are often employed to help students prepare for standardized tests. In other countries, private tutoring is much more widespread, particularly in East Asia. For instance, 60-80% of students in Korea and Singapore receive private tutoring.

### Families

Almost all of the family-focused tutoring programs involve parents acting as tutors. This approach consists of providing parents with training and materials to tutor their child in their home. Siblings and other relatives may also act as tutors, but no systematic evaluations of such efforts exist. In general, parent tutoring interventions appear to be about as effective as volunteer-based efforts ( $ES = 0.23$ ) (Ritter et al., 2009). However, fewer rigorous evaluations have assessed parent tutoring interventions than other approaches, and the effects noted in the existing evaluations are more varied.

### Peers

Peer and cross-age tutoring involves students being responsible for providing instruction to help another student (or students) learn (Robinson et al., 2005). Peer tutoring is when students tutor other students who are at their own grade level. Cross-age tutoring involves older students tutoring younger students. Notably, this strand of research tends to focus on benefits for both the tutors and their students. For this reason, peer tutoring is often considered to be a form of “cooperative learning” rather than tutoring. The present summary is not focused on evaluating what characteristics make for effective peer tutoring programs, but we share some evidence from the literature here and think peer tutoring is a promising area for encouraging student learning and social-emotional gains.

Overall, the evidence suggests that peer tutoring can be effective. Two meta-analyses found that the average effect size is about a quarter of a standard deviation ( $ES = 0.22$ ; Dietrichson et al., 2017 and  $ES = 0.26$  in elementary reading; Slavin et al., 2009). One of the most prominent peer-tutoring models that has been evaluated is the Peer-Assisted Learning Strategies (PALS), which involves a minimum of three sessions per week. An evaluation across three states found overall impacts ranging from 0.29 to 0.42 SD (Stein et al., 2008).

### Tutor Recruitment

Given the wide range of tutor types, programs will differ in the ways in which they recruit tutors. Tutors may apply based on a posted job description, or they may be actively solicited by tutoring organizations. Moreover,

tutoring programs may choose to set standards for their tutors to meet. For instance, a tutoring program may require a certain level of education or formally articulate the specific knowledge and skills necessary to be effective. Taking the hiring process a step further, tutoring programs may require an “entrance exam” that tutors must pass before actually instructing students. Understanding how to maximize the tutor supply while not lowering standards is an important future direction for research.

### Tutor Training

The type of tutor can vary, as can the training tutors receive before the intervention. In general, it would appear that more highly skilled tutors are more effective (i.e., teacher and paraprofessional tutoring resulted in greater gains than volunteer and family tutoring). But the massive effects demonstrated by teacher tutoring programs are largely driven by a program that requires intensive training and support (i.e., Reading Recovery). Thus, it is hard to differentiate whether it is tutor skill or tutor training that drives the effects. We imagine that it is a combination of both, but we do not know the maximally effective amount of tutor training. We hypothesize that tutors — particularly those with fewer pedagogical skills, like volunteer and university tutors — who receive more pre-training will be more effective than those who receive less pre-training.

### Tutor Support and Feedback

Tutor skill development does not need to end at the end of training. Programs differ in the extent to which they provide their tutors with ongoing support and feedback. Based on prior research on teaching and learning, we hypothesize that tutors who receive higher levels of ongoing support will be more effective than those who receive less.

### Accountability, Consequences, and Incentives

Tutoring programs may choose to set accountability measures for their tutors to meet. For instance, can tutors be fired if their students are not improving on selected assessments? While no studies explicitly focus on tutor consequences, research examining teacher accountability efforts might inform better and worse practices to test. Additionally, tutoring programs could offer incentives for tutors based on tutors’ behaviors (e.g., attendance) and/or student success (e.g., learning outcomes).

---

## Instruction

### Delivery

Tutoring initiatives can be delivered in-person, virtually, or via blended instruction (a combination of in-person and virtual tutoring). For the present project, we recognize that schooling can be delivered asynchronously, but we operate under the assumption that high-impact tutoring involves a synchronous learning component.

Nickow et al. (2020) focused solely on in-person tutoring interventions in their review of the tutoring literature, but there are a few recent evaluations that suggest virtual or blended tutoring may be effective. First, a virtual tutoring intervention that occurred during the pandemic-induced remote schooling in Italy resulted in increases in student academic achievement ( $ES = 0.26$ ) and social-emotional outcomes (Carlana & La Ferrara, 2021). This intervention recruited and trained volunteer college students to tutor middle school students virtually for three to six hours per week for an average of five weeks. Second, a small-scale evaluation of a math online tutoring

program (Cognition) found promising results (Roschelle et al., 2020). Students who received online 1:1 tutoring showed greater gains on a math assessment than those who did not ( $ES = 0.46$ ), which compares favorably to effects found from in-person elementary math tutoring programs ( $ES = 0.26$ ) (Pellegrini et al., 2018). Finally, preliminary results from a recent evaluation found that a blended model was equally effective at increasing student learning as the costlier in-person-only tutoring.

The most rigorous evidence comes from in-person tutoring programs. Whether virtual and blended tutoring interventions can be as effective as those conducted in-person remains an open question.

## Dosage

### Session Frequency and Length

Overall, tutoring interventions appear to be more effective as the number of tutoring sessions per week increases. Once- or twice-a-week tutoring sessions ( $ES = 0.24$ ) are less effective than interventions that involve three ( $ES = 0.34$ ) or more sessions per week ( $ES = 0.41$ ). While whether tutoring occurring four or five times a week is more effective than three-times-a-week tutoring remains unknown, the evidence suggests that once-weekly tutoring sessions have smaller impacts on learning than more frequent sessions. Drop-in and as-needed tutoring programs may also contribute to student learning, but there are few rigorous studies with estimates of their effectiveness.

Tutoring sessions also vary in the length of time tutors and their students spend in each individual session. Sessions vary from 10-15 minutes to more than 60 minutes per session, with the most common session length lasting between 30 and 60 minutes. At this point, we do not know the most effective combination of session frequency and length.

### Program Duration

Tutoring programs can last between one session (i.e., a drop-in session) and multiple school years. In general, most of the programs evaluated lasted between 10 weeks and one school year. Surprisingly, interventions that lasted for fewer than 20 weeks appeared to be slightly more effective ( $ES = 0.39$ ) than those lasting longer than 20 weeks ( $ES = 0.29$ ). However, this difference may be attributed to (the more effective) teacher tutoring programs having shorter durations compared to the volunteer programs.

The studies we report on generally assess student learning outcomes over a relatively short period of time. We know less about whether having a consistent tutor for a longer duration would impact longer-term outcomes or influence other outcomes pertaining to students' relationships, motivation, and behavior. For instance, a recent evaluation of Saga Education's tutoring model reported that tutored students enjoyed math more and reported positive relationships with their tutor.

### Student-Tutor Ratio

We define tutoring as supplemental instruction with one tutor being paired with four students or fewer. The most common tutoring interventions involved a one-to-one tutoring ratio. The effect size was largest when tutors were paired with just one student (1:1;  $ES = 0.38$ ), but the difference between programs with 2:1 ( $ES = 0.29$ ) and 3+:1 ( $ES = 0.36$ ) student-tutor ratios were statistically similar.

Understanding whether larger student-tutor ratios are as effective — or nearly as effective — as 1:1 tutoring is an important thing to learn from a cost-perspective. There may also be social benefits for students working with their peers in consistent small groups (Schacter, 2000).

### Student Grouping

If tutoring occurs in small groups (not one-on-one), the way in which students are grouped may matter. The research on supplemental educational supports suggests grouping students based on skill level may be most effective (Zimmer et al., 2010). It also may make sense to pair English language learners together, particularly if the tutor can speak their native language.

### Tutor Consistency

Many tutoring interventions keep one tutor with a set student or group of students for the duration of the program. For instance, some successful interventions — such as those facilitated by Saga Education — are relationship-driven programs that focus on building trusting relationships in addition to enhancing academic achievement. In other tutoring interventions, students might cycle through different tutors throughout the program. A program might have a set pool of tutors that rotate through students. Many volunteer tutoring programs that occur through community organizations (e.g., Boys & Girls Clubs) organize volunteer tutors to provide homework help to students on an as-needed basis.

Many studies are not totally transparent about the consistency of the tutor-student match, nor do they directly evaluate how much it matters. However, students who have teachers for multiple years — a practice referred to as “looping” — appear to have better outcomes (Hill & Jones, 2018). We hypothesize that both tutors and students benefit from working together over longer periods of time, but more research is needed to determine whether this is the case.

### Tutor-Student Relationship

There is a continuum across tutoring programs regarding how vital it is to build strong tutor-student relationships. In some programs, particularly when the student does not consistently work with the same tutor, the relational aspect is of very low importance. In other programs, like the year-long Saga Education tutoring initiatives, the relational component is central to the model.

Further investigation into the importance of the tutor-student relationship to student outcomes is warranted. None of the following characteristics have been explicitly studied yet, but programs differ in the extent to which they: (a) aim to maintain tutor consistency, (b) attempt to match students with “similar” tutors (e.g., matching on background or race), and (c) assess the quality of the tutor-student relationship. Programs also differ as to whether they provide materials and strategies for tutors explicitly focused on building strong tutor-student relationships, which are most often cultivated during the session.

### Structure

While the evidence base does not provide detailed information on the structure of specific tutoring interventions and how they affect student learning outcomes, students tend to benefit from consistent lesson structure. In addition to providing opportunities for formative assessment, routines save time, provide a space for independent practice, and facilitate behavior management (Schumaker et al., 2002). As Neitzel et al. (2020) note, successful tutoring programs almost invariably use structured, sequenced approaches, with manuals and

materials. These tutoring models may include technology, but the instruction is primarily driven by the tutor, not the technology. As an example, Reading Recovery has highly structured lessons with clear and detailed directives to teachers who are tutoring students (Sirinides et al., 2018). Similarly, Saga Education follows a standard lesson structure which often involves a “Do Now” and an “Exit Ticket” to start and end the tutoring session, respectively.

### Quality of Materials

The quality of the materials tutors have at their disposal will undoubtedly correlate with the quality of the tutoring instruction itself. As is expected in the classroom, tutoring sessions should similarly deploy high quality, grade-level appropriate, and rigorous instructional materials. These materials will also likely be more effective if they align with the classroom curriculum (as discussed in the Learning Integration section). Future studies might vary the amount or type of materials provided to tutors to learn more about how much they matter for student outcomes.

### Facilitation

One of the reasons teacher-led tutoring interventions may be so successful is that teachers are experts at facilitating learning. However, non-teacher tutors can also adopt strategies that promote student learning. For instance, tutoring programs that encourage tutors to ask students open-ended questions, provide a variety of learning tools, and offer guidance on pacing likely increase tutors’ abilities to bring about student gains. Effectively facilitated sessions may also promote more positive relationships between tutors and students (see more in the “Tutor-Student Relationship” section). Future research may choose to focus on what tools and strategies set the stage for effective instruction.

---

## Learning Integration

### Setting: During-School vs. Out-of-School Time

Tutoring programs can occur either during school or out-of-school time (OST). The OST tutoring programs reviewed were implemented after school, as opposed to during vacations or in the summer. In the aggregate, the during-school programs have an effect size that is nearly twice as large as that of the after-school programs ( $ES = 0.40$  and  $0.21$ , respectively). That said, the magnitude of this difference could be in part due to the fact that most teacher and paraprofessional tutoring programs (which tend to be more effective) are conducted during school hours.

While more research is needed, during-school programs may benefit from additional features that we hypothesize make for effective tutoring: easier alignment of tutoring sessions with students’ classroom curriculum; facilitated teacher-tutor interaction; more regular attendance with less effort (i.e., no additional efforts are needed to get students to attend OST tutoring sessions); and a likely increase in ease of take-up.

### Displacement / The Counterfactual

If tutoring is occurring during school, the session will necessarily displace something else. In the research we reviewed, not much attention was paid to what the tutoring was replacing. In some cases, interventions

specifically did not decrease time spent in the target subject of the tutoring (e.g., with math tutoring, students were not removed from their primary math classes). However, elementary reading tutoring programs that provide additional time for tutoring seemed to be no more effective than those in which tutoring replaced regular reading instruction (Neitzel et al., 2020). In many programs, tutoring often replaced non-core subjects (e.g., art, music, physical education). We do not yet have specific recommendations based on the available evidence as to the best way to schedule tutoring. But we acknowledge that the benefits of the tutoring must be weighed against the potential costs of missing out on other classes or activities.

### Take-up

Tutoring interventions can differ in how they are taken up by students. Students can be *required* by the school to receive tutoring — in this case, students tend to have tutoring sessions embedded in their school day schedule. Conversely, tutoring can be *voluntary*, where students or parents choose to enroll. Many tutoring interventions — particularly those conducted outside of school hours — are driven by parents opting their children into the program. However, voluntary tutoring can also be offered on an *opt-out* basis (where students are enrolled unless their parent actively asks that they not be) rather than an *opt-in* basis (where parents have to actively enroll their child). While these different features have not been experimentally tested, we hypothesize that required opt-out tutoring interventions would be more effective than opt-in ones because they reduce the barriers to entry (Bergman et al., 2020).

### Curricular Alignment

The content tutors teach to students may be generic (e.g., based on certain subject or grade-level standards) or aligned with the curriculum students are learning in their classrooms. Because many of the most effective tutoring programs occurred during school, it is possible that students make more learning gains when the tutoring curriculum complements and is responsive to the classroom instruction students receive. It would appear that successful tutoring programs do make efforts to align their curricula, but we still lack clear empirical support to definitively say that tutoring interventions that are aligned with the curriculum are better than those that are not.

### School and Teacher Engagement

Both during-school and OST tutoring programs can differ in the amount they enlist involvement from schools and teachers. In general, tutoring that occurs during school likely requires at least a small amount of school engagement, whereas OST tutoring programs do not necessarily require any (although many do coordinate with schools). While during-school programs appear to be more effective than OST tutoring programs, we do not know whether increased school engagement or other factors may be driving the differential impacts.

There is some suggestive evidence that tutoring programs that employ school-based coordinators may be more successful (Heinrich et al., 2014). These coordinators can handle logistics and help tutors get individualized student information. Relatedly, the amount of interaction between tutors and classroom teachers also varies greatly. We imagine that greater tutor-teacher communication would result in greater curriculum alignment, as well as facilitate sharing information on students' needs and progress.

## Family Engagement

In general, parental engagement in education is positively associated with student success. And, as with school, families can engage to different degrees in their child's tutoring experience. Parent tutoring, by definition, is an example of an intervention with high family engagement because it relies on the parent to actually take on the role of tutor. However, programs that deploy paraprofessionals, volunteers, and university students as tutors may be in a good position to both increase and capitalize on family engagement. For instance, tutors could strengthen family engagement in their child's education by regularly communicating with parents. Thus, encouraging tutors to proactively engage with families might strengthen students' learning gains.

---

## Data Use

Those conducting tutoring interventions can use data in various ways. In addition to evaluating whether an intervention works (i.e., a summative assessment), they also can use data to target which students might benefit from tutoring the most. We need more research to understand whether and how data use impacts the effectiveness of tutoring interventions, as tutoring programs differ in the extent to which they encourage tutors to use data to inform their instruction.

## Design and Improvement Process

The extent to which tutoring programs utilize data in their design and improvement process likely associates with the effectiveness of the tutoring program. Drawing on prior research to design a tutoring intervention will likely result in higher-quality tutoring. In addition, programs that capture ongoing implementation data to make course correction improvements both within and across tutoring interventions will likely eventually become increasingly effective over time. Future research might attempt to quantify how using data to design and improve tutoring programs associates with student learning and other outcomes.

## Formative Assessment

A specific type of data usage involves formative assessment, and tutors are particularly well positioned to use formative assessment when teaching students. Formative assessment — whether conducted in the classroom or in a tutoring session — provides the tutor with timely feedback on each student and allows for personalized instruction. Importantly, for formative assessment to result in more student learning, tutors need (a) time and support to review the assessment and (b) the ability to act upon it (Bennett, 2011).

## Student Progress

In addition to using formative assessments to inform tutoring instruction, tutoring programs can collect data on student progress over time. Monitoring student progress can occur at the individual, tutor, or program level, and progress can be measured by self-designed assessments, grades, or standardized test scores. Tracking progress, however, is just the first step — using the information gleaned to continuously develop tutors and programs is likely when such tracking may lead to downstream improvements in program quality and student learning.

## Program Evaluation

Arguably, tutoring organizations need to use data, or share data with external researchers, to conduct a program evaluation. However, individual programs differ significantly in the rigor with which they assess whether their program is working. Numerous factors contribute to the quality of program evaluations, among them: the research methods used (e.g., randomized controlled trials, quasi-experimental, observation, qualitative); the data collected and used (e.g., test scores, grades, self-reports); and the duration of the evaluation (e.g., short-, medium-, or long-term). In this review, we have prioritized the results from studies that evaluate tutoring interventions with randomized controlled trials.

## Tutor Effectiveness

Traditionally, measuring the effectiveness of tutoring programs has often involved collecting student achievement test scores. But no studies have rigorously examined how to identify whether a tutor is effective or not beyond that metric. Building off of the large body of academic literature on teacher effectiveness, programs could apply a “value-added” approach or use observations to rate tutors on their performance with students. It remains an open question as to whether assessing tutors’ effectiveness improves student outcomes.



## A RESEARCH AGENDA MOVING FORWARD

While the evidence suggests high-impact tutoring interventions consistently produce positive learning gains for students, there is ample opportunity for future research to deepen our understanding of what works and why. The goal of this research agenda is not to encourage evaluation of any one particular tutoring intervention — although we strongly endorse rigorous evaluations of programs being implemented with youth. Instead, we focus more broadly on two, complementary lines of suggested research, with a goal of understanding how to deliver high-impact tutoring to students in need across the U.S.

First, we aim to identify research questions that will advance knowledge on how and why tutoring interventions are effective. Research responding to these questions will uncover specific characteristics of effective tutoring programs. Second, we introduce research questions regarding how to implement tutoring effectively at scale. This research line articulates supply, demand, and implementation questions; new knowledge generated by this research will be integral to scaling high-impact tutoring.

### Research Priority #1: Identifying the Characteristics of Effective Tutoring

There are several particularly promising avenues for future research on identifying the characteristics of effective tutoring, and we highlight and outline priority research questions in detail below. These questions focus on: the cost-effectiveness of different tutoring models; tutor skills; tutor-student relationships; adoption by stakeholders; how programs increase engagement and implementation fidelity; and how best to extend the research focus. Following a discussion of these areas of focus, we summarize open research questions that we hope will be explored by researchers in the field as tutoring interventions continue to be evaluated in Table 1.

#### The Cost-Effectiveness of Different Tutoring Models

**Overarching Question:** How can programs deliver tutoring in the most cost-effective way?

Expanding high-impact tutoring to reach all students in poverty will be easier if the costs are lower. As a result, understanding the importance of the main cost-drivers is important.

#### Student-Tutor Ratios

One way to lower costs is to group more students with a single tutor. To understand the marginal impact of adding an additional student to a tutoring session, an evaluation could focus on randomly assigning students to tutoring groups with ratios ranging from 1:1 to 4:1. Smaller student-tutor ratios might allow for more personalized instruction and stronger tutor-student relationships, which could increase student learning. On the other hand, having multiple students participate in a tutoring session could improve peer relationships and motivation, which could also result in improved student learning.

Primary Research Question	How to Test	Related RQs
What is the marginal change in student learning when the student-tutor ratio increases by an additional student (from 1:1 to 4:1)?	<i>Within-Program RCT:</i> Randomly assign students to tutoring groups with ratios ranging from 1:1 to 4:1.	Do smaller student-tutor ratios facilitate more positive tutor-student relationships?  Do small group tutoring sessions facilitate more positive peer relationships than 1:1 tutoring sessions?

### Dosage

The number of hours students spend with tutors each week has major implications for program costs. The evidence suggests that one or two tutoring sessions per week is not enough, but it remains an open question as to whether three tutoring sessions per week are as effective as four or five sessions. By intentionally varying the number of tutoring hours per week within a program, we can learn the ideal tutoring dosage and at what point more time spent on tutoring is no longer necessarily better.

Primary Research Question	How to Test	Related RQs
What is the marginal benefit of increasing the number of hours students spend with tutors each week?	<i>Within-Program RCT:</i> Randomly assign students to receive different numbers of tutoring hours per week within a program.	Does the distribution of tutoring dosage matter? Could tutors and students meet for 2 hours, 1x per week and receive the same benefits as tutors and students who meet for 30 minutes, 4x per week?  Does the ideal session frequency and length differ by age and subject?

### In-Person vs. Virtual vs. Blended Tutoring

The majority of the evidence we reviewed suggests that in-person tutoring is effective. Whether virtual tutoring can be as effective remains an open question. By randomly assigning students to receive in-person versus virtual tutoring, we can learn what the marginal impact of in-person versus virtual tutoring is on student outcomes. Studies can also explore the marginal impact of a blended-learning tutoring model as compared to an in-person model, building on a recently evaluated program, which found no statistically significant difference between the two delivery types on student academic outcomes.

Primary Research Question	How to Test	Related RQs
Can virtual tutoring be as effective as in-person tutoring?	<p><i>Within-Program RCT:</i> Randomly assign students to receive tutoring in person, virtually, or through a combination of the two.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare the differences between in-person and virtual tutoring programs.</p>	<p>What is the marginal impact of in-person over virtual tutoring?</p> <p>What is the marginal impact of blended over virtual tutoring?</p> <p>What is the marginal impact of in-person over blended tutoring?</p>

### Computer-Assisted Learning Assistants

Relatedly, can tutoring programs use educational technology to effectively deploy tutors with less experience (e.g., volunteers, university students)? If tutors take on the “human” role of building relationships with students and directing student attention, they could support students as they navigate computer-assisted instruction platforms (e.g., Khan Academy) that provide content and instruction. By shifting the role of tutor to “caring adult monitor” as opposed to instructor, could tutoring leverage human interaction in a cost-effective way? Put another way, could synchronous tutoring sessions involve asynchronous instruction to effectively increase student learning while employing relatively lower-skilled labor?

Primary Research Question	How to Test	Related RQs
Could synchronous tutoring sessions involve asynchronous instruction to effectively increase student learning?	<p><i>Within-Program RCT:</i> Randomly assign students to tutoring sessions that use computer-assisted instruction and those that do not.</p>	Does adding an asynchronous component to tutoring allow for a) employing lower-skilled tutors and b) tutors to increase their caseloads?

### Tutor Type & Training

Finally, and related to the next focus area (“Tutor Skills”), it will be critical to explore whether lower-skilled tutors can become effective with the right training and support. Inherent in that question is understanding what type of training and how much training tutors of different skill levels need. While individual programs may be able to explore the effectiveness of tutors with varying skill levels, a multi-program evaluation would allow comparisons of how programs that differ in their minimum tutor requirements and associated training impact student outcomes. For instance, do programs that require a college degree or that use entrance assessments to select tutors tend to have more beneficial student outcomes?

Primary Research Question	How to Test	Related RQs
What type and amount of training do lower-skilled tutors need to improve student learning?	<p><i>Within-Program RCT:</i> Randomly vary the length of training tutors receive.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare effectiveness among programs with different types of training (e.g., length, focus, etc.).</p>	Do certain tutor types benefit more than others from training? (e.g., volunteers and university students)

## Tutor Skills

In addition to assessing the most effective tutor type and training, it will be essential to understand what skills and supports tutors need to be successful. Research in this area will contribute to both understanding the costs and raising the quality of tutoring.

## Tutor Recruitment

One tutor recruitment question involves the criteria tutoring programs establish to evaluate prospective tutors. For instance, a tutoring program may require a certain level of education for tutors, identify specific competencies they believe tutors need to be successful, or even administer a subject-based exam prior to hire. A program evaluation of multiple different tutoring programs could start to determine what criteria and standards for tutors contribute to more and less effective tutoring.

Primary Research Question	How to Test	Related RQs
How effective are tutors with different backgrounds?	<p><i>Within-Program Observational Evaluation:</i> Compare the effectiveness of different tutors within the same program.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation comparing student learning outcomes by minimum tutor qualifications.</p>	<p>Are tutoring programs more effective when they require tutors to have a minimum level of education?</p> <p>Are college students, recent college graduates, and seniors equally effective tutors?</p> <p>Are tutors who have degrees in the subjects they are teaching more effective (particularly at the high school level)?</p>

		Do entrance exams for tutors improve the effectiveness of a tutoring program?
--	--	---

## Tutor Support

Programs that provide support for their tutors are likely to be more effective, but we do not know what amounts and types of ongoing support are most effective. We hypothesize that tutors — particularly those who are lower skilled, so volunteer and university tutors — who receive higher levels of ongoing support will be more effective than those who receive less ongoing support. But, because providing support requires human and financial resources, we propose exploring the most effective support by randomly varying the amount of support tutors receive for the duration of the program.

Primary Research Question	How to Test	Related RQs
Does ongoing support enhance the effectiveness of tutoring programs?	<i>Within-Program RCT:</i> Randomly assign tutors to receive high levels of support for the duration of the program versus low levels of support.	N/A

## Tutor Oversight

A multi-program evaluation can start to answer questions surrounding the amount and type of oversight tutoring programs might implement. Oversight is necessary to ensure safety standards and protocols are being met, but it also can maintain tutor quality and effectiveness. It will be necessary to understand whether accountability measures increase tutor effectiveness and, if so, what type of measures should be deployed. Other related questions might include whether setting consequential standards and/or offering incentives for tutors based on either tutors' behaviors (e.g., attendance) or their students' success (e.g., learning outcomes) increases effectiveness.

Primary Research Question	How to Test	Related RQs
Are tutoring programs more effective when they incorporate accountability measures for their tutors?	<i>Within-Program RCT:</i> Randomly vary the types of tutor oversight (e.g., firing, standards, observations, incentives, etc.) tutors receive across tutoring sites.	Are tutoring programs more effective when they set consequential standards for their tutors?  Are tutoring programs more effective when tutors' behaviors are incentivized?

	<p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare effectiveness among programs with different types of tutor oversight (e.g., firing, standards, observations, incentives, etc.).</p>	<p>Are tutoring programs more effective when student learning gains are incentivized?</p>
--	---	---

## Tutor-Student Relationships

Increasing students’ social connections with a caring adult is one mechanism through which tutoring programs are thought to improve student outcomes, so this is a particularly promising (yet currently underdeveloped) line of research. We propose a few opportunities to delve into the relational component of tutoring and how it impacts student learning outcomes.

### Tutor Consistency

First, we propose testing the importance of having the same tutor over time by randomly assigning students to meet consistently with one tutor or interact with multiple tutors over the duration of the program. We hypothesize that a sustained relationship between students and their tutors will allow for more intentional relationship-building and positively impact student outcomes.

### Looping

Students’ educational relationships are uprooted every year when they change grades, and even more so when they transition from one school to another. By randomly assigning students to learn from tutors for multiple years (i.e., looping) versus just one year/semester, we can learn whether a consistent tutor-student relationship is more effective at promoting positive outcomes for students than introducing new tutors each year. We particularly recommend testing whether looping is effective during critical time periods and transitions (e.g., 8<sup>th</sup> to 9<sup>th</sup> grade).

Primary Research Question	How to Test	Related RQs
<p>Does having a consistent tutor for the program duration improve the effectiveness of the tutoring?</p>	<p><i>Within-Program RCT:</i> Randomly assign students to meet consistently with one tutor or interact with multiple tutors over the duration of the program.</p>	<p>Does “looping” students with tutors across semesters or years improve student outcomes?</p> <p>Is “looping” more effective during major transitions (e.g., 8th to 9th grade)?</p>

*Across-Program Evaluation:*  
Conduct a multi-program evaluation and compare the differences between those that use consistent versus varied tutors.

### Embedding Relationship-Building Content

Second, tutoring interventions could randomly assign whether they embed relationship-building content into the tutoring sessions. Based on decades of correlational research on teacher-student relationships (Roorda et al., 2011), we hypothesize that a greater focus on social connection might facilitate better relationships, more enjoyment, and learning.

Primary Research Question	How to Test	Related RQs
Does intentionally embedding relationship-building content in tutoring sessions improve student outcomes?	<i>Within-Program RCT:</i> Randomly assign tutors to embed relationship-building content into tutoring sessions.	Do improved relationships between tutors and students lead to greater student learning?

### Tutor-Student Matching

Third, based on prior research, students who share similar attributes with their teachers might perform better over time. For instance, building on studies showing positive effects from teacher-student race-matching (Gershenson et al., 2018), we propose randomly assigning tutor-student matches based on race match or not. Black students, in particular, appear to benefit from learning from Black teachers. Alternatively, both tutors and students could complete a Big 5 personality test and then be matched based on their personality traits (Cavell et al., 2020). We hypothesize that matching students with tutors who share similar attributes to their own will promote greater engagement and learning.

If possible, we propose conducting a 2x2 experiment where the following factors vary: a) whether tutoring interventions formally introduce relationship-building content, and b) whether tutors and students share certain attributes. This would help us learn the marginal impact of concertedly cultivating relationships over matching tutors and students based on attributes.

Primary Research Question	How to Test	Related RQs
Does matching tutors and students on demographics (e.g., race) improve student outcomes?	<i>Within-Program RCT:</i> Randomly assign tutor-student matches based on demographic matches or not.	What is the marginal impact of demographic-matching when programs focus explicitly on relationship-building?

## Adoption by Stakeholders

Even the best educational initiatives will not be successful if students do not actually engage in them. Tutoring interventions can be adopted at multiple levels: by the district, the school, the parents, and the students. We propose several research avenues to determine what program characteristics will be most likely to increase take-up of and engagement in tutoring initiatives.

### Targeting Students

First, we propose testing engagement when the tutoring is offered on a universal versus need-driven basis. If need-driven tutoring is interpreted as remedial by the school community, students may be less likely to engage because of the stigma associated with it. On the other hand, if all students are expected to engage in tutoring sessions, the intervention may be better received by the students who need the help the most. A universal tutoring intervention might also allow underserved students to advance beyond the classroom curriculum, which could help them become more college and career-ready. Evaluating the data from need-driven versus universal programs would help the field understand the potential trade-offs between how students are targeted from a cost-effective standpoint.

Primary Research Question	How to Test	Related RQs
<p>Are the potential benefits from universal tutoring programs too costly compared to need-driven tutoring programs?</p>	<p><i>Within-Program RCT:</i> Randomly vary the cut-off point for receiving tutoring (e.g., based on a test score or continuous grading score) to determine who receives the tutoring.</p> <p><i>Within-Program RCT:</i> Randomly assign tutoring sites to provide tutoring for a sub-group of students or make tutoring available to all students, holding other aspects of the program design constant.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare engagement across different tutoring program recruitment strategies.</p>	<p>Do universal tutoring programs improve students’ perceptions of tutoring?</p>



### Opt-in vs. Opt-out Enrollment

Second, we recommend exploring the difference between how tutoring interventions are taken up by parents by randomly assigning students to programs where their parents have to either opt-in or opt-out. This design feature may be particularly consequential for out-of-school programs. Based on a large body of research, we hypothesize that tutoring opt-out interventions reach more students and are therefore more effective at improving student learning.

Primary Research Question	How to Test	Related RQs
Are tutoring interventions that require parents to opt-in their students less effective at increasing overall student learning than those that are mandated, or have an opt-out feature?	<p><i>Within-Program RCT:</i> Randomly assign students to programs in which their parents have to either opt-in or opt-out.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare adoption between programs that use opt-in versus opt-out enrollment.</p>	N/A

### Recruitment Strategies

Finally, we can test how different recruitment strategies impact whether different tutoring sites (e.g., districts or schools) implement tutoring programs. We can explore, for example, whether leaders are more likely to take up tutoring programs when the recruitment leads with student learning gains, how tutoring goals align with accountability measures, how tutoring supports teachers, or program cost-effectiveness.

Primary Research Question	How to Test	Related RQs
Are certain recruitment strategies more effective at getting sites (districts, schools) to enroll in tutoring intervention programs?	<p><i>Within-Program RCT:</i> Randomly vary the recruitment strategies used on tutoring sites.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare program adoption across different tutoring program recruitment strategies.</p> <p><i>Survey:</i> Conduct a survey of school leaders varying the presentation</p>	N/A

	of potential tutoring program descriptions and compare expressed interest/requests for more information.	
--	--	--

### Setting: During-School vs. Out-of-School Time

While it would appear that during-school tutoring programs may be more effective than OST tutoring programs, more research is needed to disentangle whether it is actually the setting or other program characteristics that are driving the effect. Ideally, we would test this by randomly providing the same tutoring program to students during school or OST. However, if this test design is not feasible, comprehensive multi-site program evaluations may be able to control for some of the program characteristics that tend to appear more in during-school tutoring programs (e.g., aligned curricula, teacher-tutor interaction, etc.).

Primary Research Question	How to Test	Related RQs
How much more effective are during-school tutoring programs than out-of-school tutoring programs?	<p><i>Within-Program RCT:</i> Randomly vary whether students receive the tutoring sessions during school time or out-of-school time.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare the effectiveness of programs that are during school versus out-of-school.</p>	<p>How much more effective at increasing student learning are mandatory tutoring programs than voluntary tutoring programs?</p> <p>Does the increase of student learning associated with in-school tutoring justify the additional costs?</p> <p>Does during-school tutoring result in more engagement among students than out-of-school tutoring?</p>

### Curricular Alignment

Many successful tutoring programs do make efforts to align their curricula, but we still lack clear empirical support to definitively say that tutoring interventions that are aligned with the curriculum are better than those that are not. We hypothesize that a tutoring intervention will be more effective when it has a system in place for aligning and/or responding to the classroom instruction student receive than if it follows a generic curriculum.

Primary Research Question	How to Test	Related RQs
To what extent is tutoring more effective when the tutoring session content aligns with students' core curriculum?	<i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare effectiveness between programs	What is the best way to measure the degree of alignment between tutoring content and students' core curriculum?

that have standalone curricula and those that have systems for aligning content.

How can tutoring programs be responsive to the classroom instruction students receive?

## How Programs Increase Engagement and Implementation Fidelity

Simply adopting a tutoring program will not magically increase student achievement. For tutoring to be successful, various stakeholders need to be actively engaged in the program and implement the program with fidelity. Because so many different stakeholders are involved, research can test different levers for promoting engagement and study how they impact outcomes.

### School Engagement

Even when schools are involved in delivering tutoring, the extent to which they are engaged in the day-to-day tutoring can differ greatly. One way to manage — and potentially increase — school engagement is to have a school-based coordinator dedicated to handling tutoring logistics. These school-based coordinators can come from the school or the tutoring organization. Though coordinators impose an additional cost, we hypothesize that tutoring programs will be more effective when they employ them. We propose varying whether schools have a school-based coordinator or not, or exploring differences in effectiveness across many programs that do and do not have these site-based coordinators.

Primary Research Question	How to Test	Related RQs
Are tutoring programs more effective when resources are dedicated to connecting them to the school?	<p><i>Within-Program RCT:</i> Randomly vary whether tutoring sites have a school-based coordinator.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare effectiveness between programs that employ school-based coordinators or not.</p>	Does school engagement with the tutoring program improve teachers' perceptions of tutoring?

### Teacher Engagement

Across and within programs, the amount of interaction between tutors and classroom teachers also varies greatly. We hypothesize that facilitating more tutor-teacher interactions may result in more positive student outcomes. One way to test this hypothesis would involve varying whether tutoring programs intentionally open

and promote lines of communications between tutors and teachers. A multi-site program evaluation could also start to explore how different programs perform based on how they approach tutor-teacher interactions.

Primary Research Question	How to Test	Related RQs
Does more tutor-teacher interaction result in better student outcomes?	<p><i>Within-Program RCT:</i> Randomly vary whether tutoring sites facilitate communication between teachers and tutors.</p> <p><i>Across-Program Evaluation:</i> Conduct a multi-program evaluation and compare effectiveness between programs that actively connect tutors and teachers.</p>	<p>Does teacher engagement with the tutoring program improve teachers' perceptions of tutoring?</p> <p>Does teacher engagement with the tutoring program improve students' perceptions of tutoring?</p>

### Family Engagement

Tutoring programs, both those that are offered during school and OST, may benefit from engaging families. To learn more about how families can increase the impact of tutoring, we propose randomly assigning parents of students to receive information on their child's tutoring program or not. We hypothesize that greater family support of tutoring could improve students' perceptions of tutoring, their engagement, and, ultimately, their learning.

Primary Research Question	How to Test	Related RQs
Does increased family engagement with the tutoring program improve student outcomes?	<p><i>Within-Program RCT:</i> Randomly assign students to have their parent receive information on the tutoring program or not.</p>	Does parental engagement in the tutoring program improve students' perceptions of tutoring?

### How Best to Extend the Research Focus

The majority of the studies that experimentally evaluate tutoring interventions have focused on measures of student achievement, like standardized tests. However, tutoring has the potential for more widespread impacts.

### Expanding Outcome Measures

For instance, in addition to tested outcomes, tutoring could also result in improvements in student grades, increased motivation, more positive relationships, advanced course-taking, reductions in absenteeism, fewer

disciplinary actions, and lower rates of dropout. A number of studies do assess some of these additional outcomes, like student motivation. Evaluations of Saga Education tutoring programs have focused on non-tested outcomes and find promising impacts on dropout rates, grades, and enjoyment of the subject matter. Moreover, a number of these outcomes might mediate the eventual effect on student achievement tests. By expanding the outcomes measured, we can further understand how and why tutoring positively impacts students. In doing so we may also be able to capture spillover effects from tutoring interventions, like improved teacher-student relationships (perhaps because the tutor takes some of the pressure off teachers to catch students up) or benefits for students in the class who are not receiving tutoring.

To determine the feasibility of enacting high-impact tutoring at scale, future evaluations should also incorporate measures of adoption and fidelity of treatment. For instance, programs should systematically evaluate what recruitment methods are more and less effective. It will also be helpful to identify and capture markers for successful implementation. Tutor and student attendance, time on task, and curriculum reviews — to name a few — might serve as indicators of fidelity.

Primary Research Question	How to Test	Related RQs
Do tutoring programs improve outcomes beyond student test scores?	<i>Within-Program RCT:</i> Randomly assign students to receive tutoring or not and assess a wide variety of student outcomes.	Does tutoring improve motivation, result in more positive relationships, encourage advanced course-taking, reduce absenteeism, reduce disciplinary actions, and lower dropout rates, etc.?

### Long-Term Outcomes

As tutoring programs expand, we need to think more critically about long-term outcomes. The majority of the impressive results emerging from the research on tutoring programs have focused on relatively discrete amounts of tutoring (e.g., 12 weeks, one semester, one school year). However, we do not know if struggling students will need one or two years of intensive tutoring, or whether they will require ongoing support throughout their education.

Primary Research Question	How to Test	Related RQs
Does higher-dosage tutoring result in longer-term outcomes?	<i>Within-Program RCT:</i> Randomly assign students to receive different durations of tutoring (e.g., one year, two year, multi-year) and assess long-term outcomes.  <i>Across-Program Evaluation:</i>	How long do the effects from tutoring persist?  How do different tutoring program characteristics contribute to long-term outcomes?

	<p>Conduct a multi-program evaluation and compare long-term effectiveness among different types of programs.</p>	
--	--	--

**Tutor Benefits**

Noticeably absent from most studies are the effects on the *tutors*, themselves. We believe there are likely positive consequences for tutors who meaningfully interact with students in tutoring programs. In particular, volunteers and college students may experience both career and social benefits. Research suggests that acting as an advisor for students can improve the advisor’s outcomes. We propose making a more explicit effort to evaluate whether the tutor experiences positive outcomes.

Additionally, to promote better outcomes for tutors, tutors could be randomly assigned to receive specific career development or social support during their tenure as tutors. Potential outcome measures include, but are not limited to: sense of purpose, likelihood of pursuing education-related careers, time to next job, likelihood of voting, anti-racist attitudes, performance on mental health batteries, and workforce development behavior (e.g., attendance, project management tasks).

Primary Research Question	How to Test	Related RQs
<p>Do tutors experience positive social and career-related benefits?</p>	<p><i>Within-Program RCT:</i> If tutoring programs have an oversupply of eligible tutors, randomly select who becomes a tutor and collect survey and longitudinal data.</p> <p><i>Within-Program RCT:</i> Randomly assign tutors to receive specific career development and social support during their tenure as tutors.</p>	<p>Can tutoring programs intentionally promote better outcomes for tutors through curated additional programming?</p>

**Table 1. Research Questions by Tutoring Program Characteristics**

Element	Program Characteristics	Research Questions
Program Focus	Subject & Grade Level	<p>Are tutoring interventions more effective when they are subject-based or focus on students' overall schooling?</p> <p>Are subject-focused tutoring interventions effective at improving student learning outside of math and literacy (e.g., science, social studies)?</p>
	Students	<p>Are the potential benefits from universal tutoring programs too costly compared to need-driven tutoring programs?</p> <p>Do universal tutoring programs reduce the stigma that is sometimes associated with tutoring?</p> <p>Can universal tutoring programs help mid- and high-performing students who may lack opportunities reach new academic heights?</p>
Safety	Safety Protocols	How do specific safety protocols relate to student outcomes?
Equity	Language	Do ELL students benefit when tutors can speak their native language?
	Diversity of Tutors	<p>Do students benefit from a more diverse range of tutors?</p> <p>Does matching tutors and students on demographics (e.g., race) improve student outcomes?</p>
	Cultural Competency Training	<p>Does providing tutors with cultural competency training improve students' social and motivational outcomes?</p> <p>Does providing tutors with cultural competency training improve tutor-student relationships?</p> <p>Does providing tutors with cultural competency training increase student learning outcomes?</p>
	Social-Emotional Training	Does providing tutors with social-emotional training improve students' social and motivational outcomes?

		<p>Does providing tutors with social-emotional training improve tutor-student relationships?</p> <p>Does providing tutors with social-emotional training increase student learning outcomes?</p>
Cohesion	Organizational Culture	What organizational features result in more and less effective tutoring programs?
Tutor	Tutor Type	<p>Do certain tutor types benefit more than others from training (e.g., volunteers and university students)?</p> <p>Can tutoring interventions leverage educational technology to effectively deploy tutors with less experience (e.g., pair computer-assisted instruction with volunteers or university students)?</p> <p>Do certain characteristics make for more effective tutors? For example: big five personality traits, grit, growth mindset, etc.</p>
	Tutor Recruitment	<p>How effective are tutors with different backgrounds?</p> <p>Are tutoring programs more effective when they require tutors to have a minimum level of education?</p> <p>Do entrance exams for tutors improve the effectiveness of a tutoring program?</p>
	Tutor Training & Support	<p>What type and amount of training do lower-skilled tutors need to improve student learning?</p> <p>Does the type of training tutors receive impact how effective tutoring programs are?</p> <p>Does ongoing support enhance the effectiveness of tutoring programs? Potential supports: ongoing professional development; check-ins with site leaders; tutor coaches.</p>
	Tutor Oversight: Accountability, Consequences, and Incentives	<p>Are tutoring programs more effective when they incorporate accountability measures for their tutors?</p> <p>Are tutoring programs more effective when they set consequential standards for their tutors such as the possibility of being fired?</p>



		<p>Are tutoring programs more effective when tutors' behaviors are incentivized?</p> <p>Are tutoring programs more effective when student learning gains are incentivized?</p>
Instruction	Delivery	<p>Can virtual tutoring be as effective as in-person tutoring? What is the marginal impact of in-person tutoring over virtual tutoring on student outcomes?</p> <p>What is the marginal impact of blended tutoring over virtual tutoring on student outcomes?</p> <p>What is the marginal impact of in-person tutoring over blended tutoring on student outcomes?</p> <p>Could synchronous tutoring sessions involve asynchronous instruction to effectively increase student learning?</p> <p>Does adding an asynchronous component to tutoring allow for a) employing lower-skilled tutors and b) tutors to increase their caseloads?</p>
	Session Frequency & Length	<p>What is the marginal benefit of increasing the number of hours students spend with tutors each week?</p> <p>Does the distribution of tutoring dosage matter? Could tutors and students meet for 2 hours, 1x per week and receive the same benefits as tutors and students who meet for 30 minutes, 4x per week?</p> <p>Does the ideal session frequency and length differ by age level and subject?</p>
	Duration	<p>What is the most effective duration for tutoring interventions to last?</p> <p>Does the ideal program duration differ by age level and subject?</p>

<p>Student-Tutor Ratio</p>	<p>What is the marginal change in student learning when the student-tutor ratio increases by an additional student (from 1:1 to 4:1)?</p> <p>Do smaller student-tutor ratios facilitate more positive tutor-student relationships?</p> <p>Do small group tutoring sessions facilitate more positive peer relationships than 1:1 tutoring sessions?</p> <p>Can small group tutoring sessions intentionally build more positive relationships between students?</p>
<p>Student Grouping</p>	<p>What is the most effective approach to grouping students for small group tutoring?</p>
<p>Tutor Consistency</p>	<p>Does having a consistent tutor for the duration of the intervention improve student outcomes?</p> <p>Does “looping” students with tutors across semesters or years improve student outcomes?</p> <p>Is “looping” more effective during major transitions (e.g., 8th to 9th grade)?</p>
<p>Tutor-Student Relationship</p>	<p>Does intentionally embedding relationship-building content in tutoring sessions improve student outcomes?</p> <p>Does matching tutors and students on demographics (e.g., race) improve student outcomes?</p> <p>What is the marginal impact of demographic matching when programs focus explicitly on relationship building?</p> <p>Does matching tutors and students on personality traits improve student outcomes?</p>
<p>Structure</p>	<p>Are tutoring programs with structured, sequenced approaches more effective?</p> <p>Do set routines make tutoring sessions more effective?</p>

	Quality of Materials	Do standards-aligned, educator-developed materials enhance the effectiveness of tutoring?
	Facilitation	Can lower-skilled tutors (e.g., volunteers, college students) be more effective when provided with facilitation tools and strategies?
Learning Integration	Setting	Are during-school tutoring programs more effective than out-of-school tutoring programs?  Does during-school tutoring result in more engagement among students than out-of-school tutoring?  Can “vacation academies” and “summer tutoring” be as effective as during-school tutoring?
	Displacement	What is the most effective way to schedule tutoring sessions during the school day?
	Take-up	Are tutoring interventions that require parents to opt-in their students less effective than those that are required or have an opt-out feature?  Are certain recruitment strategies more effective at getting sites (districts, schools) to enroll in tutoring intervention programs?
	Curricular Alignment	Are tutoring interventions that are intentionally aligned to a student’s classroom curriculum more effective than those that are not?  Does increasing the communications between tutors and students’ teachers improve student outcomes?
	School and Teacher Engagement	Are tutoring programs with school-based coordinators more effective?  Does more tutor-teacher interaction result in better student outcomes?
	Family Engagement	Does increased family engagement with the tutoring program improve student outcomes?

		<p>Does encouraging tutors to communicate with students' families improve family engagement?</p> <p>Does tutor-induced family engagement improve student outcomes?</p> <p>Can tutors facilitate the take-up of public benefits?</p> <p>Can tutors act as "navigators" for students' families, identifying supports and opportunities that advance student success?</p>
Data Use	Design and Improvement Process	To what extent does using data to design and improve tutoring programs associate with student outcomes?
	Formative Assessment	<p>Does encouraging tutors to use formative assessments to adapt their instruction improve student outcomes?</p> <p>Does training tutors to use formative assessments to adapt their instruction improve student outcomes?</p>
	Student Progress	<p>Does the quantity of data tutoring programs collect and provide to their tutors impact student outcomes?</p> <p>Does the quality of data tutoring programs collect and provide to their tutors impact student outcomes?</p>
	Program Evaluation	Are tutoring programs that engage in more rigorous program evaluations more likely to see student learning gains?
	Tutor Effectiveness	<p>Are tutoring programs more effective when they incorporate accountability measures for their tutors?</p> <p>Are tutoring programs more effective when tutors' behaviors are incentivized?</p> <p>Are tutoring programs more effective when student learning gains are incentivized?</p>
	Expanding Outcome Measures	Do tutoring programs improve outcomes beyond student test scores?

		<p>Does tutoring improve motivation, result in more positive relationships, encourage advanced course-taking, reduce absenteeism, reduce disciplinary actions, and lower dropout rates, etc?</p> <p>Does higher-dosage tutoring result in longer-term outcomes?</p> <p>How long do the effects from tutoring persist?</p> <p>How do different tutoring program characteristics contribute to long-term outcomes?</p>
	Tutor Benefits	<p>Do tutors experience positive social and career-related benefits?</p> <p>Can tutoring programs intentionally promote better outcomes for tutors through curated additional programming?</p>

## Research Priority #2: Implementing Tutoring at Scale

In addition to identifying the characteristics of effective tutoring, researchers will need to better understand how to scale tutoring programs in order to reach the greatest number of students in need. With this goal in mind, we have begun to articulate a second line of suggested research questions that broadly address issues such as: Where will the supply of tutors come from? What is the demand for high-impact tutoring? And how are tutoring programs most successfully implemented?

In an effort to know that we are on the right track — that we are articulating the right questions — we start by developing questions regarding how to scale high-dosage, school-driven tutoring focused on a specific subject and grade. Our next step will involve determining the best approaches and strategies for answering these initial questions.

### Supply and Demand Questions

To implement tutoring at scale, we need to understand (1) where the supply of tutors would come from and (2) the extent to which districts want and are capable of implementing tutoring initiatives. This section highlights some of the key questions and associated considerations that will inform whether and how tutoring can reach students across the United States.

Question	Considerations
What skills do tutors need?	<ul style="list-style-type: none"> <li>■ Subject matter skills (may require a</li> <li>■ screening tool, particularly for math)</li> <li>■ Grade level content knowledge (e.g., high</li> <li>■ school math requires more initial</li> <li>■ competency than early literacy)</li> <li>■ Ability to connect with students</li> </ul>
When are tutors available?	<ul style="list-style-type: none"> <li>■ Full-time vs. part-time</li> <li>■ During the school year vs. summer</li> <li>■ Weekday vs. weekend</li> <li>■ During the school day vs. out-of-school time</li> </ul>
How can the tutoring be delivered?	<ul style="list-style-type: none"> <li>■ In-person vs. Virtually vs. Blended <ul style="list-style-type: none"> <li>■ In-person: geographic constraints</li> <li>■ Virtual/Blended: internet access and technical skills</li> </ul> </li> <li>■ Age of student and subject (e.g., early literacy may be best suited for in-person tutoring)</li> </ul>
What types of supervision do tutors need?	<ul style="list-style-type: none"> <li>■ Capacity of supervisors</li> <li>■ Frequency of supervision</li> <li>■ Supervisor-tutor caseload</li> </ul>
Who can be tutors?	<p><i>More promising:</i></p> <ul style="list-style-type: none"> <li>■ Current college students</li> <li>■ Pre-college gap year</li> <li>■ Recent college grads</li> <li>■ AmeriCorps members</li> <li>■ Retirees</li> <li>■ Current teachers</li> </ul> <p><i>Less promising; may be better suited for location-based early literacy tutoring:</i></p> <ul style="list-style-type: none"> <li>■ Graduate students</li> <li>■ Education majors</li> <li>■ Currently employed paraprofessionals</li> <li>■ Unemployed paraprofessionals</li> <li>■ Not in labor force paraprofessionals</li> </ul>
What is the best way to recruit tutors?	<ul style="list-style-type: none"> <li>■ Work-study through universities</li> <li>■ Social media advertising</li> <li>■ Entry-level professional work (e.g., EdBoost)</li> <li>■ Compensated service (e.g., AmeriCorps)</li> </ul>
How many students can a tutor handle?	<ul style="list-style-type: none"> <li>■ # of students per tutoring session</li> <li>■ # of students overall</li> <li>■ Tutor skill level</li> <li>■ Full-time vs. part-time</li> </ul>

How much do tutors cost?	<ul style="list-style-type: none"> <li>■ Skill level</li> <li>■ Recruitment</li> <li>■ Training</li> <li>■ Salary/Wages</li> <li>■ Supervision</li> <li>■ Coordination</li> <li>■ Overhead</li> </ul>
Where is the money for tutoring coming from?	<ul style="list-style-type: none"> <li>■ Federal Education Acts (particularly those with deadlines for usage, e.g., CARES plus follow-up)</li> <li>■ 21<sup>st</sup> Century Community Learning Centers (non-school hours only; federal funds can go to tutoring providers directly)</li> <li>■ Title I funds</li> <li>■ Title II funds (for building teacher pipeline)</li> <li>■ Title III funds (for English Language Learners)</li> <li>■ Foundations</li> </ul>
<p>What are the constituencies that might want high dosage tutoring?</p> <p>What stakeholders at different levels need to be brought in to make high-dosage tutoring happen?</p>	<ul style="list-style-type: none"> <li>■ School board members</li> <li>■ Superintendent and other members of the district leadership team</li> <li>■ Principals</li> <li>■ Teachers</li> <li>■ Teacher unions</li> <li>■ Parents</li> <li>■ Students</li> <li>■ Local non-profits</li> <li>■ Local organizations: Community centers, Day camps, Churches</li> <li>■ Other district partners</li> <li>■ Educator pipeline advocates</li> </ul>
How does state or district governance shape demand for tutoring?	<ul style="list-style-type: none"> <li>■ Existing programs</li> <li>■ Alternate options</li> </ul>
Which students receive tutoring?	<ul style="list-style-type: none"> <li>■ Need-driven vs. Curriculum-driven vs. Universal</li> </ul>

### Implementation Questions

The goal of our implementation questions is to guide researchers in building a deeper understanding of how district leaders, school leaders, teachers, parents, and students experience a tutoring initiative. Each question addresses either the motivations, opportunities, logistical needs, or barriers for delivering tutoring to students. Answering each of these individual questions will ultimately help answer the overarching question: What are the necessary conditions required to effectively implement a high-dosage, school-driven tutoring program?

## District-Level Questions

- What are the district-level constituencies that might want high dosage tutoring (School board members; superintendent and other district leadership team members; principals; teachers; teacher unions; parents; local non-profits; other district partners? What are their motivations/ pressing concerns/ interests?)
- What stakeholders at different levels need to be bought in to make high-dosage tutoring happen?
- How do varieties of district governance shape demand for tutoring?
- What are the biggest barriers to bringing high-dosage tutoring to a district? Political? Institutional? Logistical? Others?
- What sorts of funding sources would districts require to make high-dosage tutoring feasible?
- Who at the district level will be responsible for overseeing tutoring partners to assess quality of performance?
- How are schools and districts going to make time in their daily schedules for high-dosage tutoring?
- Who will manage the logistics, including scheduling, transportation, finding space, etc.?
- What type/level of support could tutoring providers give to school districts to bolster implementation?
- What type/level of support could tutoring providers give to classroom teachers to bolster implementation?
- How do districts weigh the goals of tutors vs. diversifying and growing the teacher/paraprofessional pipeline?
- How important is it that tutors come from the local area?
- How do these considerations vary if tutoring were to be implemented out-of-school time (e.g., after school, during vacations, during the summer)?

### *Contract-Specific Issues:*

- How to monitor contracts to hold tutoring providers accountable for results? (i.e., Frequency of reporting? Parent reporting?)
  - How to craft effective RFPs for tutoring work?
  - What systems are needed to easily organize and share relevant data, while still navigating various student privacy and other concerns (e.g., FERPA, IEPs, etc.)?
  - How to identify a pool of potential tutoring providers who could compete for district contracts? (Broad advertising, RFPs, etc.)



### School Administration- Level Questions

- How will student outcomes be measured?
- How do schools select who will receive high-dosage tutoring?
- What is the role for the guidance counseling staff in implementing high quality tutoring?
- What is the role for the principal in implementing high quality tutoring?
- Who at the school will be the liaison with the district in helping contribute to oversight of a tutoring provider?
- Does every school need a “site coordinator,” and where could these coordinators come from?
- Does the tutoring provider need to offer school-specific “site coordinators” that work with teachers or other school staff?
- What type/level of support could tutoring providers give to schools to bolster implementation?

### Teacher-Level Questions

- How does high-dosage tutoring change the job of the teacher? Discarding some current tasks? Taking on some new tasks?
- How can high-dosage tutoring be designed so it is helpful, not burdensome to teachers?
- How can we encourage regular communication between tutors and teachers without burdening teachers?
- What role will teachers play in contributing to progress monitoring that feeds into school-level and then district-level efforts to hold tutors accountable for performance?
- What type/level of support could tutoring providers give to classroom teachers to bolster implementation?
- What type/level of support could tutoring providers give to families to bolster implementation?

### Parent-Level Questions

- What type/ level of support could tutoring providers give to families to bolster implementation?
- How do parents feel that an in-school, high-dosage tutoring format would impact their child’s learning experience?

- Are parents comfortable with tutoring being incorporated into the school day, or would they prefer that additional academic assistance come outside of the normal school period?
- What types of services and qualifications would parents desire from the tutoring company that their child's school district hires?
- How much interaction would parents like to have with their child's tutor in terms of keeping track of the child's progress (up to the tutor or mandated by the school itself)?
- Do parents have any concerns about implementing a high-dosage tutoring system into their child's school day? If so, what are they?

### Student-Level Questions

- How would in-school tutoring provide students with beneficial academic assistance? Do students think that learning while still in school but from someone else (not the classroom teacher) could be helpful?
- What are some areas of assistance that students would like a tutor to help with? (Catching up on previous grade-level material, reinforcing current material, overall mentorship/ academic success skill help, etc.)
- Does learning in group or individual settings work better for individual students, and when learning remotely, does this change?
- If receiving assistance from a tutor, what would students like their tutor's role to be in relation to their normal academic instruction with their teachers? Would students like their tutor to work in tandem with their teacher as a team to develop a curriculum that fits their needs, or would they rather work solely with their tutor on extra practice?
- Would having extra academic assistance from anyone be beneficial, regardless of whether it is a new tutor or a teacher? (i.e. is it the current method of instruction with which the student is struggling, or is it the concepts themselves that the student wants to improve on overall)
- What are some qualities that students would look for in a tutor?
- Have students worked with any tutor/received additional academic support in the past (pull out classes, after school help, enrichment classes, summer school, etc.)? If so, how did students find those experiences worked for them?
- Would an in-school tutoring format work well for students, and/ or would a summer program work better?

## REFERENCES

- Baye, A., Inns, A., Lake, C., & Slavin, R. E. (2019). A synthesis of quantitative research on reading programs for secondary students. *Reading Research Quarterly, 54*(2), 133-166.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in education: principles, policy & practice, 18*(1), 5-25.
- Bergman, P., Lasky-Fink, J., & Rogers, T. (2020). Simplification and defaults affect adoption and impact of technology, but decision makers do not realize it. *Organizational Behavior and Human Decision Processes, 158*, 66-79.
- Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289-328.
- Borges, D. R., & McCarthy, M. P. (2015). *Leduc Center for Civic Engagement: America Reads\* Counts Program Evaluation, 2014-15*. Public Policy Center, UMass Dartmouth.
- Carlana, M., & La Ferrara, E. (2021). Apart but Connected: Online Tutoring and Student Outcomes during the COVID-19 Pandemic.
- Cavell, T. A., Mutignani, L. M., Alfonso, L., & Marie Smith, A. (2020). Attachment Tendencies, Big 5 Personality Traits, and Self-Efficacy as Predictors of Mentors' Relationships with Aggressive Children. *American journal of community psychology, 66*(1-2), 130-143.
- Cook, P. J., Dodge, K., Farkas, G., Fryer, R. G., Guryan, J., Ludwig, J., Mayer, S., Pollack, H., & Steinberg, L. (2015). Not too late: Improving academic outcomes for disadvantaged youth. *Institute for Policy Research Northwestern University Working Paper WP-15, 1*.
- Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A.-M. (2017). Academic Interventions for Elementary and Middle School Students With Low Socioeconomic Status: A Systematic Review and Meta-Analysis. *Review of Educational Research, 87*(2), 243-282. <https://doi.org/10.3102/0034654316687036>
- Fitzgerald, J. (2001). Can minimally trained college student volunteers help young at-risk children to read better? *Reading Research Quarterly, 36*(1), 28-46.
- Fitzgerald, J., Morrow, L. M., Gambrell, L., Calfee, R., Venezky, R., Woo, D. G., & Dromsky, A. (2002). Federal policy and program evaluation and research: The America Reads example. *Literacy Research and Instruction, 41*(4), 345-370.
- Fryer Jr, R. G. (2014). Injecting charter school best practices into traditional public schools: Evidence from field experiments. *The Quarterly Journal of Economics, 129*(3), 1355-1407.

- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., DeSelms, J., Seethaler, P. M., Wilson, J., & Craddock, C. F. (2013). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology, 105*(1), 58.
- Gershenson, S., Hart, C., Hyman, J., Lindsay, C., & Papageorge, N. W. (2018). *The long-run impacts of same-race teachers* (0898-2937).
- Gersten, R., Rolfhus, E., Clarke, B., Decker, L. E., Wilkins, C., & Dimino, J. (2015). Intervention for first graders with limited number knowledge: Large-scale replication of a randomized controlled trial. *American educational research journal, 52*(3), 516-546.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900-1902.
- Heinrich, C. J., Burch, P., Good, A., Acosta, R., Cheng, H., Dillender, M., Kirshbaum, C., Nisar, H., & Stewart, M. (2014). Improving the implementation and effectiveness of out-of-school-time tutoring. *Journal of Policy Analysis and Management, 33*(2), 471-494.
- Hill, A. J., & Jones, D. B. (2018, 2018/06/01/). A teacher who knows me: The academic benefits of repeat student-teacher matches. *Economics of Education Review, 64*, 1-12.  
<https://doi.org/https://doi.org/10.1016/j.econedurev.2018.03.004>
- Jacob, R., Armstrong, C., & Willard, J. (2015). Mobilizing volunteer tutors to improve student literacy: Implementation, impacts, and costs of the Reading Partners Program.
- Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational researcher, 49*(4), 241-253.  
<https://doi.org/10.3102/0013189x20912798>
- Markovitz, C. E., Hernandez, M. W., Hedberg, E. C., & Silbergitt, B. (2014). Impact Evaluation of the Minnesota Reading Corps K-3 Program. *Corporation for National and Community Service*.
- Mattera, S., Jacob, R., & Morris, P. (2018). Strengthening children's math skills with enhanced instruction: The impacts of Making Pre-K Count and High 5s on kindergarten outcomes. *New York: MDRC, March*.
- Neitzel, A., Lake, C., Pellegrini, M., & Slavin, R. (2020). A synthesis of quantitative research on programs for struggling readers in elementary schools. *Manuscript submitted for publication*.
- Nickow, A., Oreopoulos, P., & Quan, V. (2020). The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence. *National Bureau of Economic Research Working Paper Series*(w27476).
- Parker, D. C., Nelson, P. M., Zaslofsky, A. F., Kanive, R., Foegen, A., Kaiser, P., & Heisted, D. (2019). Evaluation of a math intervention program implemented with community support. *Journal of Research on Educational Effectiveness, 12*(3), 391-412.
- Pellegrini, M., Lake, C., Neitzel, A., & Slavin, R. (2020). Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis. *Manuscript submitted for publication*.

- Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research, 79*(1), 3-38.
- Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review, 17*(4), 327-362.
- Roorda, D. L., Koomen, H. M., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students’ school engagement and achievement a meta-analytic approach. *Review of Educational Research, 81*(4), 493-529.
- Roschelle, J., Cheng, B. H., Hodkowski, N., Neisler, J., & Haldar, L. (2020). Evaluation of an Online Tutoring Program in Elementary Mathematics. *Online Submission*.
- Schacter, J. (2000). Does individual tutoring produce optimal learning? *American educational research journal, 37*(3), 801-829.
- Schumaker, J. B., Deshler, D. D., & McKnight, P. (2002). Ensuring Success in the Secondary General Education Curriculum Through the Use of Teaching Routines.
- Sirinides, P., Gray, A., & May, H. (2018). The Impacts of Reading Recovery at scale: Results from the 4-year i3 external evaluation. *Educational evaluation and policy analysis, 40*(3), 316-335.
- Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*(4), 1391-1466.
- Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., Yen, L., Fuchs, L. S., & Compton, D. L. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational evaluation and policy analysis, 30*(4), 368-388.
- Worthy, J., Prater, K., & Pennington, J. (2003). “It’s a Program That Looks Great on Paper”: The Challenge of America Reads. *Journal of Literacy Research, 35*(3), 879-910.
- Zimmer, R., Hamilton, L., & Christina, R. (2010). After-school tutoring in the context of no child left behind: Effectiveness of two programs in the Pittsburgh public schools. *Economics of Education Review, 29*(1), 18-28.