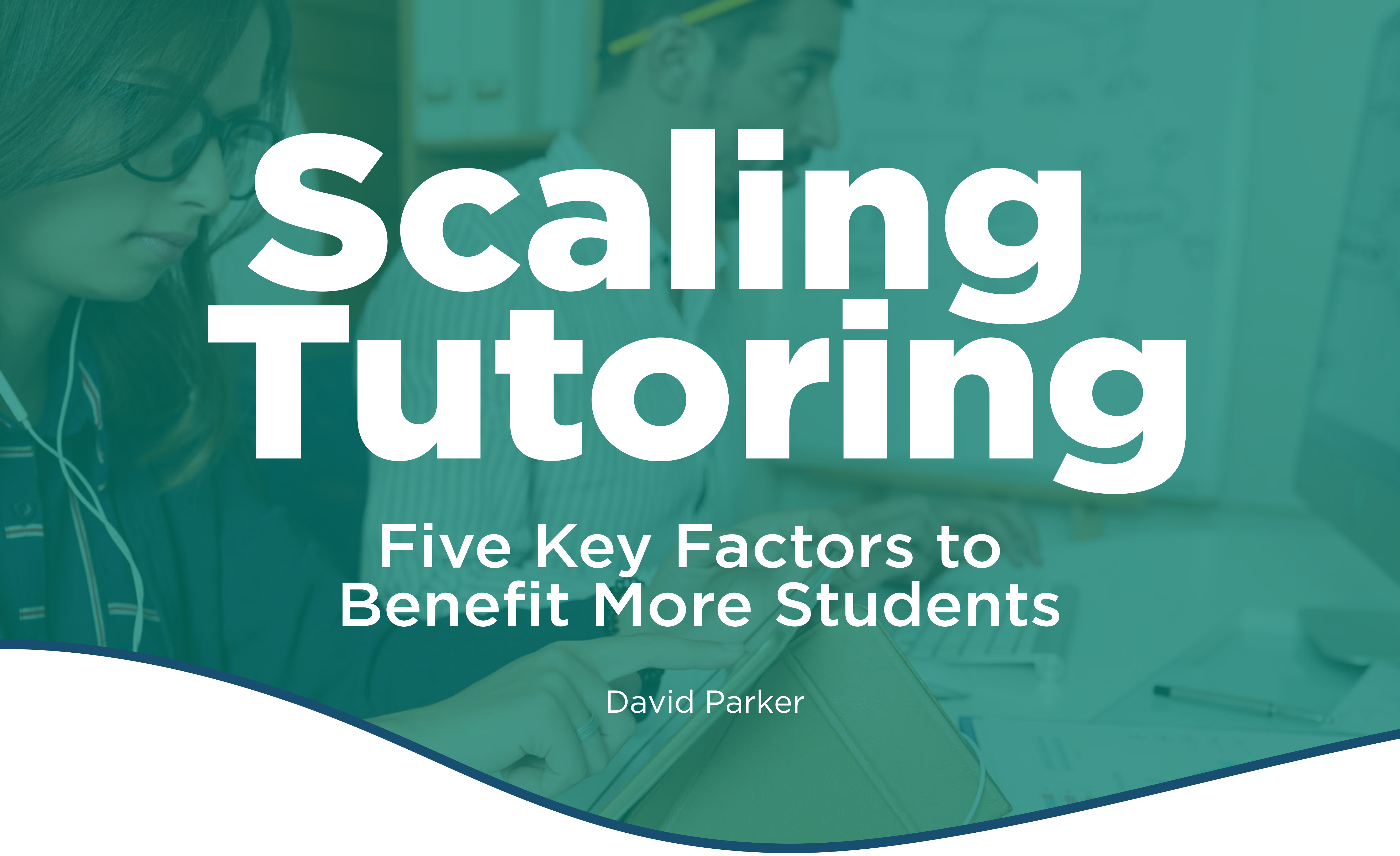


Scaling Tutoring

Five Key Factors to Benefit More Students

David Parker

ACCELERATE LEARNING

IDB

ACCELERATE LEARNING

Scaling Tutoring

Five key factors to benefit more students

David Parker Ph.D.
Vice President of Research and Development at ServeMinnesota

The preparation of this brochure was supported and reviewed by the following members of the Education Division of the Inter-American Development Bank: **Andrea Prieto, Gonzalo Almeyda.**

Illustration and graphic design: **Juan Sebastián Fonseca**

Introduction

When it comes to tutoring, “no other intervention in all of educational research has a comparable impact” (p. 5, Slavin, Madden, Neitzel, & Lake, 2020). Accordingly, education advocates across the world have called for scaling tutoring to help address learning inequities exacerbated by the global pandemic. The apparent optimism for scaling tutoring rests on two assumptions. First, that tutoring at scale will remain effective. Second, that tutoring at scale is feasible.

This paper describes five key factors that are essential to scaling tutoring, each of which must be thoroughly considered to ensure tutoring models remain effective as they scale. They are not an exhaustive list, but each factor represents a core element of tutoring that coordinated efforts across practice and policy must consider to ensure tutoring delivers on its promise. Without them, even successful scaling efforts will ultimately fail.

Fortunately, scaled models for tutoring exist. Every school year, a single tutoring program serves over 50,000 students, coast-to-coast across the United States (see Ampact.us). Each student receives evidence-based reading or math instruction tailored to their needs. Each student has data collected about their tutoring exposure and learning. Each student is making progress toward recouping a half-to-full year of learning. It may be that no other tutoring program in the world serves that many students with that level of effectiveness.

The program leverages AmeriCorps as a national infrastructure to activate local resources. Community members complete a year of service under AmeriCorps, are trained and supported via a centralized implementation apparatus, and complete their service term in their local community’s schools. The program achieves its powerful effects via its application of proven educational innovations, such as using data to make student-centered decisions and using instructional methods known to accelerate learning.

Such features distinguish the program from other tutoring models, some of which have scaled, some of which have evidence, and some of which are economically accessible to all children and schools, but none of which have achieved all three. Many look different in different schools, communities, or cities. Others attempt to support such a wide variety of learning outcomes—from basic reading skills to advanced coursework in science—that they fail to deliver effective, consistent support to students. Most lack evidence of impact on the outcomes they purport to target, while still others use unproven instructional methods. Almost all come with costs that burden families, educators, or schools. In sum, most tutoring initiatives fail to meet basic assumptions of effectiveness and feasibility and thus have limited potential for realizing the great promising of tutoring delivered at scale (Nickow, Oreopoulos, & Quan, 2020).

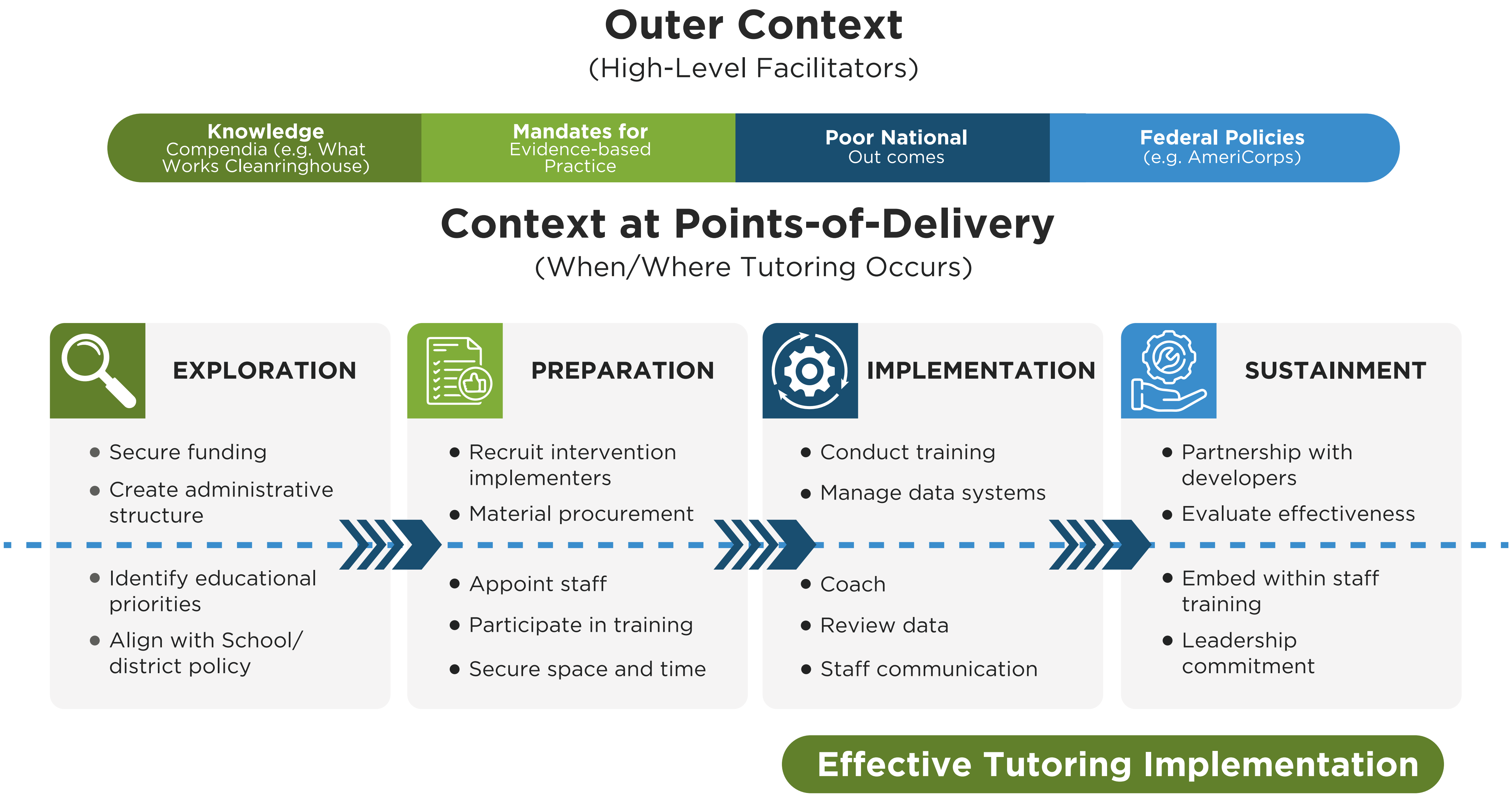
The five key factors described here can be organized into contextual and practical factors. Two contextual factors—(1) a well-specified theory of change and (2) a dedicated organizational structure—influence the way in which tutoring is delivered (referred to here as "points of delivery"). Of the five factors, they are the most influenced by policy. They do not reflect a comprehensive view of all policy-related considerations for scaling tutoring (see chapter 5 for more information). Rather, they must be considered as facilitators for the practical realities of tutoring. The three practical factors—(3) using evidence-based instructional methods, (4) making data-based decisions, and (5) programming for accurate model delivery—are core elements of tutoring that are nonnegotiable for maximizing learning.

This brief follows a two-part structure. The first part discusses problems that each factor is designed to address. The second part focuses on solutions, providing additional descriptions and examples for each factor.

Problem

The challenge behind scaling effective innovations like tutoring is so complex and sophisticated that an entire scientific discipline dedicates itself toward its study. Called implementation science, it is a necessarily broad field of inquiry that channels the methods of scientific inquiry into understanding processes and mechanisms that will lead to improved adoption of evidence-based practices (Eccles & Mittman, 2006). Figure 1 captures a cross-section of this complexity.

Figure 1: Cross-section of factors and processes that influence effective implementation of evidence-based practices like tutoring.



The figure incorporates one of several research-based models (Aarons, Hulburt, & Horwitz, 2011) that have emerged from implementation science to make explicit the phases that are necessary to bring an effective innovation into practice and sustain it over time. Such models recognize that effective implementation tends to follow phases, from exploration and preparation, to actual implementation and sustainment. At the same time, the figure recognizes the practical implications of multiple levels of context, including policy-level facilitators that establish public awareness and activate resources, as well as a diverse set of considerations that are proximal to the points of delivery, each of which contributes to effective implementation and outcomes. The five key factors for scaling tutoring recognize and address much of this complexity. Table 1 presents a brief overview of each factor, why it is important, and how it aligns with implementation science.

Table 1: Five key factors for successfully scaling tutoring programs.

Factors	Why is it critical for tutoring?	How does it align with Implementation Science?
	Contextual Factors	
Require a Well-Specified Theory of Change	**It enables a program to be feasibly implemented and productively evaluated at scale.**	**It is a core assumption for any evidence-based innovation.**
Use a Dedicated Organizational Structure	**It is unlikely that existing points of delivery can successfully implement tutoring programs independently**	**Implementation science recognizes that organizational partnerships are essential for scaling effective innovations.**
	Practice Factors	
Use Evidence-based Instructional Methods	**Learning science has clearly identified instructional processes that work.**	**Factors like high-quality materials, training, and coaching should focus on these methods.**
Make Data-based Decisions	**Data enable optimal decision-making.**	**Without good data, implementation is unknown and evaluation is impossible.**
Program for Accurate Delivery	**Ongoing support is necessary for desired implementation and outcomes**	**This is a direct recommendation from the implementation science literature base.**

Well-Specified Theory of Change

A precise and well-defined theory of change is a core assumption for effective implementation. The theory of change drives what the program focuses on, who does what with whom, and how well the program outcomes can be measured. For example, in Figure 1, the logistical requirements of material procurement, training, data systems, and coaching all necessitate a specific theory of change; in an unspecified theory of change the precise details for implementation remain unknown. Evaluating effectiveness is also rendered difficult – and sometimes impossible - when lack of specificity complicates how a program defines and measures outcomes. Table 2 lists distinctions between an unspecified theory of change and a well-specified theory of change using literacy tutoring to illustrate.

Table 2: Distinctions between unspecified and well-specified theories of change for literacy.

	Unspecified theory of change	Well-specified theory of change
Tutoring Focus	**Reading**	**Phonics and fluency**
Target Participants	**Children**	**Early primary school students.**
Training Focus	**Tips on how to engage with children**	**How to deliver systematic, explicit phonics instruction**
Coaching	**Call-center for tutor questions**	**Live observations, monthly**
Data	**Extant (e.g., school-collected)**	**Reliable and valid assessments for foundational reading skills**
Evaluation	**Parent satisfaction surveys**	**Improved phonics and fluency skills measured against essential benchmarks**

The practical implications between an unspecified and well-specified theory of change are profound. In this example, any component of literacy for students of any age could be the focus of the unspecified tutoring model and yet would still align with the focus on "reading". A corresponding tutoring initiative operating at scale could expect implementation partners to interpret their tutoring initiative differently. In one site, they might adopt shared book reading with preschool children. In another, they might support high school literature. When these discrepancies emerge, the tutoring program will fail to meet expectations whether or not it has scaled. The tutoring initiative has become a distribution channel for generic educational resources loosely organized under "reading". A similar scenario could be envisioned for an environmental initiative that seeks to "address climate change", an important and aspirational goal that is accessible to public opinion, but one that is impossible to scale effectively when it could legitimately include focus areas as diverse as public transportation investments to rainforest preservation to clean water initiatives.

Dedicated Organizational Structure

Schools around the world tend to have broad mandates: keep children safe, foster a joy for learning, develop social-emotional skills, teach core academic skills, build knowledge across domains (i.e., history, social studies, science). In addition, many schools have to address societally-mandated requirements, such as specializing education for disabled students, developing sports and other enrichment programs, and serving as community-centers. Despite lofty expectations, resources tend to lag and staff are stretched thin.

For these reasons, implementation scientists have identified the need for intermediary organizational structures to support implementation for evidence-based innovations (Franks & Bory, 2015). Such organizations are designed to offset the logistical demands and expertise requirements that innovations need. They can help with model development, quality assurance, ongoing improvement efforts, evaluation, training, and policy advocacy. They may also be better positioned to hold ultimate accountability for implementation and outcomes, depending on how policy and funding structures are established. As contrasted with other approaches to implementing evidence-based innovations--which rely on assumptions that practitioners read, care and translate into practice innovations that are described in scientific journals (Greenhalgh, Robert, MacFarlane, Bate, and Kyriakidou, 2004)—intermediary organizations help ensure action.

The tutoring program described at the beginning of this article is an intermediary organization that reduces the capacity demands that tutoring programs require of schools and staff. The organization behind the program develops the tutoring methods, recruits all the tutors, creates all materials and training content, provides the coaching, develops the data systems, and supports all program evaluation. Centralizing each of those requirements enacts multiple economies of scale. One set of evidence-based materials for reading can be used in thousands of schools. One training module for how to interact with students in a culturally-responsive way can be viewed by thousands of tutors. One data system can house data for tens of thousands of students and hundreds of thousands of data points.

Evidence-based Instructional Methods

Much is known about how children learn. Explicit instruction to build vocabulary knowledge builds students' overall understanding of words and their meaning (Neuman & Marulis, 2010). Building conceptual understanding of rational mathematical concepts (e.g., fractions) supports overall math development (Butler, Miller, Crehan, Babbitt, & Pierce, 2003). Facilitating practice opportunities and providing feedback accelerates learning of discrete skills, from word learning to math facts (e.g., Varma & Schleisman, 2014). Instructional methods that leverage this knowledge are effective. They are considered evidence-based.

Those that are not evidence-based will be less effective. For example, sitting down with a student and reading them a story does not promote learning. Doing so repeatedly over many years may have an empirical association with better literacy outcomes, but the causal mechanisms are scientifically unknown and thus the associated outcomes are more likely an artifact of the familial and community resources those children access throughout their school years. Many variables must align—from parental employment to community infrastructure—for a child to be read a book every day. Tutoring initiatives that misinterpret the empirical association between home reading practices and increased learning as a need to provide access to books would fail. When compared with an evidence-based instructional method such as dialogic or shared book reading (Mol, Bus, deJong, & Smeets, 2008), the advantage is clear. This point is similarly true across academic learning domains. Word problem solving using a simple and consistent schema-based approach is unquestionably superior to an ad hoc approach (Peltier & Vannest, 2017). Thus, any tutoring initiative seeking to scale should employ known and evidence-based instructional methods.

Data-based Decision-Making

The importance of data within tutoring initiatives becomes apparent the moment a decision needs to be made. Every decision—from which students need tutoring to determining if the tutoring program works—is made better by using data. See Table 3 for a list of the questions that data answer and issues that emerge if data are unavailable.

Table 3: Key questions about tutoring and what happens if they are answered without data.

Key Tutoring Questions	What Happens Without Data
Which students need tutoring?	**Students that need support don't get it. Students that don't need support get it. Students with the strongest advocates get support (furthering educational inequities).**
What should their tutoring focus on?	**Concepts and skills that are already acquired. Concepts and skills that are too difficult. Concepts and skills that are less optimal to overall learning (e.g., instruction on decimal placements vs. understanding fractions).**
Are they improving? How much?	**Improvement is unknown. Anecdotes or intuition dictate changes. A need to make a change is missed.**
Can they stop receiving tutoring?	**Students stay in tutoring too long. Students are not in tutoring long enough. Attainment of benchmarks/goals is unknown.**
Was tutoring beneficial?	**Ineffective tutoring programs scale. Effective tutoring programs do not scale.**

In the absence of data, decisions about tutoring are subject to inaccuracy, bias, and insufficient detail. In most communities, some students need tutoring more than others, and this is particularly important to consider when resources do not allow for tutoring every student. Objective, simple, and accurate data help identify who is eligible and their degree of need. Relatedly, data help avoid scenarios in which adult biases or perceptions lead them to over- or under-advocate for certain students. Data allow the nature and type of each student's need to inform equitable decisions about accessing tutoring.

Another decision-point in tutoring concerns what content to focus on with students identified for tutoring. Even within a well-specified theory of change, data are needed to clarify and optimize the instructional focus for individual students. In math, for example, tutoring may be indicated for a student to improve conceptual understanding and proficiency in working with whole and rational numbers (National Mathematics Advisory Panel, 2008), but that is insufficiently diagnostic for a tutor to begin instruction. A student could need support in understanding operations (e.g., subtraction or multiplication), a whole-number concept (e.g., regrouping during addition), or a rational-number concept (e.g., multiplying fractions). Data help tutors make that decision.

Program for Accurate Delivery

Tutoring, whether remote or in person, presumes that an adult will interact with a student. It is also presumed that those interactions will be effective and lead to improved learning outcomes for the student. As noted above, the instructional methods used in those interactions are what yield improved learning. Unfortunately, research is clear that adults do a poor job of implementing new instructional methods. In fact, the vast majority—up to 95%—of trainees on a new instructional routine are not expected to use it after training (Joyce & Showers, 2002). Conversely, when provided regular, ongoing support up to 95% will use the new method for a sustained period of time. The implications of this research are clear. Training is important, but it is not the agent of change that many assume it to be. Coaching and ongoing support is.

The implication for scaled tutoring initiatives is that for the desired outcomes to be achieved, investments in personnel and ongoing support and coaching must be made. This is likely a significant hidden and unexpected cost consideration, but it would be like purchasing a car without thinking about fuel and maintenance costs. Any pragmatic analysis of the facts makes it clear that ongoing resources are necessary to realize the goals.

Solutions

The five factors described in this chapter work in concert to ensure scaled tutoring programs remain effective. A well-specified theory of change forms the blueprint for all aspects of the tutoring model while ensuring focus and consistency across geographies. It prescribes what evidence-based instructional methods to select, what data to use, what training and support structures are needed, and ultimately how the dedicated organizational structure will operate. The evidence-based instructional methods and data for decision-making are reciprocal in the sense that the data inform what instructional methods to use and also how students progress. Finally, accurate model delivery is best focused on each tutor's delivery of evidence-based instructional methods and the collection of accurate assessment data. Table 4 below summarizes how to effectively address each of factor.

Table 4: Summary of how to address each factor and corresponding expectations.

Factors	How to address the factor successfully.	What can be expected if successfully addressed.	Important Consideration(s)
		Contextual Factors	
Require a Well-Specified Theory of Change	**Precisely describe who, what, when, where, and how for the tutoring model.**	**A tutoring program that will improve student outcomes in a measurable way.**	**It's better to have multiple well-specified tutoring programs than one unspecified program.**
Use a Dedicated Organizational Structure	**Coordinate local and regional policy investments to build a single implementation organization.**	**Local partners need (and will appreciate) the dedicated expertise and capacity.**	**Starting with early-adopting local partners will naturally build scaling momentum.**
		Practice Factors	
Use Evidence-based Instructional Methods	**Anchor instructional methods in empirically-rigorous research studies; not expert opinion**	**Greater student growth than if unproven instructional methods are used.**	**These methods do not preclude or prevent strong student-tutor relationships from developing.**
Make Data-based Decisions	**Build data systems and use technically-adequate educational assessments**	**More efficient and effective decision-making**	**Optimal data systems allow tutors—and their coaches—to access student data in real time.**
Program for Accurate Delivery	**Plan to invest time and money in ongoing coaching for tutors**	**Effective implementation of the tutoring program**	**The cost for ongoing support make economic sense in light of stronger outcomes.**

Well-Specified Theory of Change

Perhaps the most significant barrier to ensuring tutoring programs have a well-specified theory of change is the fact that most students have concomitant educational needs, even within an academic skill area. A student might struggle in math, but within math they may struggle with conceptually understanding concepts like adding fractions, applying geometric principles, conducting data analysis, and completing word problems. Complicating matters further is the fact that the student may avoid completing their math homework. It is tempting in these scenarios to construct tutoring programs that provide "on-demand" homework assistance with whatever the student needs to complete in the moment. Such an approach would be understandable but it would fail.

Experts recommend calibrating tutoring initiatives toward 'keystone' skills or concepts (e.g., National Mathematics Advisory Panel, 2008), and doing so aligns with a need to have a well-specified theory of change. In math, such skills form a foundation for building other skills and thus contribute greatly to future math learning. In the brief list of math struggles mentioned in the paragraph above, conceptually understanding what it means to add fractions is a keystone skill that benefits all the others. Adding any lengths of shapes that are not whole numbers, analyzing data in percentages, and solving word problems that involve fractions would all be facilitated by a stronger understanding of adding fractions.

A well-specified theory of change directly informs details to a degree that is necessary to both implement the program effectively and achieve the desired results. Simply put, a well-specified theory of change will also address questions related to What, Who, Where, When, and How? Table 5 illustrates how the well-specified theory of change described earlier leads to clear specification of the necessary details of a reading tutoring program.

Table 5: Example of how a well-specified theory of change informs implementation details for a tutoring program.

Well-specified theory of change	
Tutoring Focus	**Phonics and fluency**
Target Participants	**Early primary school students.**
Training Focus	**Systematic, structured phonics instruction; Reading fluency**
Coaching	**Live observations, monthly**
	Reliable and valid assessments for foundational reading skills
Evaluation	**Improved phonics and fluency skills measured against essential benchmarks**

Example details of well-specified program	
What?	**Systematic explicit phonics instruction and practice (Ehri et al., 2001). Repeated Reading fluency instruction (Suggate, 2016)**
Who?	**Students—Early primary school students eligible based on teacher recommendation and scores <25%ile on curriculum-based phonics measure. Tutors—paid volunteers**
Where?	**At school.**
When?	**During the school day; when student is not engaged in teacher-led instruction. Daily 20 min sessions**
How?	**Virtual sessions over zoom. Tutor trained in three, 2hr training sessions before school year. Tutor coached once a month via staff member joining zoom session.**

Dedicated Organizational Structure

A dedicated organizational structure should achieve one simple goal: ensure the evidence-based tutoring program is implemented successfully across all sites. The need for a dedicated organization to accomplish that goal is made clear by the deceptive complexity that can arise at points of delivery. For example, funding, readiness for implementation, internal leadership, content knowledge, trust and understanding in evidence, availability of implementation technology are all significant barriers that each delivery point for tutoring needs to address (Franks & Bory, 2015). Yet each delivery point for tutoring needs to also fulfil their primary, pre-existing requirements. If that is a school, the school must fulfil all of its local mandates. If it is another community resource (e.g., library, public center) it must fulfil those duties.

Implementation scientists in healthcare and education have recently worked to identify extensive lists of strategies that delineate and organize how evidence-based innovations are implemented. The Expert Recommendations for Implementing Change (ERIC) project in healthcare has been converted into the list of School Implementation Strategies, Translating ERIC Resources (SISTER) for education (Cook, Lyon, Locke, Waltz, & Powell, 2019). Of the 75 identified strategies that implementers need to account for, none were rated as more than "moderately feasible", and most were only "somewhat feasible" (Lyon et al., 2019). Table 6 shows a sample listing of 10 of the SISTER strategies, its unaided feasibility rating, and the contribution that intermediary organization can make.

Table 6: Contribution of intermediary organization for implementing evidence-based strategies.

	SISTER Strategy	Mean Feasibility w/o Intermediary Organization[a]	Contribution of Intermediary Organization (Specific Example from Tutoring Program)
1	Asses for readiness and identify barriers/ facilitators	3.21	Assesses implementation capacity as part of application. (e.g., Can at least one staff member provide day-to-day support to tutor?)
2	Audit and provide feedback	3.14	Collects implementation data; review reports with school administrators. (e.g., Share implementation dosage and fidelity data with principal.)
3	Conduct cyclical small tests of change (Pilot or trial practice)	2.91	Monitors data and adapts program for fit and scale within each school site. (e.g., Place more or less tutors based on data.)
4	Conduct local needs assessment	3.24	Identifies student need as part of program application. (e.g., What percentage of students are proficient on local reading assessment?)
5	Develop detailed implementation plan or blueprint	3.18	Creates and implements standard implementation blueprint. (e.g., Tutoring is for grades K-3, focused on foundational literacy skills, and occurs during the school day.)
6	Develop and organize quality monitoring system	3.15	Collects and manages implementation and outcome data in program-specific database (e.g., Database reports include minutes tutored and weekly literacy scores.)
7	Develop instruments to monitor and evaluate core components	3.15	Conducts regular fidelity observations with corresponding tools. (e.g., Tutor observations show mean fidelity of 90% across seven observations.)
8	Obtain and use student and family feedback	3.27	Facilitates sharing tutoring content with families. (e.g., Reading passages are sent home with instructions for further practice.)
9	Monitor the progress of the implementation effort	3.27	Tracks and organizes program-specific implementation data. (e.g., Reporting systems show number of students served per grade.)
10	Stage implementation scale up	2.83	Increases program size across and within years. (e.g., Additional tutors can be added, based on local need.)

[a] Feasibility ratings from Lyon et al. (2019); possible range of 1-5, with higher values denoting greater feasibility.

The critical point to make is that existing staff and structures are unlikely to successfully implement an effective tutoring program both within and across locations. A dedicated organizational structure helps make implementation feasible, consistent, and successful. It also operates in a way that allows broader policy initiatives, and the resources they may activate (as reflected in the outer ring of Figure 1), to have a coordinated distribution channel. The dedicated organizational structure collects resources, can focus them on how to augment implementation, and then leads distribution in partnership with stakeholders at local points of delivery.

Evidence-based Instructional Methods

Evidence-based instructional methods are distinct from other instructional methods that lack strong empirical support and are thus unlikely to be as impactful. Instructional methods that have been subjected to many rigorous, controlled empirical studies—which can in turn be collectively analyzed in what is known as "meta-analysis"—are the most likely to make a profound impact on learning. Figure 2 shows a simple relationship between the empirical rigor used to test a method and its corresponding likelihood of success. Those that are based solely on expert opinion are least likely to influence learning, whereas those with a positive effect derived from quality meta-analysis are most likely to influence learning. Fortunately, educational research has provided many examples of instructional methods that have been rigorously evaluated and shown to maximize learning (see Table 7). Methods like these should be considered the default starting point for any effort to scale tutoring.

As is common in academia, the definition of what constitutes "evidence-based" is debated (Wadhwa, Zheng, & Cook, 2023), but that should not deter efforts to scale tutoring initiatives that prioritize known and effective instructional methods. Academic debates discuss important nuances with methodology, like how sampling and randomization were executed, but scientific reviews of instructional procedures apply rigorous methodological and conceptual criteria to the studies they review. For example, the aforementioned schema-based approach to improving math problem solving skills has been vetted scientifically by multiple reviews: one focused on the extent to which 18 studies addressed indicators of scientific quality (e.g., accounting for attrition; measuring outcomes after immediate post-test; Jitendra et al., 2015), whereas another incorporated these quality indicators into a comprehensive meta-analysis of 21 additional studies involving more than 3,400 students (Peltier & Vannest, 2017), but the practical and policy implications did not vary across the two reviews. The instructional technique of using schemas to organize, understand, and solve math word problems is powerfully effective.

Figure 2: Relationship between the empirical rigor used to evaluate an instructional method and its likelihood of success.

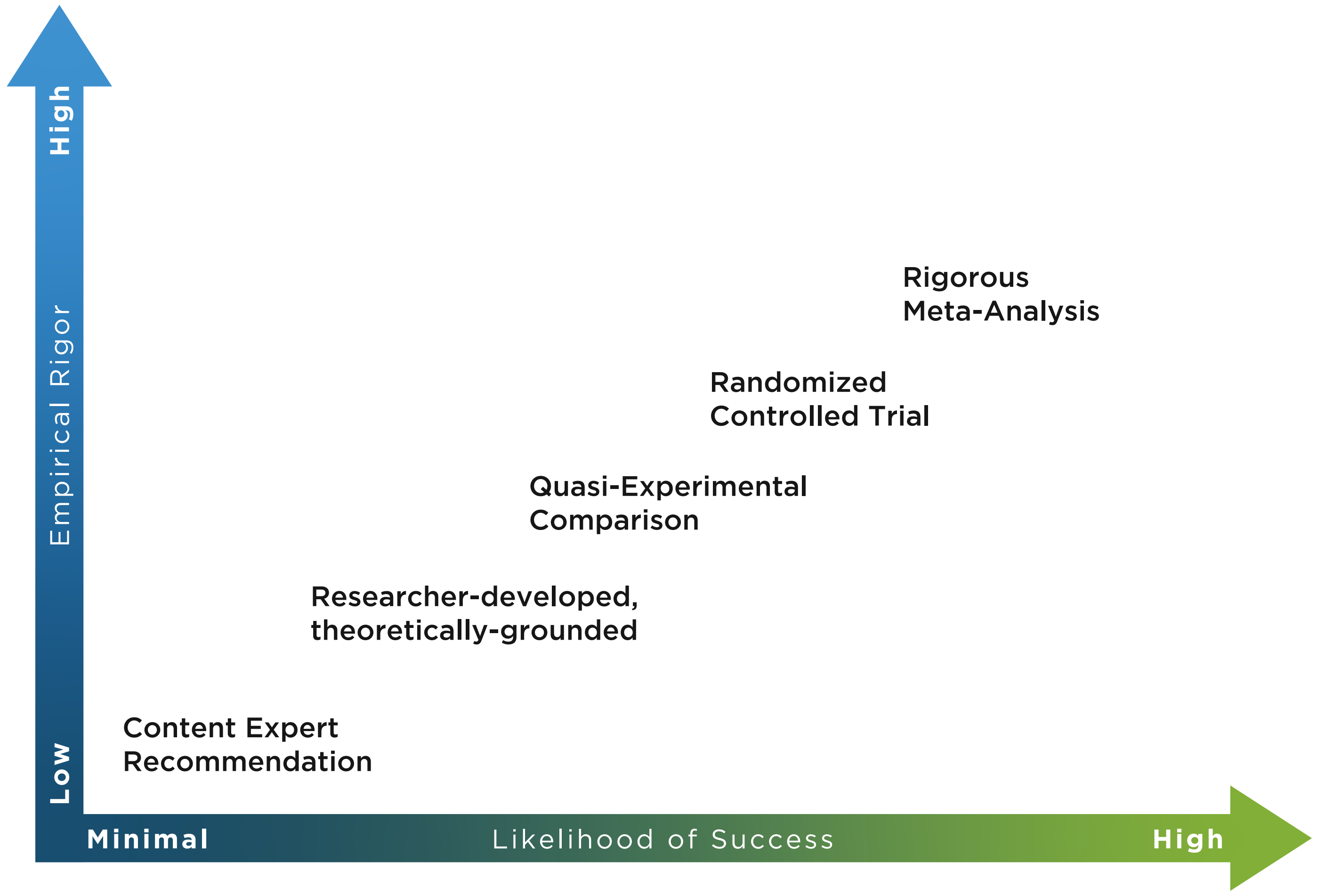


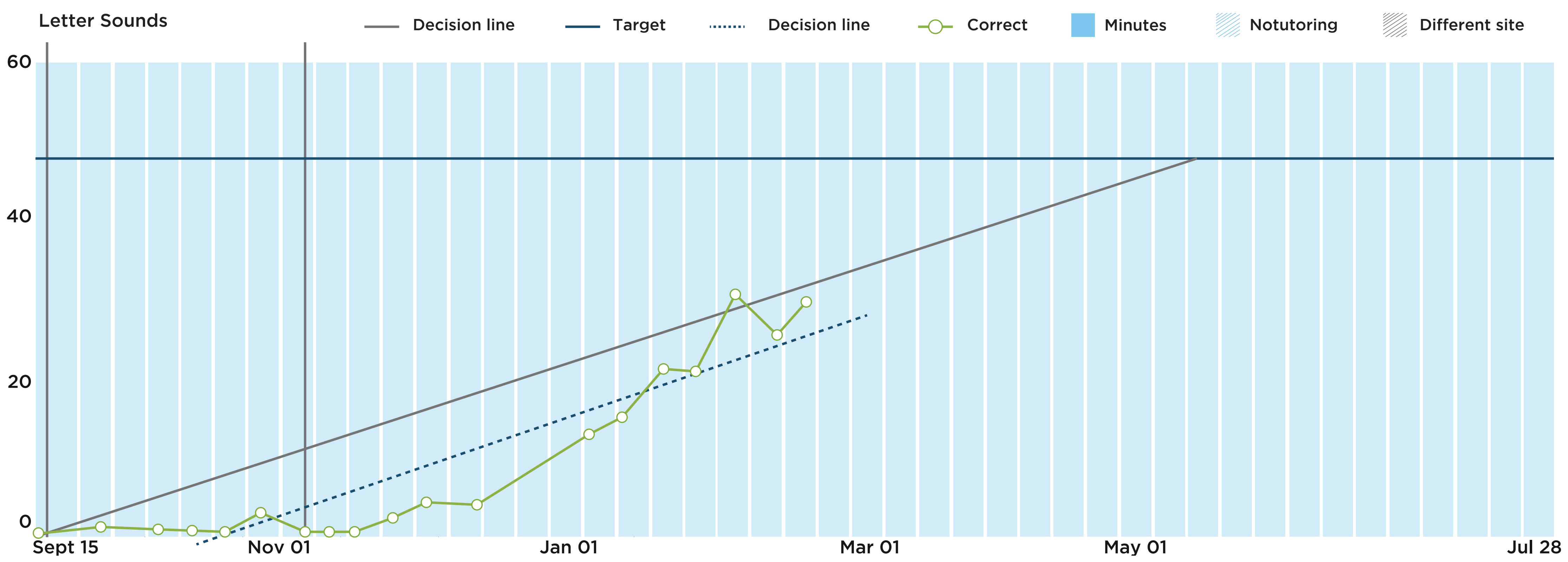
Table 7: Evidence-based instructional methods across various learning outcomes.

Learning Outcome	Instructional Method	Sample Research
Math (Problem-Solving)	Schema-based instruction	Peltier & Vannest (2017)
Math (Conceptual Understanding)	Concrete/Representational-abstract instruction	Carbonneau, Marley, & Selig (2013)
Math (Procedural)	Cover-Copy-Compare procedure	Joseph et al. (2012)
Math (Early Math)	Early numeracy instruction	Nelson & McMaster (2019)
Reading (Sound Awareness)	Phonemic awareness training	Bus & Van IJzendoorn (1999)
Reading (Phonics)	Explicit phonics instruction	Ehri, Nunes, Stahl, & Willows (2001)
Reading (Fluency)	Repeated reading	Therrien (2004)
Reading (Comprehension)	Inference instruction	Elleman (2017)
Reading (Vocabulary)	Dialogic Reading	Marulis & Neuman (2010)
Writing (Composition)	Self-regulated strategy development	Graham, McKeown, Kiuhara, & Harris (2012)
Writing (Handwriting)	Legibility and fluency instruction	Santangelo & Graham (2016)

Data-based Decision-Making

Data for tutoring programs need to fulfil multiple purposes. Data inform instructional decisions and also support program implementation and evaluation. Yet at the core of all data collection and analysis should be student-level performance on the skills and instructional targets specified in the theory of change. If, for example, the focus is on secondary math (e.g., algebraic concepts), then a data collection apparatus focused on those skills is essential. The figure below can be used to demonstrate the breadth of utility data bring to tutoring programs.

Figure 3: Sample figure showing student-level data necessary for effective tutoring.



Though simple, this figure contains a wealth of information regarding student-level data for reading. First, it shows the student had zero initial skills with respect to knowing letter sounds, which is concerning—but not uncommon—with early primary schooling. Second, it shows two black vertical lines that reflect the instructional methods that were used. Direct and explicit phonics instruction (letter-sound correspondence) was tried first (Ehri et al., 2001), but to no effect. Despite the seemingly obvious match with a student having no initial skills, and the strong empirical support for that instructional method, the data were helpful in clarifying that the student had underlying needs, particularly with developing sound awareness skills. Thus, phoneme blending (PB) activities were added as an additional instructional method (Buz & Van IJzendoorn, 1999), and the student made much more rapid progress. A third piece of information provided by this graph is that the student's rate of growth (dashed line) is progressing rapidly toward the goal (blue horizontal line), a key indicator of progress toward age-level expectations. Combined, data like these help tutors make sense of the inherent complexity of learning and instruction.

Importantly, these data provide utility well beyond the simple but critical student-level decisions that tutoring programs need to make. By their mere presence—or absence—they provide valuable insights into program implementation. In this case, a weekly data point is proof-positive that a tutoring session occurred, whereas the absence of data points over a period of time—in this case the month of January—indicate clear need for program oversight. The student may be absent, the tutor may be absent, or some structural challenge (e.g., school closure(s)) may be interfering with tutoring.

Similarly, the student-level data in Figure 3 incorporate all the information necessary to evaluate the program's overall effectiveness. By nature of having data to create such a graph, each participating student has a pre-test score, final score, and a slope reflecting average growth while receiving tutoring. Moreover, each student's final score can be measured against the pre-established target for year-end performance. Collectively, data like these permit robust programmatic evaluation. For tutoring programs, the net result of prioritizing data-based decision-making is threefold: stronger student-level decisions, improved ability to monitor program implementation, and robust program evaluation.

Program for Accurate Delivery

In most social services—from health care to education—there is a persistent and pernicious research-to-practice gap. What is known about a proven and effective practice takes years—sometimes decades—to be incorporated into practice (See, Gorard, & Siddiqui, 2016). As a solution to the current educational crisis, tutoring needs to be an exception to that rule. Facilitating ongoing support for the tutor is among the most important implementation considerations, and was highlighted in research that reviewed over 500 studies and concluded there was "strong empirical support that…implementation affects outcomes" (p. 327, Durlak & Dupre, 2008).

Two broad types of ongoing support might be conceptualized for ensuring tutoring is accurately and effectively implemented at scale. Both are needed to sustain adult behavior change involving the use of effective instructional methods. The first type is general technical assistance. Implementation scientists draw a distinction between the points of delivery (sometimes referred to as delivery systems) and the support systems that can be built around them (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). Support systems exist to help with myriad factors known to affect implementation, from establishing local champions, to informing training processes, to facilitating change management. Each factor is a barrier that individuals working within a delivery system must address, and their ability to access technical support can make the difference between success and failure. The provision of such support speaks to the need for a dedicated organizational structure that can develop and house the capacity to provide general, responsive supports wherever and however tutoring may be provided.

The second type of ongoing support is specific to the tutor and the tutoring process. This type of support ensures the tutoring program remains effectively focused on the instructional methods that help students learn. For this purpose, scaled tutoring initiatives need to recognize and plan for the infrastructure necessary to support ongoing coaching for the tutor. Implementation scientists refer to this type of support as "instructional fidelity" or "instructional integrity". Figure 4 shows a simple checklist that can be used in a coaching session, whether virtual or in person.

Figure 4: Sample checklist for coaches to use in observing tutoring integrity.

INTERVENTION SEQUENCE	Yes	No
1. Tutor has a list of words for blending formatted in a size the student can read		
2. Tutor explains task and gives rationale to student at least briefly every session.		
3. Tutor models task with al least two words every session.		
4. Tutor uses appropriate hand signaling during a model.		
5. Tutor initiates practice.		
6. Tutor uses appropriate hand signaling for each word during practice phase.		
7. Tutor follows error correction procedure immediately for every error. also mark "Yes" if the student did not make any errors during the observed session.		
8. Tutor maintains brisk pace of presentation.		
Total (maximum of 8 possible):		

Items of strength: **Items for improvement:**

In this example, the coach is focused on a set of instructional steps that will help students build proficiency with phonetic relationships at the word level (Ehri et al., 2001). It may be simple, but the simplicity is powerful. The purpose of coaching is to maximize the likelihood of rapid student gains by ensuring tutors closely approximate evidence-based instructional methods. Even the simplest consideration—"tutor initiates practice"—can have outsized relevance, as adults have a tendency to minimize student practice, and opportunities to practice are strongly associated with learning (Van Camp, Wehby, Martin, Wright, & Sutherland, 2020). Scaled tutoring programs that provide ongoing general support to the individuals responsible for point-of-service delivery, including for the tutors themselves and the instructional methods they use, are better positioned to increase student learning in a meaningful way.

Conclusion

The hopes for tutoring are high for good reason. Students need access to high-quality tutoring now more than ever, and there has never been a more robust knowledge base for tutoring and its critical factors. From instructional practices to implementation drivers to data-based decision-making, the blueprint for successfully scaling tutoring is known.

Yet efforts to scale tutoring will have limited success if they ignore the contextual and practical considerations that make tutoring effective. The five key factors presented here capture and organize much of the complexity. When and where they are adequately addressed, tutoring will be successful, changing the learning outcomes of children across the Americas and realizing the great potential of tutoring in the process.

References

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. Administration and policy in mental health and mental health services research, 38, 4-23.

Bus, A. G., & van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. Journal of Educational Psychology, 91, 403–414.

Butler, F. M., Miller, S. P., Crehan, K., Babbitt, B., & Pierce, T. (2003). Fraction instruction for students with mathematics disabilities: Comparing two teaching sequences. Learning Disabilities Research & Practice, 18(2), 99-111.

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. Journal of Educational Psychology, 105(2), 380.

Cook, C. R., Lyon, A. R., Locke, J., Waltz, T., & Powell, B. J. (2019). Adapting a compilation of implementation strategies to advance school-based implementation research and practice. Prevention Science, 20, 914-935.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. American journal of community psychology, 41, 327-350.

Eccles, M. P., & Mittman, B. S. (2006). Welcome to Implementation Science. Implementation Science, 1, 1.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. Review of Educational Research, 71, 393–447

Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review. Journal of Educational Psychology, 109(6), 761.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2012). Implementation research: a synthesis of the literature. 2005. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication# 231), 119.

Franks, R. P., & Bory, C. T. (2015). Who supports the successful implementation and sustainability of evidence based practices? Defining and understanding the roles of intermediary and purveyor organizations. New Directions for Child and Adolescent Development, 2015(149), 41-56.

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. Journal of educational psychology, 104(4), 879.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: systematic review and recommendations. The milbank quarterly, 82(4), 581-629.

Jitendra, A. K., Petersen-Brown, S., Lein, A. E., Zaslofsky, A. F., Kunkel, A. K., Jung, P. G., & Egan, A. M. (2015). Teaching mathematical word problem solving: The quality of evidence for strategy instruction priming the problem structure. Journal of Learning Disabilities, 48(1), 51-72.

Joseph, L. M., Konrad, M., Cates, G., Vajcner, T., Eveleigh, E., & Fishley, K. M. (2012). A metanalytic review of the cover-copy-compare and variations of this self-management procedure. Psychology in the Schools, 49(2), 122–136. doi:10.1002/pits.20622

Joyce, B., & Showers, B. (2002). Student achievement through staff development (3rd ed.). Alexandria, VA: ASCD.

Lyon, A. R., Cook, C. R., Locke, J., Davis, C., Powell, B. J., & Waltz, T. J. (2019). Importance and feasibility of an adapted set of implementation strategies in schools. Journal of school psychology, 76, 66-77.

Mol, S., Bus, A., deJong, M., & Smeets, D. (2008). Added value of dialogic parent–child book readings: A meta-analysis. Early Education and Development, 19, 7–26.

National Mathematics Advisory Panel. (2008). Foundations for success: The final report of the National Mathematics Advisory Panel. Washington, DC: U.S. Department of Education

Nelson, G., & McMaster, K. L. (2019). The effects of early numeracy interventions for students in preschool and early elementary: A meta-analysis. Journal of Educational Psychology, 111(6), 1001.

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. Retrievd from https://www.nber.org/system/files/working_papers/w27476/w27476.pdf

Peltier, C., & Vannest, K. J. (2017). A meta-analysis of schema instruction on the problem-solving performance of elementary school students. Review of Educational Research, 87(5), 899-920.

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. Educational Psychology Review, 28, 225-265.

See, B. H., Gorard, S., & Siddiqui, N. (2016). Teachers' use of research evidence in practice: A pilot study of feedback to enhance learning. Educational Research, 58(1), 56-72.

Slavin, R. E., Madden, N. A., Neitzel, A., & Lake, C. (2020). The national tutoring corps: Scaling up proven tutoring for struggling students. Retrieved from https://www.successforall.org/wp-content/uploads/2020/12/Scaling-up-Natl-Tutoring-Corps.pdf.

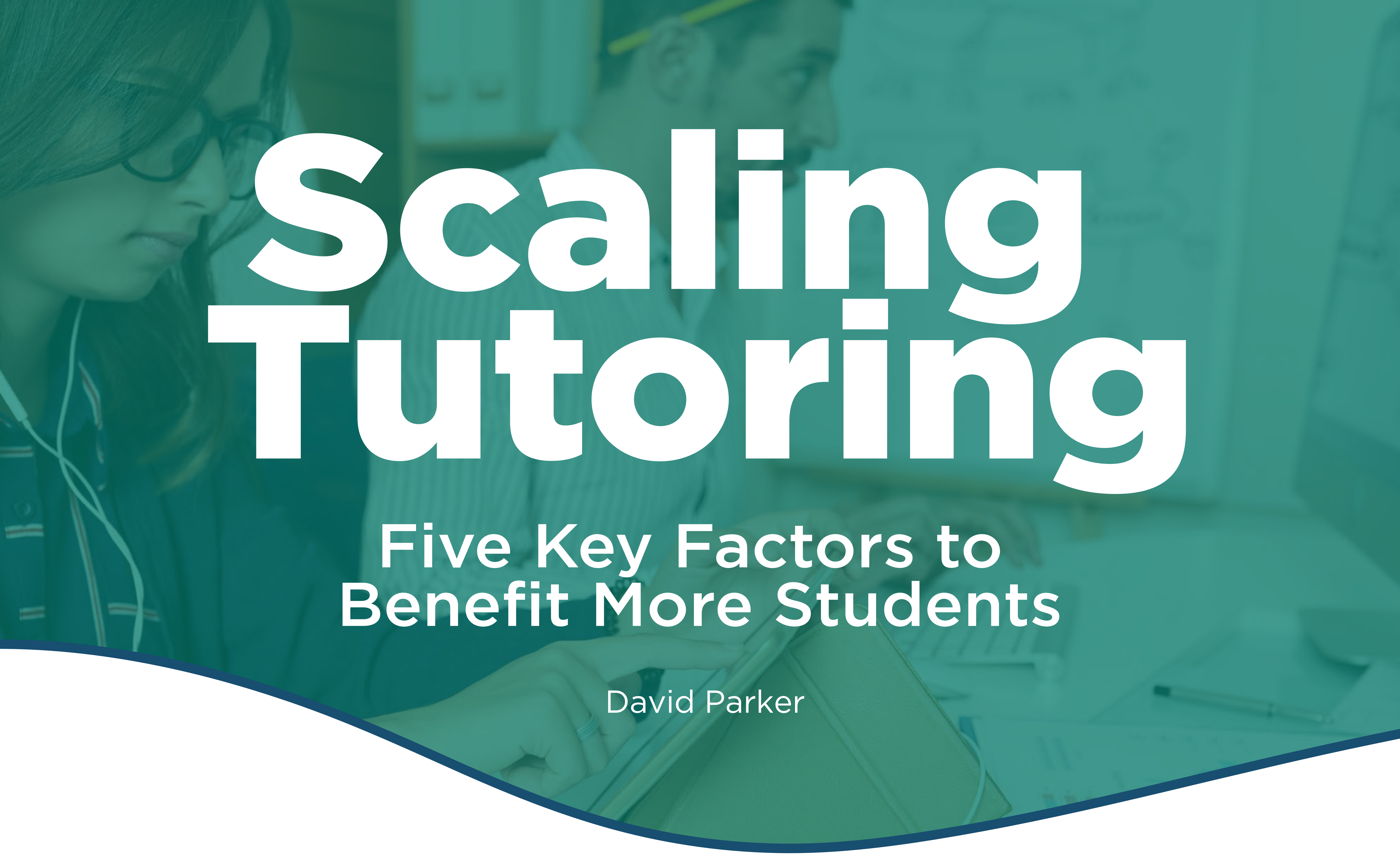
Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. Journal of learning disabilities, 49(1), 77-96.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. Remedial and special education, 25(4), 252-261.

Van Camp, A. M., Wehby, J. H., Martin, B. L. N., Wright, J. R., & Sutherland, K. S. (2020). Increasing opportunities to respond to intensify academic and behavioral interventions: A meta-analysis. School psychology review, 49(1), 31-46.

Varma, S., & Schleisman, K. B. (2014). The cognitive underpinnings of incremental rehearsal. School Psychology Review, 43(2), 222-228.

Wadhwa, M., Zheng, J., & Cook, T. D. (2023). How Consistent Are Meanings of "Evidence-Based"? A Comparative Review of 12 Clearinghouses that Rate the Effectiveness of Educational Programs. Review of Educational Research, 00346543231152262.



Scaling Tutoring

Five Key Factors to Benefit More Students

David Parker

ACCELERATE LEARNING

IDB