TABLE OF CONTENTS

Recovering from COVID-Learning-Loss with a Platform to Support Human Tutoring

The team from The ASSISTments Foundation (TAF) has developed a tutoring support platform, **TutorASSIST**, that acts as an extension of their proven ASSISTments intervention. In a partnership combining the artificial intelligence (AI) abilities of Worcester Polytechnic Institute (WPI), the educator professional learning work of The Friday Institute for Educational Innovation (The FI), and the evaluation abilities of WestEd, TAF proposes a mid-phase Education and Innovation Research (EIR) grant to develop and scale-up a cost-effective implementation of high-dose human tutoring, a proven intervention, in high-needs schools. This implementation will target and help to rectify the learning loss associated with COVID-19 for high-needs students.

Human tutoring is the single most talked about intervention to help alleviate COVID-19 learning loss (Safran, et al.;Schneider, 2020; MIT Poverty Action Lab, 2021; Brown, 2021, Kraft & Falken, 2021; Slavin, 2020; Guryan et al. 2021). Multiple meta-analyses of ~100 well-done randomized controlled experiments on tutoring estimate that tutoring interventions tend to have large and reliable effects (Dietrichson et al., 2017; Fryer, 2017; Nickow et al., 2020b). Never in our professional lives has the field shown such a united front for a single educational intervention: high-dosage human tutoring can help students learn effectively.

[redacted] and [redacted], the co-founders of ASSISTments, have both run volunteer human tutoring programs in the past and have been employed as expert tutors. They know tutoring and they know how to innovate technology. ASSISTments is the most widely used research-based platform for assigning independent practice online from the two most common

and highly rated (EdReports, 2021) core math curricula. This scale-up project will expand upon ASSISTments's success in increasing student learning by scaling-up the TutorASSIST Program that allows schools to manage their own tutors.

During the pandemic, ASSISTments (already funded by an EIR mid-Phase due to having qualified with a What Works Clearinghouse study rating of "positive without reservation" that showed ASSISTments caused reliable learning gains [Roschelle, et al, 2016]) went from supporting 800 teachers to 20,000 teachers and their 500,000 associated students. During this period, 93% of ASSISTments' teacher-users were using either EngageNY or Illustrative Math, showing a tremendous demand for a free digital solution for these curricula. With ASSISTments, students are assigned problems online (important during distance learning) selected by their teachers. Teachers access real-time data on student performance via data dashboards (Appx. J.1.A.a). Students receive immediate feedback as they work (Appx. J.1.B.a and b).

In the spring of 2020, anticipating the urgent crisis and severe learning loss that would result from COVID-19, TAF built a new tool, TutorASSIST, that would enable the data collected by teachers via ASSISTments to be repurposed to support human tutoring. With funding from the Bill and Melinda Gates Foundation, TAF partnered with Cignition, a tutoring company with promising evidence of increased student learning from tutoring (Roschelle, et al., 2020), to create the TutorASSIST tutor dashboard (Appx. J.1.C.a and b) and pilot it in Baltimore City, MD (see Letter of Support in Appx. C showing commitment to this project) and Savannah, GA. This led to a contract with the State of Louisiana which was seeking a tool that would allow both math teachers and local tutors to have access to powerful, curriculum-aligned formative assessment data to address learning loss.

Our proposed intervention, TutorASSIST, aligns with the school math reform movement that focused on ensuring that students are using grade-level content. Too often, low performing students face what has been termed the "opportunity gap" (EdGlossary, 2013). Instead of being presented with rigorous, grade level materials, students are given content that is below grade level, causing many activists to charge that even using the term "achievement gap" is misleading as poor students were never even presented with the same opportunity to challenge themselves (Carter & Welner, 2013). We will now develop a powerful Artificial Intelligence (AI) tool that will **not** replace the human tutor but will instead make that tutor more efficient and aligned tutoring content with core on-grade-level curriculum presented in class. This AI-Agent will provide problems for the tutor to work on with the student. Problems will be aligned to what the student needs to be successful in their main math classroom and will support students in both on-grade-level tasks and prerequisite support material.

The overall goal of the proposed project is to implement, refine, and evaluate a human tutoring program that can support the delivery of tutoring three times a week for 20 weeks with a tutoring ratio of 1 tutor to 2 students in various schools with high-needs student populations. To scale up the implementation, the team will improve the TutorASSIST Program for school administrators and tutors. For school administrators we will provide administrative views in the platform for management and oversight, including video recording of all virtual tutoring sessions, and we will implement training and support. For the tutors we will help them run effective tutoring sessions by providing an AI-Agent that will suggest problems for the tutor to assign, and train tutors on simple methods that emphasize self-directed learning in students. In Year 3 we will study the effects of the intervention by conducting a large-scale, student-level

randomized controlled trial to test the efficacy of the TutorASSIST Program on student learning outcomes (See logic model in Appx. G).

ABSOLUTE AND INVITATIONAL PRIORITIES

Absolute Priority 1 - Moderate Evidence: High-dosage tutoring--generally defined as one-to-one tutoring or small-group tutoring by teachers, paraprofessionals, or trained volunteers occurring at least three times per week--has repeatedly been effective in increasing learning. A meta-analysis of 96 randomized evaluations of tutoring programs found that high-dosage tutoring consistently led to large improvements in learning outcomes for students (in 93 of 96 evaluations), with a pooled effect size of 0.37 standard deviations (Nickow et al., 2020a). The Evidence Form allows us to identify 2 citations (ServeMinnesota, 2018; Cook et al., 2015) showing examples of such studies that overlap with the proposed project. Three additional studies are listed in Appx. J.4 (Guryan et al., 2021, Rutt et al., 2014, Styers et al., 2011).

Absolute Priority 3 - Promoting STEM Education: Since the 2 studies cited in our Evidence spreadsheet are in mathematics, they also qualify for this STEM absolute priority.

Invitational Priority 1 - Underserved students most impacted by COVID-19: Tutors will be assigned to students who missed learning due to COVID-19 and needing additional support.

A. SIGNIFICANCE

A.1 The National Significance of the Proposed Project

Underperformance in math learning among underserved students worsened when schools were forced to close due to the COVID-19 pandemic. The need to improve students' math achievement is well-recognized. According to the National Assessment of Educational Progress (NAEP), only 34% of American public school 8th-grade students score at or above "proficient"

in math (NCES, 2020b). Two-thirds of American 8th-graders from low-income families failed to acquire basic proficiency in reading, writing, and math and the picture is worse for students of color (NCES, 2020a, 2020b). Mounting evidence documents how the pandemic has resulted in substantial academic content learning loss and increased educational inequity (Bacher-Hicks et al., 2020; Engzell et al., 2020; Kraft & Goldstein, 2020; Maldonado & De Witte, 2020; Stein, 2020). According to the NWEA computer testing systems, math achievement dropped during the pandemic much more than reading (Kuhfeld et al., 2020) accentuating the need for math tutoring.

While high-dosage tutoring has been proven effective, it is also very costly. During this pandemic, families that could afford it paid for tutoring. Companies like Varsity Tutors, Tutor.com, Wyzant, and Chegg have reported huge spikes in sales; it is estimated that global private tutoring went from being a $173.4 billion/year business to $279 billion, a staggering growth rate of 61% for one year (Intrado, 2020). Sadly, the massive expenditure by individuals who can afford to hire tutors is only likely to worsen our nation's already large achievement gaps (Schneider, 2021). Schools with high needs students must also have ways to offer tutoring to their students.

We propose leveraging volunteer tutors to mitigate high costs while improving student learning. We understand that licensed teachers would be more competent tutors but the cost is prohibitive. Ritter et al., (2007) did a meta-analysis of 21 research articles and reported that volunteer tutor programs had a very meaningful effect size of 1/3 of a standard deviation which is the equivalent of raising achievement an extra year of education (Lipsey et al., 2012). Nickow et al. (2020b) did a meta analysis on tutoring studies where the tutoro were teachers (18 studies),

Paraprofessionals (46 studies) and volunteers (24 studies). He concluded "while teachers may often make effective tutors, it is far from clear that effectiveness differentials between trained teachers and paraprofessionals outweigh the cost differentials. … Where suitable volunteer pools are available, programs utilizing them as tutors are likely to prove productive" (Nickow et al., 2020b). We are not the only ones who think harnessing the treasure trove of potential volunteers to help one or two students at a time is a great idea. Jon Baron, who is running for governor of Maryland, as an evidence-based candidate, has volunteer tutoring as the first thing he will implement if he becomes governor (Appx. J.3). In summary, the national significance of this project is attested to by 1) the great need on behalf of students due to COVID-19 learning loss, 2) the great science showing tutoring can work, and 3) the cost-effective idea of using volunteers.

A.2 The Potential Contribution of the Proposed Project

Given the huge amount of tutoring that will likely occur in the near future, our nation needs a collection of evidence-based strategies meant to increase collective understanding of the modernization of tutoring programs in general and volunteer tutoring programs in particular. So many industries have become more efficient because of technology. Take banking as an example, first it was the ATM and then it was online banking and now we have Venmo and PayPal. This project if funded will contribute to improving school based tutoring by offering a well-designed software platform and the associated workflows.

The software will deal with administrative tasks such as scheduling tutoring sessions and connecting tutors with students, while also allowing for oversight of the efficacy of tutoring. TAF and WPI, with their work on ASSISTments, have already made substantial contributions in related areas, including the use of technology to solve the systematic lack of immediate feedback

as students work on independent practice, paired with resulting data to help inform teacher instruction (Roschelle, et al., 2016). We are uniquely positioned to do the same for tutors.

A.2.a The contribution of recorded zoom tutoring sessions.

It is safe to say that the COVID-19 pandemic strengthened the use of video conferencing. Pairing a tutor with students online for virtual tutoring is now a viable solution. Video conferencing has improved tutor efficiency, allowing them to work with students before, during, or after school from their home or office. There is a major problem that prevents schools from allowing tutors to work with students online. Schools lack the technology required to effectively monitor these virtual tutoring sessions for the safety of the students; Schools do not allow volunteers to be alone with students in person, so how could they let them be alone on a zoom call? Appx. C documents that school leaders (see Appx. C: schools for letters from Dr [redacted], Dr [redacted] & Ms [redacted], and the letter from iMentor) want automated video recordings to deter bad behavior. At a corporate level, Zoom does not immediately plan to address this issue. Therefore, this project will create an add-on for GoogleMeet and Zoom that provide this functionality using Zoom and GoogleMeet's API's.

B. STRATEGY TO SCALE

B.1 Strategies That Address Barriers To Scale

For the impacts of human tutoring to scale, schools need support in managing and sustaining tutoring activities. How can we support schools in recruiting, training, monitoring, and affording tutors while also supporting tutors in the complicated task of increasing student learning? Exhibit 1 presents three **barriers** to schools implementing tutoring at scale, and our proposed strategies for overcoming them**.** Note the second and third barriers are of our own making due to using

volunteer tutors as our solution to the first barrier.

Barrier	Strategy
1. The high **cost** of hiring licensed teachers to tutor.	Support schools in implementing a **volunteer** tutoring program.
2. Tutoring by volunteers can be inefficient and ineffective and not aligned with core instruction.	Implement the AI-Agent and the self-directed method of tutoring.
3. The **extra time** invested in **oversight** & **administration** of volunteer tutoring.	Provide supportive administrative **tools** that are effective and easy to use.

Exhibit 1: Barriers and Strategies to Scale

Strategy 1 - Support schools in implementing a volunteer tutoring program.

Recruiting volunteers to serve as tutors will lower costs while still yielding positive gains. The supply of such volunteers exists. The US Census in 2017 conducted a stratified sample of households, and 29.6% of respondents reported they volunteered and one fourth of those claimed they tutored as their volunteer activity (US Census, 2017). We intend to capitalize on that interest -- people want to help and are interested in tutoring. When US Representative Alexandria Ocasio-Cortez recently asked for volunteers to tutor students in the Bronx via Zoom during the pandemic, 13,000 volunteers signed up within days (Zimmer, 2020).

Tutoring can provide immediate satisfaction for volunteers due to building a personal relationship with students and as mentioned in section C.1 volunteers can be effective in increasing student learning. ExperienceCorps is a program for adults over 60 to provide tutoring, and has some amazing results (Fried et al. 2003). Fried et al. showed in their well done randomized controlled trial that adults get healthier when they volunteer. This same study also shows improvements in students' achievement and behaviors. Interestingly teacher morale and even retention and more generally school climate became better as well. A lot of school districts recruit volunteer tutors (Somerville, 2011, Boston, 2011). People from all walks of life can be

tutors. AmeriCorps, a US volunteer program, is for young people and as mentioned above ExperienceCorps is for older people. Our TutorASSIST Program will have the ability for anyone who wants to volunteer, no matter how far they are from the school, to be able to do so as a tutor. We allow tutors to work in person or virtually. So an out of state grandparent could volunteer at a school if they want.

Strategy 2 - Implement the AI-Agent and the self-directed method of tutoring.

We will use artificial intelligence to create a mechanism, the AI-Agent, that will select which problems the tutors will work on with the students. The idea for the AI-Agent was motivated by the inefficiencies we noted in the Saga Model of tutoring where the lesson plan for a tutoring session had 44 problems to choose from and multiple pages of information about the topic (Saga, 2021). Just picking the problem to use would take an hour due to all the choices. If the tutor is a volunteer we could never be sure they could do this.

Exhibit 2 shows how the student data, including performance and timing data on every problem a student has ever done in ASSISTments will feed into the AI-Agent. Each of those problems is tagged with a Common Core standard and each problem from one of the OER 6th, 7th, or 8th grade curriculum will have a *similar but not the same problem* (SNS) (built as part of this project see example in Appx J.2.B.d) that will be part of all the problems available. The AI-Agent will use the data to select problems. Two simple examples are (1) if a student does poorly on a problem for the teacher the AI-Agent will select a SNS problem for the student to solve with the tutor or (2) if a student does poorly on a problem then the AI-Agent will select a problem from the previous grade level that covers a prerequisite skill (using the Prerequisite hierarchy that is already in the system). This will allow the tutor to give support for prerequisite

knowledge while the teacher stays on grade level continuing the strategy of supporting students as they work at an age appropriate level in the main math classroom.

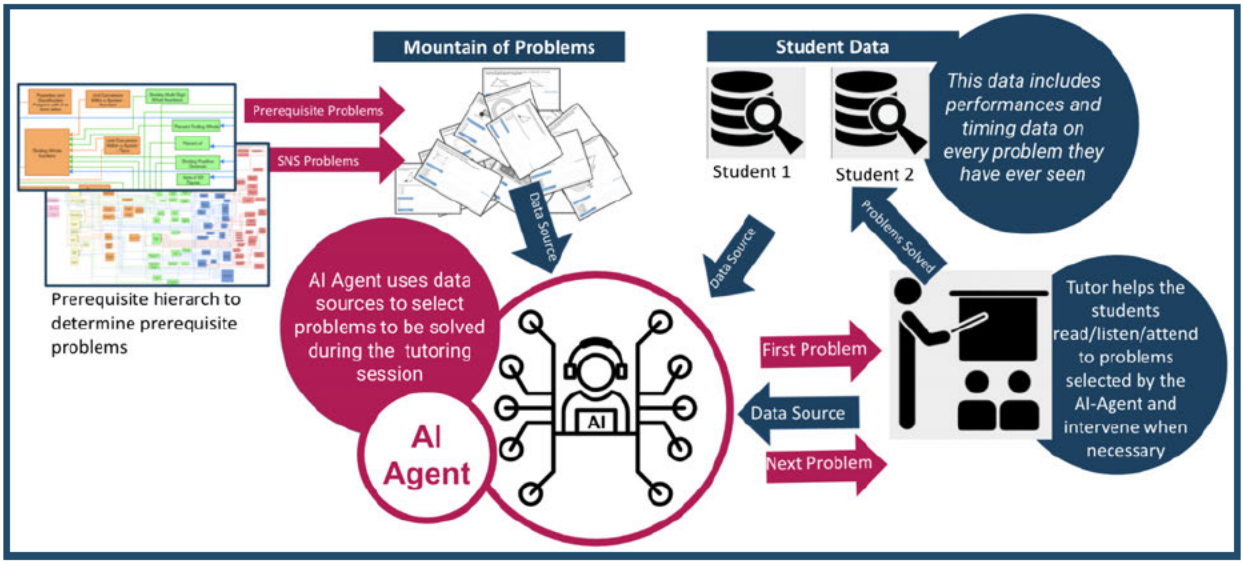

Exhibit 2: AI-Agent suggests problems from it's inputs, see Appx J.2.B.b and c

The next thing we will do in order to support volunteer tutors who may be novices in math instruction is to train them in a self-directed method of instruction. During their *Foundations of TutorASSIST* course tutors will learn how to respond to a student who is working on the problem that has been delivered by the AI-Agent. The tutor will learn to not just jump in and help with the math, they instead will be trained to initiate a self-directed strategies such as (a) having the student look at an online hint provided through ASSISTments or (b) asking the student what they would do first if they were working alone then (c) suggesting the student draw a diagram or write down some details from the problem or (d) having the student read the hint out loud. A list of these strategies can be found in Appx. J.5.C. This strategy of using AI and a simple teaching method will allow us to overcome the barrier of having tutors who may not have any formal training in math teaching be effective and inspired tutors.

Strategy 3 - Tools for schools to implement tutoring - the TutorASSIST Program

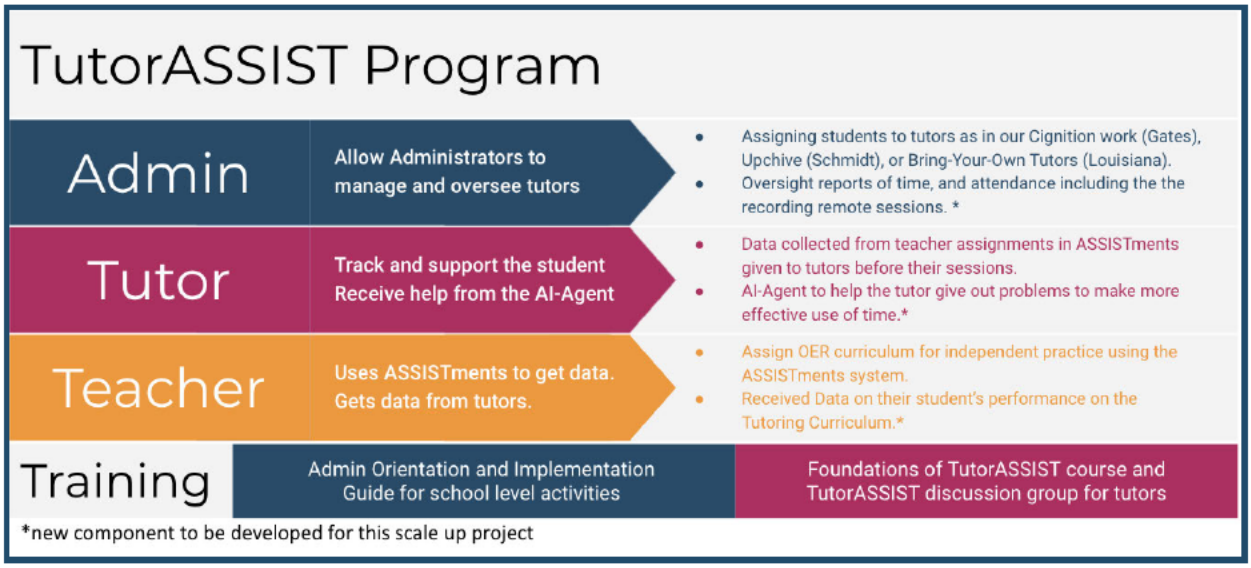

Exhibit 3: TheTutorASSIST Program is composed of components that help 1) adminators, 2) tutors and 3) teachers. It also comes with training (shown at the bottom)

It is one thing to say tutoring is the solution to learning loss, it is quite another to run a tutoring program. Managing all the details of recruiting, training and managing multiple tutors, even if they are paid and provided to your school by an agency is a difficult job. We will provide a system to support the school with this workflow. Exhibit 3 shows the three consumers of the TutorASSIST Program. The admin portion of the program provides support for the management and oversight of the tutors. In the diagram the portions of the program that are currently working are shown as are those that will be developed as part of this scale up project (these are marked with an *). It is key that the program will provide the admin with an orientation and extensive implementation guides to support their adoption of the program. A detailed description of the design of the Admin portion of the TutorASSIST Program is given in section C.3.a.

B.2 Dissemination for Further Development and Replication

TAF will be able to widely disseminate the TutorASSIST Program to the quickly growing user base of over 5,000 schools that are already using ASSISTments. While ASSISTments is

forever free, we plan to sell a package that would include access to the software platform, the training course, and the initial training and implementation guide for the administrators described in section C.3.e. Building on our Louisiana model, we would price TutorASSIST as an annual school subscription, enabling the flexibility to provide tutoring to all students in the building if needed. Our market research has shown that if we do sell it as part of a package we need to price it at around the $10,000 range for a school. This allows schools to purchase the program without a request for proposal (RFP). Our professional development is, and has always been, a separate add-on. We have seen success with this model in Louisiana.

Given the rapid growth of ASSISTments, we have a ready market for the adoption of the TutorASSIST Product. Growth of ASSISTments fuels growth of TutorASSIST. ASSISTments has scaled to reach over 20,000 teachers in the past year, and TAF has developed a partnership with the state of Louisiana as well as large urban districts like Baltimore and Savannah. We will continue this growth through advertising, strategic communications, partnerships that grow awareness, and by adding curricular content that expands our market. We have formed formal partnerships with curriculum publishers Kendall Hunt, Jump Math, and Open Up Resources that open the doors to district adoption. In Louisiana alone, we have a ready market for statewide growth of TutorASSIST as it is, with the funding and demand to enable its reach to grow.

C. QUALITY OF THE PROJECT DESIGN

C.1 Conceptual Framework Underlying the Proposed Research

The quality of our project design revolves around the objects in our **conceptual framework**. It begins with the recommendations given by the Poverty Action Lab that are directly aligned with the research on what makes for high-quality high-dosage tutoring(J-PAL, 2021). Exhibit 4

shows their 5-framework recommendations and how our program will address these. Our logic model that accompanies this framework is in Appx. G.

The J-PAL Conceptual framework and how the TutorASSIST Program addresses the framework	
1	Tutoring sessions should occur at schools, during the school day or immediately before or after school.
Our tutoring sessions will only occur during the school day or immediately before or after the school day. In the Admin panels of TutorASSIST, the administration will only be able to select tutoring times from these options.	
2	Tutoring should occur a minimum of three sessions per week.
Tutors will commit to tutoring three times a week.	
3	Students should meet with a consistent paid and trained tutor supported by ongoing oversight and coaching.
Students will be paired for a whole semester with one tutor, and the tutors will participate in the Foundations of TutorASSIST course as well as sign up for the TutorASSIST tutor discussion group. The course will be for 10 hours with the second half being after the tutoring sessions begin so that it will involve reflection of the actual tutoring. There will also be a summative evaluation halfway through the year conducted by the school administrator or cooperating teacher with assistance from the tool. Constant oversight by the school administrator will be made easier through the development of online tools.	
4	Tutoring sessions should be informed by data.
Tutors will begin their session with both data from class as well as data from previous sessions and AI-enabled data driven activities provided by the AI-Agent.	
5	Tutoring materials should be aligned with research and state standards.
The tutors will be trained in the Self-Directed tutoring method powered by the AI-Agent and directly tied to the SNS problems that are by design aligned to the high quality research driven curriculum as well as problems that come from the common core state standards prerequisite hierarchy.	

Exhibit 4: The Conceptual Framework of the TutorASSIST Project

This framework of the MIT Poverty Action Lab is high quality and universally respected. The group's meta-analysis reviewing all of the tutoring studies (Nickow, Oreopoulos, & Quan, 2020a & 2020b) are already heavily cited even though they were published just last year. The Poverty Action Lab garnered a great deal of respect; the founders won the Nobel Prize in 2019 (Finkelstein, et al. 2019). The current head of the US Department of Education's research arm, Mark Schnieder, recently co-authored an OpEd with Dr. Dadisman, the head of the education division at the Poverty Action Labs, highlighting his agreement with their framework. (Dadisman & Schneider, 2000).

C.2 The Extent to Which Goals, Objectives, and Outcomes Are Specified and Measurable

The four objectives of this project will guide the scaling up and evaluation of this improved version of TutorASSIST. The four activities in *Objective 1* support the iterative improvement of our overall strategy of supporting the school administrators and the tutors participating in the tutoring program. By achieving *Objective 2*, we will have implemented a sustainable program that can be scaled and replicated. *Objective 3* supports our desire to assess the efficacy of the project while *Objective 4* supports our dissemination and sustainability. All objectives, activities, outcomes, and measures are shown in Exhibit 5.

Activities	Outcomes	Measures
Objective 1. Iteratively improve and further develop the methods and materials to support school based volunteer tutoring		
1.1 Build a workflow infrastructure to help school administrators easily manage the tasks of leading a tutoring program with volunteer tutors.	A software infrastructure that allows school personnel to efficiently onboard tutors and match them with students is created and used by schools to manage the tutoring workflow.	By the end of year 2 administrators in the partner schools will have an online record of tutors and students, including parent permission slips and tutor student pairings and schedules.
1.2 Iteratively improve and develop data needed for the tracking and oversight that will go on the Admin dashboard.	Data on the Admin dashboard such as Zoom recordings when available, attendance, tutor actions and student actions as well as post session messages from tutors to teachers.	Through surveys and interviews we will collect information from the school administrators that they are satisfied with their ability to have oversight of the tutoring sessions.
1.3 Create the AI-Agent and the SNS problems that iteratively improve the inputs and outputs of the AI-Agent to support the self-directed tutoring method.	Creation of SNS problems (for all 6th,7th, and 8th grade problems in Engage NY and IM) and the AI-Agent. Support for the self-directed tutoring method.	Count of the SNS problems and survey and interview data from tutors and students on the effectiveness of the AI-Agent. Satisfaction on post session survey from tutors and students.
1.4 Iteratively improve and further develop volunteer tutor training materials and support for the tutors and administrative structure. Including creating an implementation guide for the program for school administrators.	A prepackaged volunteer tutor program that schools can implement in their schools with minimal administrative burden. This will include the Foundations of TutorASSIST Course and the TutorASSIST Program Implementation Guide.	Trained volunteer tutors working with students and receiving positive reviews by students, teachers and school administrators. Administrator satisfaction reported in feedback surveys, cooperating teacher satisfaction reported in feedback surveys.
1.5 Formatively evaluate the design and execution of the tutor training and tutor program administration and materials for process, outcomes,	Interim memos and annual evaluation reports for program improvement - mixed methods approach based on CIPP Evaluation	Interim memos and annual formative evaluation reports in SY1 and SY2 that are useful to the program team for design and

and scalability.	Model and research framework for scaling innovations	implementation improvements.
Objective 2. Implement scalable model of volunteer tutoring in middle schools		
2.1 Support 70 participating schools in recruiting and training the number of tutors needed to support the students designated for tutoring.	For each of SY3 and SY4 200 tutors with stipends and 600 without to tutor 2400 students at 70 schools as part of the BYOT system.	Count of tutors for the designated number of schools who completed the Foundations of TutorASSIST course.
2.2 Support the 70 schools and the 800 recruited tutors as they implement the TutorASSIST Program in SY3 and SY4.	4800 students in SY3 and SY4 will be tutored using the self-directed method of tutoring.	Since all parts of the TutorASSIST program involve signing into ASSISTments and TutorASSIST, we will measure tutoring by action logs.
Objective 3. Rigorously evaluate the impact of the TutorASSIST Program		
3.1 Randomly assign 6th and 7th grade eligible students to treatment condition (immediate tutoring) or control condition (delayed tutoring) to receive tutoring in SY3 or SY4.	Samples of treatment and control students with baseline equivalence in prior performance with appropriate parent consent and student assent.	List of students (and their associated schools and teachers) assigned to each condition. Report on the baseline equivalence.
3.2 Prepare data collection instruments and administration procedures and collect data on tutor practices and student learning outcomes.	Data on main outcome measures, tutor/teacher/administrator practice, data collected and student migration tracked; participation rate and attrition rate calculated according to WWC standards.	State standardized test, tutor logs, protocols for analyzing tutoring video recordings, tutor and school administrator interview protocols, teacher survey.
3.3 Measure and analyze implementation fidelity and explore the relationship between implementation and its effects on learning.	Fidelity and compliance levels analyzed and their relationship with intervention effect analyzed; factors that hinder or facilitate implementation documented.	ASSISTments and TutorASSIST log data, tutor logs, PD records. Implementation fidelity report from WestEd according to the implementation compliance thresholds.
3.4 Assess the impact of TutorASSIST on student learning outcome.	Data on main learning outcome measures collected and analyzed, following WWC-approved approach.	Study reports of findings on group differences in student math learning outcomes, with subgroup effects.
3.5 Collect cost data and analyze cost-effectiveness.	Findings on cost and cost-effectiveness ratio.	Memo of findings regarding the cost and cost-effectiveness of TutorASSIST.
3.6 Refine materials and procedures.	Improved materials and procedures for all aspects of the project.	Records from team meetings; annual report on revisions made to materials and procedures.
Objective 4. Disseminate findings and Support Sustainability		
4.1 Broadly disseminate learnings of how best to support schools' tutoring programs and final results of impact study.	Increased awareness among school personnel, researchers and policy makers of how to support tutoring programs, results of the study.	Number of presentations at conferences, blog posts, webinars, and publications.
4.2 Develop a business plan for the "bring your own tutor (BYOT)" TutorASSIST Program to sell to schools.	Develop a business plan and a pricing model for the BYOT TutorASSIST Program.	Sold the TutorASSIST Program to at least five schools in 3 different states.

Exhibit 5: Objectives, Activities, Outcomes, and Measures

C.3 The Design of the Proposed Project Addresses the Needs of the Target Population.

The TutorASSIST Program will be a complete easy-to-use support for a school that is using one of the highly rated OER, Eureka Math/EngageNY or Illustrative Math/OpenUp curricula to carry out a volunteer tutoring program. Once a school administrator decides to implement math tutoring in their school and to adopt TutorASSIST to support that program, they have everything they need for success. By school years 3 and 4 we will be prepared for a scaled rollout of the TutorASSIST Program. We will now explain how it is designed to address the needs of struggling students through the workflows of the three different TutorASSIST users: a) **administrators** who have to oversee and schedule the tutoring, b) the **teachers** that have students being tutored and c) the **tutors** who are carrying out the self-directed tutoring method with the support of the AI-Agent.

C.3.a How ADMINS are supported by the design of TutorASSISST to address student's needs

The program starts with an implementation guide and initial training for the school administrator. They will be given advice on and materials for recruiting volunteers and a tool for onboarding those volunteers (Appx J.2.A.a). A school administrator will work with classroom teachers to select the students to be tutored. There is the option to use ASSISTments data to identify these students (Appx. J.2.A.b), but the school can use any data of their choice. Administrators will record which students are selected to be tutored (Appx. J.2.A.c). Finally the tool allows for the school administrator to match students with their tutors (Appx. J.2.A.d).

Once the tutoring begins, the school administrator can monitor attendance and activity in the system on a dashboard (Appx. J.2.A.e). The tutoring sessions will be conducted just before, during or just after school (consistent with our conceptual framework Section C.1). We also

expect a combination of virtual tutoring and in-person tutoring. The combinations of these will be as varied as the schools adopting the TutorASSIST Program.

Both the virtual and the in-person methods have been piloted by TAF. The Cignition project is conducted with students at school but with a virtual tutor with a ratio of 1:2. The Louisiana program will begin August 2021 with tutors in person in the school building.

C.3.b How TEACHERS are involved with the design of TutorASSISST to address student's needs

The key way that teachers participate in the TutorASSIST Program is through ensuring their students are doing their daily assignments—classwork and homework—in ASSISTments, allowing the AI-Agent to have current learning data on the students as an input. While this grant is not about ASSISTments, it is hard for the reader to understand how TutorASSIST can even function without a cursory understanding of ASSISTments. This is the workflow for the 20,000 teachers using ASSISTments following four steps (Appx. J.1.A.a): teachers find their textbook sections (Appx. J.1.A.b) and then assign online (Appx. J.1.A.c); students complete assignments and get immediate feedback, using hints as necessary (Appx. J.1.B.a and b); teacher does not need to grade and instead uses that time to analyze student learning results from reports (Appx. J.1.A.d); and data is used to inform instruction and is discussed with students (Appx. J.1.A.e).

This grant will add one additional dashboard for teachers. At the end of the tutoring session, the tutor will write a brief, two or three sentence summary of the session that the teacher can use to reach out to the student with an informed comment about the tutoring (Appx. J.2.C.a).

C.3.c How TUTORS are supported by the design of TutorASSISST to address student's needs

Recruited tutors will be entered into TutorASSIST by the school administrator and begin their training. The training is detailed in section C.3.e. Each tutor will be assigned to two students who they will meet with 3 times a week for 20 weeks.

The tutors can prepare for their session by reviewing the student's prior week's work in math class using the tutor dashboard (Appx. J.1.C.a and b). They will begin their session by greeting the students, and after establishing some rapport, they will begin by clicking the "Start Tutoring Session" button (Appx. J.2.B.a). This will inform the AI-Agents to send the first selected problem to the students. The tutor will then pay attention to the students as they solve the problems then practice the self-directed method of tutoring (Appx. J.5.C). A dashboard will collect the tutor's observations (Appx. J.2.B.f).

C.3.d How the AI-Agent addresses students needs

The AI-Agent decides which problem the students should work on with their tutor. It picks these problems from the "mountain of problems" available to assign (see Exhibit 2). This includes the *similar but not the same* (SNS) problems, hint messages and problems from earlier grades. Inputs from problem solving and the tutors end of session survey will power the AI-Agent. The team at WPI is well positioned to create this AI-Agent.

C.3.e Training of the Volunteer Tutors and Supporting School Administrators

To train the volunteers to be effective tutors for high needs students, The FI--in collaboration with TAF--will build an online course called the "Foundations of TutorASSIST." designed to build the capacity of the volunteers to understand and implement the TutorASSIST Program. The module will be designed around research-based approaches to professional learning (Darling-Hammond et all., 2017; Powell and Bodur, 2019), anchored in four core design principles: (a) learning from multiple voices (b) peer-supported learning, (c) job-connected

learning, and (d) self-directed learning (Kleiman, Wolf, & Frye, 2015). The module will predominantly be delivered asynchronously so that it may be more successfully scaled. The volunteers will arrive at the program with a disparate range of experience, and the module will be designed to account for these varying use cases. The FI will also design a competency-based summative assessment system that will be used to verify a tutor's knowledge and skills for being an effective TutorASSIST tutor. A TutorASSIST tutor discussion group will also be provided.

To onboard and support the school administrator, The FI will design and develop an onboarding module and an implementation guide for school-level TutorASSIST program administrators. A list of the materials to be created can be found in Appx. J.5.A.

The FI will formatively evaluate the tutor training and tutor program administration using feedback surveys and focus groups with tutors, students, school administrators, cooperating teachers, student learning measures, and program artifacts (see Appx J.5.B).

D. Adequacy of Resources and Quality of Management Plan

D.1 The Applicant's Capacity to Bring the Proposed Project to Scale on a National Level

The project team is uniquely qualified to carry out the proposed work. The team brings extensive expertise and experiences in technology development, tutoring support, math education, instruction design, user experience and learning engineering, and rigorous and mixed-method research and evaluation. Exhibit 6 details key personnel's role and their experience for a complete listing of personnel see Appx. J.6.

Experience and Primary Responsibilities
The ASSISTments Foundation Personnel
[redacted], *Executive Director,* has a strong background teaching and coaching middle school math, including seven years of teaching in middle schools in various settings. She served as a math teacher in Gabon as a Peace Corps volunteer and understands public service volunteerism. After college, she worked as a full time private tutor and has a keen understanding of the tutoring world. Ms. [redacted] earned a Master's Degree in Math Education working with Dr. [redacted] author of one of the best-selling books on instruction. Starting in 2005, Ms. [redacted] collaborated with [redacted] on ASSISTments. She has served as senior personnel on dozens of NSF, IES, EIR,

Gates, Hewlett, Schimdt and other grants. In 2019, she and her husband convinced two major philanthropic organizations to give $3 million to start the ASSISTments Foundation (TAF) to take ASSISTments outside of WPI. She now serves as the head of that organization as the executive director. TAF has 9 different grants worth over $5 million. TAF has the sole license to all the intellectual property of the ASSISTments platform. She continues to collaborate with [redacted]. Where he does the AI work, she oversees all operations, programs, and teacher trainers. For this project she will: ● Oversee the project and coordinate with all partners. ● Manage TAF's role in the project. ● Oversee the programmers. ● Write the annual report and disseminate the findings.
Others leaders: [redacted] - VP of Product and Programs and [redacted] - VP of Product*.
Worcester Polytechnic Institute Personnel
[redacted] *Computer Science Professor at WPI and Director of Learning Sciences & Technologies PhD program*. In high school, he started a peer tutoring program. For three years at Amherst College, he ran a tutoring program similar to the one described in this grant; every week, he drove his troup of tutors to Holyoke, MA. He had worked to convince 70 undergraduate students to commit 1.5 hours a week to tutoring a 5th grader. In retrospect, he wishes he had something like the AI Agent, as he often felt the most frequent question he received from his tutors was, "What should we work on next?" Later, he taught math in Teach for America in Baltimore, where he met [redacted]. From there, he studied at CMU and began building intelligent tutoring systems. For his dissertation, he studied human tutors and got computers to mimic some of the strategies human tutors used. At WPI, he quickly became one of its most prominent professors (in terms of citations, best paper awards, and the largest amount of funding at WPI). He teaches AI and has over 100 papers in venues like Artificial Intelligence in Education, Learning at Scale and Educational Data Mining. ● Will oversee the creation of the AI-Agent
Friday Institute for Educational Innovation Personnel
[redacted], *Development & Formative Evaluation Director,* is Senior Director for the Program Evaluation and Education Research (PEER Group) and Assistant Teaching Professor in Learning Analytics at The FI. Since 2011, Dr. [redacted] has led comprehensive research, evaluation, and capacity-building projects centered on the use of technology to make learning more equitable, engaging, and effective for students and educators. ● Direct the development of tutor training and school-level program administration structure and materials. ● Direct the formative evaluation of the tutor training and school-level program administration structure.
Others leaders: [redacted] - Lead Formative Program Evaluator [redacted] - Lead Instructional Designer*
Evaluation Key Personnel WestEd
[redacted], *Evaluation Director*, is a Senior Research Associate with the STEM program at WestEd. She leads several large-scale implementation and impact studies of technology-based education interventions, using a mixed-method approach. She also conducted research in design and development of intelligent tutoring systems and carried out analysis of system data. ● Direct the evaluation and the impact study. ● Lead implementation fidelity measures and analysis.
Others leaders: [redacted] - Lead formative evaluation, [redacted] - Lead Methodologist*
*For more information see Appendix J.5

Exhibit 6: Roles, Experience and Primary Responsibilities of Key Personnel

D.2 The Adequacy of the Management Plan

This management plan with its timeline and tracking of milestones will achieve the objectives on time and within budget. The timeline in Exhibit 7 shows when each portion of the

project will be conducted and completed. Exhibit 8 details the years in which each activity will be accomplished as well as who is responsible.

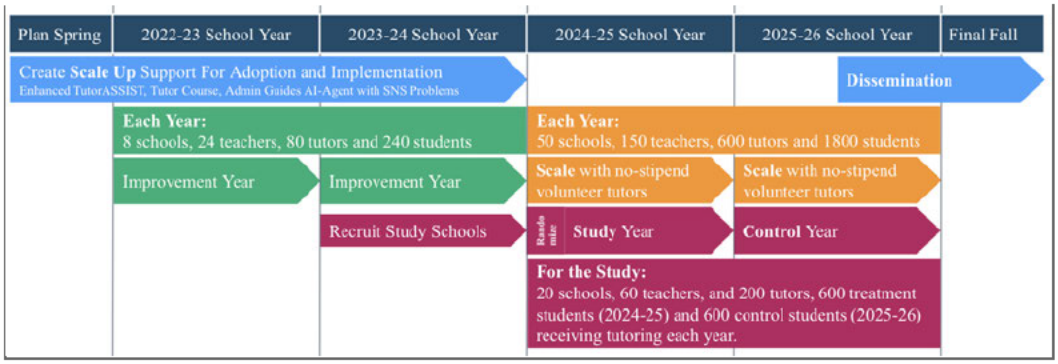

Exhibit 7. Timeline

Milestones	Sp 0	SY1	SY2	SY3	SY4	Fall 5	Who is responsible?	Others
Objective 1. Iteratively improve and further develop the methods and materials to support school based volunteer tutoring								
Admin Oversight Dashboard	x	x	x				**TAF,** WE, FI	PA
Exploring Data Needs	x	x	x				**TAF,** WPI, FI, WE	PA, PT, CT
AI-Agent and SNS problems Development	x	x	x				**WPI,** TAF, FI, WE	PT, SNS
Create Training of Volunteer Tutors and Supports for Administrators	x	x	x				**FI**, TAF, WE	PA,PT
Evaluate tutor and admin supports		x	x				**FI**, TAF, WE	PA,PT, CT
Objective 2. Implement scalable model of volunteer tutoring in middle schools								
School recruitment for scale			x	x	x	x	**TAF,** WE, FI	SA, CT
Tutor recruitment and training at scale				x	x	x	**FI**, TAF, WE	ST, SA
Monitored tutoring of students at scale				x	x	x	**TAF,** WPI, FI, WE	ST, SA, CT
Objective 3. Rigorously evaluate the impact of the TutorASSIST Program								
Random assignment				x			**WE**	SA
Prepare instruments and collect data			x	x	x		**WE**	-
Measure implementation fidelity				x	x		**WE**	-
Assess intervention effects					x	x	**WE**	ST, SA, CT
Collect data and assess cost				x	x	x	**WE**, TAF	SA
Refine materials and procedures			x				**TAF**, FI, WE	ST, SA, CT
Objective 4. Disseminate findings and Support Sustainability								
Present Findings				x	x	x	**WE**, WPI, FI TAF,	ST, SA, CT
Sell TutorASSIST Program to schools						x	**TAF**	-
Key Institutions and Teams: lead in bold, TAF - The ASSISTments Foundation; WPI – Worcester Polytechnic Institute; WE - WestEd; FI - Friday Institute; SNS - Similar Not the Same Builders; PA - Pilot School Administrator; SA - Study/Scale Administrators PT - Pilot Tutors; ST - Study/Scale Tutors; CT - Cooperating Teachers (for a description of teacher partners see Appx J.6.E)								

Exhibit 8. Milestones, Timeline and Responsibilities

D.3 The extent to which the costs are reasonable and potential significance.

We believe that it is highly feasible to replicate and expand the implementation of TutorASSIST and its school support model in a variety of settings and populations and at a reasonable cost for several reasons. First, TutorASSIST is developed to work with any tutors the school chooses. For instance, should schools decide to use ARP funding to contract with higher quality (and higher priced) tutors, they should have that option. The BYOT service should work the same as if they were volunteers. When the funding cliff occurs from the ARP, the volunteer model will still be there for the schools to use.

Interestingly, the head of IES has remarked that gaps were widening between lower performing students and higher performing students even before COVID-19, suggesting the need to close gaps existed long before COVID-19 worsened these gaps (Schneider, 2021). Even though the initial investment in high dose tutoring for all Title 1 students could cost $120 billion, this relatively small investment of $8 million could unleash the ability of millions of volunteer tutors to help millions of students. At this cost this is a very reasonable way to increase student learning.

E. PROJECT EVALUATION

An independent team of highly experienced researchers from WestEd will provide formative and summative information throughout the five-year project. The evaluation will consist of two components. The **formative component of evaluation**, aimed at augmenting TAF's development and continuous improvement of the TutorASSIST Program, will be led by Dr. [redacted] who is an expert in mathematics education. She directed the research of the National Center on Cognition and Mathematics Instruction and has extensive experience with developing

education technology, formative assessment, and studying teacher experience. The formative evaluation will be guided by research questions around both the process and progress of the study. Process-related questions include: *Are the samples of teachers and students in the user-experience studies representative and adequate? Are the newly developed tutor professional development materials, the AI-Agent, & new reports to teachers appropriately vetted and achieving their objectives?* Progress-related questions include: *Are project milestones being met as intended? Are the iterations of development improving quality and ease of use?* To address these questions, and to provide formative feedback during School Years 1 and 2, Dr. [redacted] will participate in regular meetings with the project directors, interview collaborators across institutions, teacher participants and co-developers, and other key stakeholders, conduct a comprehensive review of project materials and iterative development plans, and offer timely recommendations.

The **summative evaluation component** will include a rigorous randomized controlled trial (RCT) to measure the impact of TutorASSIST on 6th and 7th-grade math learning, as well as quality and fidelity of implementation of the intervention in 20 schools of diverse population and settings. The study will be led by Dr. [redacted] who is a WWC-certified reviewer and served as co-PI on the completed RCT that examined the efficacy of ASSISTments in Maine (IES R305A120125). She is currently serving as PI of the replication study of ASSISTments in North Carolina (IES R305A170641). She will be supported by Dr. [redacted] who is a WWC-certified reviewer and served as co-PI on I3 and EIR evaluations. Building on prior efficacy trials (see Appx. J.4.A), the study is designed to 1) yield impact estimates of TutorASSIST on learning outcomes of students who are identified as low-achieving; 2) study whether the Foundation of

TutorASSIST course for tutors, TutorASSIST, and tutor oversight are suitable for replication or testing with different populations, 3) examine implementation of key components in more detail, namely the TutorASSIST admin, tutor and teacher tools and their usage during and after tutoring sessions, the professional development for tutors and coordinators, and the AI-Agent, and 4) provide guidance to the scaling and replication of the innovative intervention in other settings.

E.1. Methods Designed to Meet WWC Standards Without Reservations

WestEd will conduct an independent impact study to address research questions that are aligned with the TutorASSIST logic model (see Appx. G), about the program's impact, implementation and cost effectiveness. RQ1 addresses the effects of TutorASSIST on students' mathematics achievement. RQ2 focuses on the differential effects of TutorASSIST on different types of students. RQ3 focuses on assessing implementation fidelity and quality.

Research Questions	Outcome	Data / Measures
(*RQ1) Outcomes*. Does receiving in-school tutoring via the TutorASSIST program improve students' mathematics learning outcome?	Students' mathematics achievement	State 6th- and 7th-grade assessments scores; Student and school demographics; student prior year test score
(RQ2) Moderators. Do the effects of TutorASSIST vary for students with different prior achievement, at different grade levels, and for students of different socio-economic status, race/ethnicity, or with other policy-relevant characteristics?		
(RQ3) Implementation fidelity and effects of implementation on student learning. a. Do participating schools implement TutorASSIST as intended by the developer? How many tutoring sessions occur? To what extent are each features used? b. What are the effects of implementation fidelity and dosage on learning? c. What are the factors that hinder or facilitate implementation?	Implementation fidelity and quality	Computer log data of usage and tutoring session interaction; Video recordings and analysis protocols; Tutor interviews and logs; School administrator interviews; Teacher survey.

Exhibit 9: Summative Research Questions, Outcomes, & Measures**.**

Study design: The study will use a student-level and delayed-treatment randomized controlled experimental design to meet What Works Clearinghouse standards without reservations (WWC, 2020). The RQs will be addressed with data collected during one school

year. Within each participating teacher, eligible 6th or 7th grade students will be identified and randomly assigned to a treatment condition to receive small group tutoring (1:2) from the volunteer tutors who receive a stipend for 20 weeks with 3 sessions per week during the study year. In the control condition, students will continue their regular classroom learning (business-as-usual), including any use of other online tools or services they are eligible for, but not tutoring from the TutorASSIST program. After the experiment is completed, students in the control group will also receive the same amount of tutoring via TutorASSIST in the next school year. Participating students will be identified within two weeks of the start of the school year and randomly assigned to conditions after appropriate parent consent and student assent is secured. Students who join study schools during the school year will be excluded from the evaluation sample. The team will track both overall and differential student-level attrition from both conditions (see Appx J.7 for Timeline of Impact Evaluation).

Sample and power: The team will recruit 20 public schools in Baltimore City Public Schools that have adopted ASSISTments and the OER curriculum for mathematics (See letters of agreement in Appx. C and demographic and state assessment proficiency overview of the partner district in Appx. J.8). We will gather data on student characteristics, such as ethnicity, ELL status, IEP, math and ELA proficiency on prior tests, socioeconomically disadvantaged) to improve precision and sample balance across conditions. We will focus on high needs (a.k.a., low-performing schools and lower-socioeconomic status) student bodies. Students whose state assessment performance levels are below or far below grade level will be eligible to be included in the study. A power analysis was conducted using PowerUp! (Dong & Maynard, 2013) for 2-level random effects blocked individual random assignment designs with assignment at the

student level. Assuming an average of three math teachers would be recruited at each school in 6th or 7th grade with 20 eligible students each, the estimated sample size of the study is 1,200 students and the proposed study has sufficient power to detect effects of 0.110 standard deviations on student achievement (see Appx J.9 for justification).

Data Analysis Plan: To estimate the impact of TutorASSIST tutoring on student achievement, the study team will conduct an intent-to-treat (ITT) analysis. The analysis of TutorASSIST impact will depend on the random assignment research design as its primary source of inference. Because the study design is a blocked trial, a two level hierarchical linear model will be used to analyze the treatment effect. Scaled state test scores in 6th and 7th grade will be used as the outcome measure. Analyses will test the overall impact of the intervention for 6th- and 7th-grade students, as well as the interventions' differential effects on students with different characteristics. To improve the precision of the impact estimate, each model will adjust for blocking (by teacher), prior year's state test score as a baseline measure, student grade level, other student and teacher-level characteristics (see details of the data analysis plan in Appx. J.10).

E.2. Generation of Guidance About Effective Strategies Suitable for Replication

The proposed study is designed to generate insightful guidance for successful replicating and scaling TutorASSIST in other districts broadly in four aspects. First, WestEd will conduct interviews with school administrators to **document the context of implementation**, such as district priorities, the community and broad socio-economic factors that may affect student learning outcomes, adoption status of tutoring services and supplemental education technology

programs, school-level policy, support and cultures of volunteer involvement, tutor collaboration, and professional learning.

Secondly, the evaluation included **differential impact analysis** (RQ2) to assess to what extent the impact of TutorASSIST is moderated by the characteristics of students (e.g., ethnicity, English-language learners, eligibility for free or reduced-price lunch, IEP). We will characterize schools (e.g., school size or type—K-8 vs. 6-8, locale, demographic composition, years adopted volunteer program) and identify for which settings or populations the intervention is particularly effective. To help interpret the effects between conditions, we will also collect data regarding the types of additional instructional services that our sample students receive, such as Tier-2 instruction, participation in a 504-plan, etc. during the study year.

Thirdly, the study builds in **analysis of implementation data from multiple sources** (RQ3). The essential features of implementation that should be followed by future efforts to replicate or scale this intervention include 1) tutors and school administrator participation in professional development training; 2) volunteer tutors providing tutoring sessions to students regularly following the designed tutoring model, leveraging supports from the AI-Agent; 3) administrators overseeing tutoring sessions and making adjustments accordingly; 4) teachers making assignments in ASSISTments so the tutors and the AI-Agent will have input data. Accordingly, data will be collected on all four aspects. We will examine implementation fidelity based on professional learning records, ASSISTments and TutorASSIST system backend logs of usage, video recordings of tutor sessions, as well as tutor logs (see Appx. J.11 for details and Appx. J.12 for data collection timeline). In particular, we will focus on the **fidelity of the scaling-up approaches**, following the framework proposed by Clarke & Dede (2009) and Coburn (2003)

and gather data around the five dimensions of scale that reflect different aspects of making an intervention effective in one setting useful across a wide spectrum of contexts, namely depth, sustainability, spread, shift, and evolution, and identify facilitators and barriers to implementation and scaling of each key component of the intervention.

Lastly, policymakers and educational administrators require information on the cost of the resources needed to adopt, implement particular instructional interventions with fidelity, and to sustain the interventions. To provide information about whether TutorASSIST is more cost effective, compared to business-as-usual tutoring practices, WestEd will conduct a cost analysis and calculate a cost-effectiveness ratio. Costs will be gathered systematically using the "ingredients method" (Levin and McEwan, 2001; Levin et al., 2017). Once all of the ingredients are identified, we will use *CostOut*--the CBCSE Cost Tool Kit (Hollands et al., 2015-19) to facilitate the estimation of costs and cost-effectiveness. (see Appendix J.13 for details of cost and cost-effectiveness analysis plan).

E.3. Valid and Reliable Performance Data on Relevant Outcomes

The evaluation team seeks to triangulate multiple data sources to reinforce validity of the findings. As shown in the logic model, the key student outcome of TutorASSIST is student academic achievement. The **student learning outcome measure** will include 6th and 7th-grade statewide mathematics assessment scores from spring 2025, using the relevant accountability measures of math achievement within each state. WestEd will secure agreements with participating districts to obtain the extant mathematics state test scores in 5th (as covariate), 6th (as covariate and outcome) and 7th (as outcome) grades, as well as demographics.

E.4. Components, Outcome, and Measurable Threshold for Implementation

The design of the evaluation is informed by the key components and student learning outcome as illustrated in the logic model in Appx. G. The key components of TutorASSIST on the tutoring side include (a) professional development and tutor community to help tutors learn about TutorASSIST and address their questions and concerns in a collaborative manner, providing AI-Agent support and pre-built content to make it easy for tutors to select relevant and timely practices during tutoring sessions, (b) supporting school administrators' in oversight of tutoring sessions. Key intervention components for students include scheduled high-dosage tutoring over the school year and supporting students' math learning with problem-solving help, SNS practices and encouragement on using online support.

The system records of all student, tutor and coordinator actions will serve as a primary source of data in assessing the extent to which TutorASSIST is used according to the specified use model. WestEd will also collect data on tutors and school administrators' attendance of professional learning sessions. Tutors will be asked to complete a short online log each time they deliver the tutoring to a student pair. The log will ask about the length of the session, content covered, practice assignments made in TutorASSIST, and a quick note of any issues. The logs will be designed to take 5 minutes to complete. We will analyze the logs together with the system record data to determine the dosage (the total number of sessions, how much content is covered overall, and the total amount of time students are engaged in tutoring). Additionally, we plan to interview 20 percent of tutors by videoconference, asking them to reflect on implementation issues, challenges and successes, students' response, supports or barriers that facilitate or hinder use of the tutoring, and what messages and tangible support they receive from school leaders or classroom teachers. The tutors will also be asked about their experience with the TutorASSIST

platform, students' capability of completing practice assignments, and if and how they integrate information from the reports of students' work into their tutoring practices.

WestEd will measure fidelity of implementation, using the five-step procedure for assessing intervention fidelity proposed by Nelson et al., (2012). Referring to the framework by Weiss, Bloom, & Brock (2014), WestEd will develop implementation metrics and set associated thresholds of implementation fidelity for each metric. There will be metrics for three types of **indicators of implementation compliance**— one on the extent to which tutors and school administrators complete in professional development and participate in community activities, one based on intensity and continuity of delivery of tutoring sessions, and a third on the extent to which tutors use the TutorASSIST reports and AI-Agent support (tutors) and school administrators' use of tutor oversight tools (administrators). The metrics will cover four categories of implementation fidelity associated with key components of the intervention, namely adherence, exposure, quality of delivery, and uptake (Carroll et al., 2007). The **minimum acceptable implementation** thresholds are defined as follows and will be adjusted as necessary after SY2: Tutors are expected to participate in the Foundations of TutorASSIST course accumulating at least 10 hours of professional learning time; School administrators are expected to attend the orientation of the TutorASSIST program; Tutors are expected to deliver at least 45 tutoring sessions (75% of the total expected sessions) to their committed tutee pairs; School administrators are expected to review attendance at least once per week and online sessions as needed. During the study year, WestEd will calculate the implementation metrics at the individual tutor level monthly and share the results with the developers to inform continuous improvement.

References

Aleven, V. & Koedinger, K. (2002) An effective metacognitive strategy: Learning by doing and explaining with a computer‐based cognitive tutor. *Cognitive science 26(2)*, 147-179. https://onlinelibrary.wiley.com/doi/pdfdirect/10.1207/s15516709cog2602_1

Bacher-Hicks, A., Goodman, J., Mulhern, C. (2020). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time (Working Paper 27555). National Bureau of Economic Research.

Baron, J. (2021). Jon Baron's gubernatorial campaign website whose first plank in his platform is to support volunteer tutoring programs. See https://jonbaron.com/

Bhatt, M. P., Guryan, J., Ludwig, J., & Shah, A. K. (2021). Scope challenges to social impact. (No. w28406). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w28531/w28531.pdf

Boston (2021). http://breakthroughgreaterboston.org/volunteer/

Brown University. (2021). Student Support Accelerator. https://studentsupportaccelerator.com/

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*. 2007;2:40

Carter, P. & Welner, K. (2013). *Closing the Opportunity Gap: What America Must Do to Give Every Child an Even Chance*. Oxford University Press; 0199982996.

Chi, M., Leeuw, M., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, Volume 18, Issue 3*, pp.439-477. ISSN 0364-0213, https://doi.org/10.1016/0364-0213(94)90016-7.

Clarke, J., & Dede, C. (2009). Robust designs for scalability. In L. Moller, J. B. Huett & D. M. Harvey (Eds.). *Learning and instructional technologies for the 21st century* (pp. 27-48) Springer, US.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational researcher*, 32(*6*), 3-12.

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-15-01. Retrieved June 10, 2021 from https://scholar.harvard.edu/files/fryer/files/not_too_late._improving_academic_outcomes_for_disadvantaged_youth_2015.pdf

Dadisman, K., & Schneider, M. (2020, December 2). *Dadisman & Schneider: Tutoring Can Help Reverse COVID-Related Learning Loss. 6 Principles for Doing It Right*. https://www.the74million.org/article/dadisman-schneider-tutoring-can-help-reverse-covid-related-learning-loss-6-principles-for-doing-it-right/

Darling-Hammond, L., Hyler, M. E., Gardner, M. (2017). *Effective Teacher Professional Development.* Palo Alto, CA: Learning Policy Institute.

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.

Dong, N., & Maynard, R. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies. *Journal of Research on Educational Effectiveness*, *6*, 24–67. https://doi.org/10.1080/19345747.2012.673143

Dweck, C. S., Walton, G. M., & Cohen, G. L. (2014). *Mindsets and skills that promote long-term learning.* Seattle, WA: Bill & Melinda Gates Foundation.

EdGlossary. (2013). Opportunity Gap.The Glossary of Education Reform. Created by the Great Schools Partnership.

EdReports. (2021). Open Up Resources 6-8 math and Eureka Math (2015) are both highly rated. See https://www.edreports.org/compare/results/math-k-8

Engzell, P., Frey, A., & Verhagen, M. D. (2020, October 29). Learning inequality during the COVID-19 pandemic. SocArXiv Papers. https://doi.org/10.31235/osf.io/ve4z7

Finkelstein, A. Katz, L., Bates, M. A. (2019). The 2019 Nobel Prize-winning scientific movement that can transform US policymaking. MIT. https://www.povertyactionlab.org/blog/12-10-19/2019-nobel-prize-winning-scientific-movement-can-transform-us-policymaking

Fried, L. P., Carlson, M. C., McGill, S., Seeman, T., Xue, Q. L., Frick, K., ... & Rebok, G. W. (2013). Experience Corps: a dual trial to promote the health of older adults and children's academic success. *Contemporary Clinical Trials*, *36*(1), 1-13. https://doi.org/10.1016/j.cct.2013.05.003

Fryer Jr, R. G. (2017). The production of human capital in developed countries: Evidence from 196 randomized field experiments. In *Handbook of Economic Field Experiments* (Vol. 2, pp. 95-322). North-Holland. https://scholar.harvard.edu/files/fryer/files/handbook_fryer_03.25.2016.pdf

Guryan, J., Ludwig, J., Bhatt, M., Cook, P., Davis, J., Dodge, K., Farkas, G., Fryer, R., Mayer, S., Pollack, H., & Steinberg, L. (2021). Not Too Late: Improving Academic Outcomes Among Adolescents. National Bureau of Economic Research, Working Paper Series, 28531. http://www.nber.org/papers/w28531

Haimovitz, K., Wormington, S. V., & Corpus, J. H. (2011). Dangerous mindsets: How beliefs about intelligence predict motivational change. *Learning and Individual Differences, 21,* 747-752.

Hollands, F. M., Hanisch-Cerda, B., Levin, H. M., Belfield, C.R., Menon, A., Shand, R., Pan, Y., Bakir, I., & Cheng, H. (2015-19). *CostOut - the CBCSE cost tool kit.* New York: Teachers College, Columbia University, Center for Benefit-Cost Studies of Education. www.cbcsecosttoolkit.org

Intrado GlobeNewswire. (2020, July 22). Global Private Tutoring Market to Reach US$279. 3 Billion by the Year 2027. Amid the COVID-19 crisis, the global market for Private Tutoring estimated at US$173.4 Billion in the year 2020, is projected to reach a revised size of US$279. https://www.globenewswire.com/news-release/2020/07/22/2065601/0/en/Global-Private-Tutoring-Industry.html

J-PAL, Abdul Latif Jameel Poverty Action Lab. (2021). *Tutoring Pre K-12: Evidence and Resources.* https://www.povertyactionlab.org/tutoring

Kleiman, G., & Wolf, M. A. (2015). *Going to scale with online professional development: The Friday Institute MOOCs for Educators (MOOC-Ed) initiative.* North Carolina State University College of Education.

Kraft, M. A. & Goldstein, M. (2020, May 21). Getting tutoring right to reduce COVID-19 learning loss. Washington, DC: Brookings Institution, Brown Center Chalkboard.

Kraft, M. & Grace F. (2021). A Blueprint for Scaling Tutoring Across Public Schools. (EdWorkingPaper: 20-335). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/dkjh-s987

Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020). Learning during COVID-19: Initial findings on students' reading and math achievement and growth. NWEA. https://www.nwea.org/content/uploads/2020/11/Collaborative-brief-Learning-during-COVID-19.NOV2020.pdf

Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., & Busick, M.D. (2012). *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms.* (NCSER 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf

Levin, H. M. (1975). Cost-effectiveness analysis in evaluation research. In M. Guttentag, & E. L. Struening (Eds.), *Handbook of evaluation research* (volume 2). Beverly Hills, CA: Sage.

Levin, H. M. (1983). *Cost-effectiveness analysis: a primer*. Beverly Hills, CA: Sage Publications.

Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: methods and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Levin, H.M., McEwan, P.J., Belfield, C., Bowden, A.B., & Shand, R. (2017). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. Thousand Oaks, CA: Sage Publications.

Maldonado, J. E., & DeWitte, K. (2020). The effect of school closures on standardized student test outcomes. Discussion Paper Series DPS20.17, Katholieke Universiteit (KU) Leuven.

MIT Poverty Action Lab. (2021). Tutoring Pre K-12: Evidence and Resources. https://www.povertyactionlab.org/tutoring

National Center for Education Statistics. (2020a). NAEP report card: Reading. National achievement-level results (Data compiled from National Assessment of Educational Progress, various years, 1992–2019 Reading Assessments). Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2020b). NAEP report card: Mathematics.National achievement-level results (Data compiled fromNational Assessment of Educational Progress, various years, 1990–2019 Mathematics Assessments). Institute of Education Sciences, U.S. Department of Education. https://www.nationsreportcard.gov/mathematics/nation/achievement?grade=8

Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *The Journal of Behavioral Health Services & Research, 39*(4), 374–396.

Nickow, A. J., Oreopoulos, P., & Quan, V. (2020a). The Transformative Potential of Tutoring for Pre K-12 Learning Outcomes: Lessons from Randomized Evaluations | The Abdul Latif Jameel Poverty Action Lab. *Abdul Latif Jameel Poverty Action Lab (J-PAL)*. https://www.povertyactionlab.org/sites/default/files/publication/Evidence-Review_The-Transformative-Potential-of-Tutoring.pdf

Nickow, A. J., Oreopoulos, P., & Quan, V.(2020b) The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence. NBER Working Paper No. 27476. July 2020. JEL No. I2,J24. https://www.nber.org/system/files/working_papers/w27476/w27476.pdf

Powell, C. G., & Bodur, Y. (2019). Teachers' perceptions of an online professional development experience: Implications for a design and implementation framework. *Teaching and Teacher Education*, 77, 19-30.

Ritter, G., Barnett, J, Denny, G & Albin, G. (2007). The Effectiveness of Volunteer Tutoring Programs for Elementary and Middle School Students: A Meta-Analysis. *Review of Educational Research, Volume: 79 issue: 1,* page(s): 3-38. https://journals.sagepub.com/doi/10.3102/0034654308325690

Roschelle, J., Cheng, B. H., Hodkowski, N., Neisler, J. & Haldar, L. (2020). Evaluation of an online tutoring program in elementary mathematics [Project Report]. San Mateo, CA: Digital Promise. Retrieved from: http://hdl.handle.net/20.500.12265/95

Roschelle, J., Feng, M., Murphy, R., & Mason, C. (2016). Online mathematics homework increases student achievement. *AERA Open Oct 2016, 2 (4)* 2332858416673968; DOI: https://doi.org/10.1177/2332858416673968

Rutt, S., Easton, C., & Stacey, O. (2014). Catch Up® Numeracy: Evaluation report and executive summary. London, UK: Education Endowment Foundation. https://www.nfer.ac.uk/publications/EFCU01/EFCU01.pdf

Safran, A., Gutierrez, A. J., Dadisman, K. (2021). Everyone's talking about a national tutoring corps; here's what we need to know to do it well. *The Hill*. https://thehill.com/blogs/congress-blog/education/539786-everyones-talking-about-a-national-tutoring-corps-heres-what-we

Saga. (2021). Sample Lesson & Activity. https://studentsupportaccelerator.org/sites/default/files/Sample_Lesson_Activity_Saga_Education.pdf

Schneider, M. (2021, May 4). Catching Up When You Started Behind. https://ies.ed.gov/director/remarks/5-4-2021.asp

Schneider, M. (2020, October 26). Introducing *Operation Reverse the Loss*. https://ies.ed.gov/director/remarks/10-26-2020.asp

Seel, N.M. (2012). Attribution Theory of Motivation. In: Seel N.M. (eds.) *Encyclopedia of the Sciences of Learning*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-1428-6_917

ServeMinnesota. (2018, November 2). A Year-Long State-Wide RCT of the Minnesota Math Corps: Final Report to Laura and John Arnold Foundation. https://osf.io/7s4bv/

Slavin, R. (2020). Slavin: An Open Letter to President-Elect Biden—a Tutoring Marshall Plan to Heal Our Students. https://www.the74million.org/article/slavin-an-open-letter-to-president-elect-biden-a-tutoring-marshall-plan-to-heal-our-students/

Somerville (2021). Somerville Public Schools. Retrieved from https://somerville.k12.ma.us/district-departments/volunteering/start-volunteering

Stein, P. (2020, October, 30). Washington Post: In D.C., achievement gap widens, early literacy progress declines during pandemic, data show. The DC Line: News, Views and Events. https://thedcline.org/2020/10/30/washington-post-in-d-c-achievement-gap-widens-early-literacy-progress-declines-during-pandemic-data-show/

Styers, M. & Baird-Wilkerson, S. (2011). A final report for the evaluation of Pearson's focusMATH Program. *Magnolia Consulting*. https://moam.info/a-final-report-for-the-evaluation-of-pearsons-focusmath-program_5a24daa11723ddae6a5a7f69.html

US Census. (2017). Current Population Survey, September 2017: Volunteering and Civic Life Supplement Technical Documentation. https://www2.census.gov/programs-surveys/cps/techdocs/cpssep17.pdf and https://www.icpsr.umich.edu/web/NADAC/studies/37303/datasets/0001/variables/PES16?archive=NADAC

VanLehn, K., Siler, S., Murray, R. C., Yamauchi, T., & Baggett, W.B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction, 21(3)*, 209-249. http://www.public.asu.edu/~kvanlehn/Stringent/PDF/03CI_KVL_SS_CM_TY_WBB.pdf

Weiss, M., Bloom, H. S., & Brock, T. (2014). A conceptual framework for studying the sources of variation in program effects. *Journal of Policy Analysis and Management, 33(3),* 778-808. https://doi.org/10.1002/pam.21760

What Works Clearinghouse (WWC). (2020). *What Works Clearinghouse: Procedures and Standards Handbook, Version 4.1.* Washington, DC: Author. Retrieved from: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Zimmer, A. (2020, December 22). Online tutors are helping NYC students catch up. But expanding these programs remains costly. ChalkBeat NY City. https://ny.chalkbeat.org/2020/12/22/22193054/nyc-free-tutoring-aoc-stuyvesant