TABLE OF CONTENTS

Scaling ASSISTments for Algebra Readiness with Curriculum-based Professional Learning and Coaching

The ASSISTments Foundation (TAF), the non-profit behind ASSISTments--an online math intervention tool—in partnership with Teaching Lab and WestEd, proposes a mid-phase Education and Innovation Research grant to implement, refine and evaluate ASSISTments for Algebra Readiness (ASSISTments4AR). ASSISTments4AR will be designed to directly impact the Algebra Readiness of high-need 6th and 7th-grade students. Algebra I is one of the most important on-track indicators of students' future success. Students who do not complete Algebra I have only a 20% chance of graduating high school (Gates Foundation, 2022). To be ready for Algebra I by 8th grade, 6th and 7th-grade students need to be proficient in grade-level standards. Yet too often, high-need students face what has been termed the "opportunity gap" (EdGlossary, 2013). Instead of being presented with rigorous, grade-level materials, teachers give these students content that is below grade-level, causing many experts to charge that even using the term "achievement gap" is misleading as high-need students were never even presented with the same opportunity to challenge themselves (Carter & Welner, 2013).

We will partner with school districts that have adopted either EngageNY or Illustrative Math, two high-quality instructional materials (HQIM) (EdReports, 2021) used by almost half of the middle school math teachers in the country (Doan, et al., 2021). Our program combines three evidence-based components that will drive student success with this rigorous grade-level content: (1) Teacher-use of ASSISTments to provide students just-in-time targeted support with grade-level standards, (2) Evidence-based teacher professional learning (PL) needed to successfully implement their HQIM and (3) Capacity-building for school coaches to more effectively support teachers, enhanced with data from ASSISTments-Coach. As school districts

increasingly offer Algebra I in 8th grade, our focus is on ensuring 6th and 7th-grade students are ready (U.S. Department of Education, 2018 November).

Component 1 centers on a forever free tool we call ASSISTments-Teacher, already funded by an EIR mid-Phase, that showed reliable learning gains in a study with a What Works Clearinghouse rating of "positive without reservation" (Roschelle et al, 2016). Teachers will use ASSISTments-Teacher to assign students problems online from their HQIM (Appx. J.1.A.1-3). Teachers will access real-time data on student performance via data dashboards (Appx. J.1.A 4 & 5). Students will log in to their learning management system (LMS) (Appx J.1.B.1) click on the link to the assignments and receive immediate feedback as they work (Appx. J.1.B.2 and 3).

While TAF has strong experience preparing teachers to use ASSISTments-Teacher successfully as we grew to 20,000 users in SY2021, the evidence suggests successful use of HQIM additionally requires curriculum-specific professional learning (PL). This second crucial component comes through our partnership with Teaching Lab, a national nonprofit that has developed best-in-class PL content aligned to Illustrative Math and EngageNY that has earned them praise from national experts and the curriculum authors themselves (Miller & Partelow, 2019; Wiener & Pimentel, 2017; The Teaching Lab, 2020). Teaching Lab will focus on building pedagogical content knowledge, and empowering teachers with evidence-based instructional strategies that promote student success with grade-level work (Component 2). In tandem with teacher PL, Teaching Lab Coaches will build the capacity of School Coaches to sustain this high-quality support (Component 3). This Coach-the-Coach model will be further enhanced with data from ASSISTments-Coach, a recently developed companion to ASSISTments-Teacher that provides real-time data to coaches on teacher and student progress (Appx. J.2).

In Years 1 and 2, we will pilot and refine the intervention in partnership with a high-need school district that has adopted either Eureka or Illustrative Mathematics (See Letter of Support Appx. C). In Years 3 and 4 we will study the effects of the intervention by conducting a large-scale, school-level randomized controlled trial to test the efficacy of ASSISTments4AR on student learning outcomes (See logic model in Appx. G). Year 5 focuses on scaling the impact.

ABSOLUTE AND COMPETITIVE PRIORITIES

Absolute Priority 1 - Moderate Evidence

Teachers' use of ASSISTments to support their instruction has been demonstrated with moderate evidence to be effective at improving student achievement in middle school math, especially for students with lower prior achievement. SRI conducted an independent evaluation of 43 schools, 87 teachers, and 2,769 seventh grade students and reported three major findings: 1) Students who consistently completed assignments in ASSISTments performed significantly better than those whose teachers followed regular practices, and did not use ASSISTments, on a standardized end-of-year math achievement assessment (effect size g = 0.22), representing an additional two-thirds of a year of gain in math learning, 2) Students with lower prior mathematics scores experienced a greater benefit from the intervention (effect size $g = 0.29$), meaning the use of ASSISTments also narrowed the achievement gap, and 3) ASSISTments, when paired with just 3 days of professional development over a school year, reliably changed teacher practice. Teachers in the intervention group more effectively targeted class time to student difficulties and errors ([redacted]). This study has been reviewed by the WWC (What Works Clearinghouse, 2019) and was designated with "positive effect without reservations" making this meet the evidence standards (See the Evidence Form

for details). Given this evidence, ASSISTments received a Strong Rating from Evidence for ESSA (Evidence for ESSA, 2020). ASSISTments is only 1 of the 3 interventions in middle school math to receive this rating, making it a highly attractive choice for school districts looking for evidence-based solutions that will mitigate the impacts of COVID-19 on high-need students in their district. Appx. J.5 has additional evidence of the effectiveness of ASSISTments as well as professional development with coaching.

Absolute Priority 3—Promoting Equity in Student Access to Educational Resources and Opportunities in STEM

ASSISTments for Algebra Readiness (ASSISTments4AR) will build on the demonstrated impact of our evidence-based intervention, now called ASSISTments-Teacher, with our coaching tool (ASSISTments-Coach), and teacher professional learning and coaching. This program is specifically designed to impact the math achievement of my high-need 6th and 7th-grade students by addressing a key source of inequality in math education: Limited opportunity to engage with grade-level content.

While Common Core standards established what students need to master at each grade level, and there is greater access to aligned instructional materials, multiple analyses show that high-need students are not provided assignments that give them the opportunity to achieve on grade level (EdReports, 2018). TNTP's Opportunity Myth found students spend more than 500 hours per school year on assignments that weren't appropriate for their grade, the equivalent of 6 months of lost learning time (TNTP, 2018). Ed Trust's examination of more than 1,800 classroom assignments shows only 1 in 10 middle-school math assignments required high levels of cognitive demand (Education Trust, 2018). Teachers with the best of intentions tend to focus

on lower-level content with the belief that high-need students, given unfinished learning, are incapable of being successful with grade-level content.

The evidence also suggests when given access to grade-level content, students rise to the occasion. In the same study, TNTP also found that for students with unfinished learning, access to grade-level tasks increased learning by as much as 7.3 more months (TNTP, 2018). Recent research leveraging data from over two million students using Zearn Math had a similar finding: classes that showed the strongest learning gains received 27% more grade-level lessons than the comparison group (NTP, 2022).

ASSISTments4AR is a program in which every element is designed to increase student access and success with rigorous grade-level content, which we refer to as High-Quality Instructional Materials (HQIM). HQIM are materials that assess grade-level content, giving all students extensive work with grade-level problems, and are coherent and consistent with rigorous standards (EdReports, 2020). The impact of using ASSISTments-Teacher has already demonstrated an impact on equity. Roschelle et al. (2016) showed that while all participating students benefited from ASSISTments, students scoring below the median on the prior year's state math tests gained the equivalent of more than two years' worth of knowledge while using ASSISTments for one year (following Lipsey et al, (2012) recommendation on how to convert an effect size into more readily interpretable forms). Since this study, we have integrated two HQIM, Illustrative Math and EngageNY, into our platform, and now 80% of ASSISTments teachers report using our tool as an enhancement to these materials.

Increasing access to HQIM alone, even through ASSISTments, will not lead to the dramatic achievement gains we seek. In a multi-state effort to measure textbook efficacy since the

implementation of the Common Core, researchers found that just 25% of teachers use their designated curriculum in nearly all their lessons (Blazar et al., 2019). In short, at current levels of usage, curriculum choice alone does not improve student achievement. A key recommendation is to increase the amount of PL teachers receive on their HQIM (the study found teachers only receive an average of 1.1 days) (Blazar, et al, 2019).

This makes sense, given decisions about what to teach starts with the teacher, based on their expectations for high-need students, and their capacity to use HQIM effectively. Teaching Lab's PL and coaching model centers on shifting teacher mindsets and behaviors to center on access to rigorous learning content as a matter of equity. To date, Teaching Lab has worked with educators from over 40 schools and school districts that serve over 1 million students combined (63% of the students Teaching Lab impacts are eligible for free/reduced price lunch, and 38% identify as Black and 35% as Hispanic (Evidentally, 2022).

More students in Teaching Lab partner sites receive opportunities to complete grade-level assignments compared to students nationwide. While TNTP's Opportunity Myth (2018) found that four out of ten classrooms with a majority of students of color never received a single grade-level assignment, analyses of over one thousand student work samples from partner schools, serving student populations which are 65% students of color, revealed that math teachers receiving Teaching Lab professional learning assigned grade-level tasks 98% of the time (Teaching Lab, 2020).

Competitive Preference 1 - Promote equity and adequacy in educational opportunity and outcomes

Crucial to the ASSISTments4AR model is the use of technology, specifically ASSISTments-Teacher and ASSISTments-Coach, to tailor instruction to student needs in

real-time. With ASSISTments-Teacher, teachers will collect crucial data on student understanding of pre-requisite standards, by assigning curriculum-aligned diagnostics before teacher grade-level content. Rather than spending whole class periods re-teaching missed concepts from prior years in isolation, they will use this data to strategically integrate review of the most important prerequisite concepts into their lesson, at the time when students will most need them to access grade-level work.

Even with this approach, after teaching the lesson, it's very likely specific students, or the entire class will still require additional support to master the grade-level standards. For this reason, teachers will assign the aligned practice problems, exit tickets, and homework from their HQIM through ASSISTments. Teachers will get immediate data they can use to plan and target their lesson the next day. Teachers have the flexibility to assign additional practice from ASSISTments content library of standards-aligned problems to students as needed for ongoing practice and progress monitoring.

Teaching Lab Coaches, as well as School Coaches, will be able to better support teachers in this planning, by viewing the data via ASSISTments-Coach (see Exhibit 1 which is part of Appx. J.2.E.3). Looking at the data in Exhibit 1, a coach would be able to see Mr. Robinson's Class A is struggling with key concepts within Lesson 2, "Using the Number Line", of their current unit.

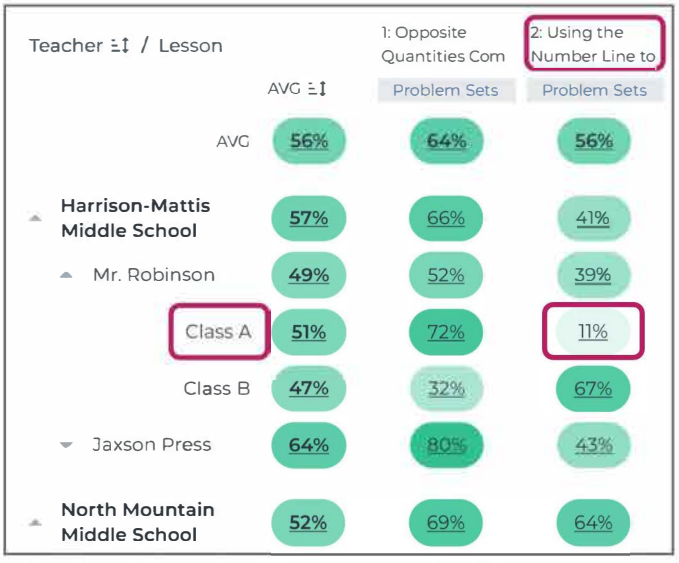

Exhibit 1: Example of the data a coach will use to support a teacher.

Given the majority of the class is struggling, they would focus on whole-class re-engagement strategies for that standard. Whereas for Class B, the coach would help Mr. Robinson plan a small group lesson, or consider those students for additional intervention outside of core instruction (e.g. tutoring). The ASSISTments data allows for all stakeholders to be more efficient and targeted in meeting student needs, making it more feasible to engage students in remedial learning without compromising access to grade-level work.

Competitive Preference 2 - Underserved students most impacted by COVID-19:

Dorn et al. (2021) estimate that without addressing student learning gaps now, the learning and economic impact of COVID-19 could be devastating and could last for decades. 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). While the pandemic has created greater urgency, the need to improve students' math achievement, specifically for high-need students is well-recognized. Two-thirds of American 8th-graders from low-income families failed to acquire basic math proficiency and the picture is worse for students of color (NCES, 2020a, 2020b).

ASSISTments4AR will be designed to specifically address the impacts of COVID-19, with an essential focus on empowering and strengthening the instructional practices of teachers and the coaches that support them. Teachers are one of the most important school-based factors that impact student achievement (Goldhaber, 2021). Students of effective teachers are more likely to experience favorable long-term outcomes, such as attending college and earning higher salaries

(Chetty et al., 2014). In our independent evaluation, teachers were able to impact math achievement with ASSISTments after three days of PL. At the same time, our sample was a majority white students (93%), and 39% received free and reduced lunch. In this program, we seek to serve a sample of at least 60% high-need students, including substantially more students of color. Additionally, we know COVID-19 has deeply impacted student achievement. Teaching is more complex than ever and well-designed teacher PL and coaching is essential for ensuring teachers have the mindsets, knowledge, and practices necessary to improve student learning (Guskey & Yoon, 2009; Lynch et al., 2019).

For this reason, Teaching Lab's PL and coaching are crucial to our model, providing teachers with evidence-based instructional strategies directly relevant to their HQIM (Lynch et al., 2019). Teaching Lab will design a thorough and cohesive plan for providing PL and instructional coaching to support teachers and school coaches that includes 1) Teacher PL, consisting of Bootcamp (3 days) and 3 Cycles of Inquiry (2 days each) and 2) 15 days of capacity building support for School Coaches on providing effective feedback to teachers. Key learning goals include improving teachers (1) knowledge of the mathematics needed to prepare students for Algebra; and (2) ability to analyze student work and achievement, and adjust instructional strategies, assessments, and materials based on such analysis.

Teaching Lab will customize the support based on a comprehensive Assets and Needs Assessment conducted at the beginning of the year at the school and district levels (see Appx. J.6 for the complete needs assessment). This assessment includes collecting baseline data on system-wide investment in and commitment to HQIM as a vehicle for educational equity, and the systems in place for sacred time (and resources) for PL and observation, feedback and coaching

cycles for teacher growth. As part of the Assets and Needs Assessment, Teaching Lab will identify subgroups of students most at risk and, together with stakeholder input, propose specific strategies to support these students and their families.

A. SIGNIFICANCE

A.1 The National Significance of the Proposed Project

Algebra I is one of the most important on-track indicators for college and career readiness, and results are in decline. As one example, in Texas, in comparing 2019 to 2021, the percentage of students scoring on grade level in Algebra I fell dramatically among Black students, from 53% to 28%; Latino students, from 64% to 34%; and students considered "economically disadvantaged" by the state from 59% to 31% (Carpenter, 2021). The benefits of passing Algebra I in 8th grade have been well documented, and districts are increasingly offering this course early (US Dept of Ed, 2018). Given this, our program is specifically designed to ensure student proficiency on 6th and 7th grade Common Core standards, in turn preparing them for Algebra I.

The model is timely given the appetite for technology solutions is at an all-time high (80% of teachers report they will need to continue to use different technologies in their work) (CDT, 2021). It is critical we understand how proven tools like ASSISTments can be paired with, and further enhance, teacher PL and coaching, so teachers are empowered to address the daunting challenge of catching students up while continuing to teach grade-level content.

A.2 Extent to which Proposed Project Develops a Promising New Strategy

ASSISTments4AR will integrate multiple evidence-based components into a highly cohesive and aligned intervention, with HQIM as the foundation. HQIM is known to be an effective solution to improving achievement (Walker, 2017). ASSISTments-Teacher, proven effective,

infuses two learning science principles into classrooms: immediate feedback for students, and actionable data for teachers to target instruction. Teaching Lab's PL and coaching address crucial well-documented barriers to advancing achievement in STEM through the implementation of HQIM, including lack of investment in teacher PL, poor content preparation, and poor content delivery and methods of assessment. (Ejiwale, 2013).

The addition of ASSISTments-Coach creates an opportunity to understand how to harness technology to enhance teacher PL and coaching, a new strategy with some initial evidence. Researchers from the Education Development Trust conducted six case studies on the design and delivery of technology-enabled PL, finding promising results for teachers and student learning. Researchers highlighted the importance of high-quality impact data and careful piloting of solutions, many of the same strategies which we are incorporating into this proposed project (McAleavy et al., 2018).

A.3. The Potential of the Proposed Project to Increase Knowledge and Understanding of Educational Problems

This project, and the resulting study, would be an important contribution to the knowledge base on the features of PL and coaching models that support the effective implementation of HQIM to impact high-need students with significant learning loss in math. According to the Evidence for ESSA, there are three times as many professional development interventions (35) with strong evidence in language arts (Evidence for ESSA, 2022a) as compared to math professional development programs (12) (Evidence for ESSA, 2022b), and little is known about professional development specifically supporting the effective use of HQIM.

As one example feature, the specific scope and sequence Teaching Lab is designing for this program will include an emphasis on the strategies and approaches outlined in the What Works

Clearinghouse Practice Guide which yielded ***Tier 2 Moderate Evidence***, "Teaching Strategies for Improving Algebra Knowledge in Middle and High School Students," such as: teaching students to recognize and generate strategies for solving problems and teaching students to utilize the structure of algebraic representations (Star et al., 2015). This will shed additional light on the impact of professional learning and coaching that embed these strategies to impact student math achievement in the context of HQIM.

B. STRATEGY TO SCALE

B.1 Strategies That Address Barriers To Scale

We have identified these three barriers we must overcome in order for ASSISTments4AR to impact Algebra I readiness for high-need middle school students at scale. Exhibit 2 outlines the three barriers, with strategies to address each barrier explained in greater detail below.

Barrier	Strategy
1. ASSISTments-Teacher lacks features needed for true scalability in alignment with HQIM	Enhance ASSISTments-Teacher with features that support scalability with HQIM
2. PL providers like Teaching Lab and school coaches lack real-time data needed to effectively support the implementation of HQIM to support Algebra I readiness	Enhance ASSISTments-Coach to provide Teaching Lab and School coaches crucial data to support the implementation of HQIM for Algebra I Readiness
3. Teachers lack access to evidence-based, curriculum-aligned PL and coaching that tightly aligns to Algebra I readiness and leverages their classroom data	Develop a PL and coaching scope and sequence within TLs existing framework with an emphasis on use of ASSISTments data

Exhibit 2: Barriers and Strategies to Scale ASSISTments as a comprehensive solution that addresses Algebra I Readiness

Strategy 1 - Enhance ASSISTments-Teacher with features that support scalability with HQIM

ASSISTments was designed for teachers based on learning science principles, and has demonstrated efficacy, but was not designed specifically in alignment with Illustrative

Mathematics and EngageNY. For the impacts of ASSISTments-Teacher to scale broadly in school districts that have adopted one of these HQIM, we have identified two specific high-leverage product improvements.

Ungraded Answer Types- Currently, ASSISTments-Teacher has two distinct types of questions within the platform: Auto-scoring (Appx.J.1.B 2 & 3) and Open Response (Appx.J.1.B.4). Auto-scoring items are powerful in that they are able to provide students immediate feedback, and students are able to try again. Teachers save time grading and are able to dive into actionable data to plan instruction. This grant will support the engineering needed to add additional auto-graded answer types to ASSISTments-Teacher in alignment with the types of problems found in each of these curricula, roughly halving the number of open-response items from 40% to 20% (see examples in Appx. J.3.A.a & b). In addition, we will add more dynamic and interactive ways for students to show their work on the platform. Whether for open response or auto-scorable questions, we know seeing students thinking is invaluable, and want to make this as easy as possible. This is particularly important given that analyzing student work samples is a key evidence-based practice within Teaching Labs' PL.

Compatibility with any Learning Management System- The final goal of this project is to develop ASSISTments4AR as a promising new program for Algebra Readiness that will be available to all schools that need it. In order to do this, we need to improve ASSISTments' compatibility with other technology systems districts use. Currently, ASSISTments can only be used by schools that have adopted Google Classroom and Canvas (see images of our existing product Appx. J.3.B that shows how we will add support for new LMSs). The lack of access to all LMS options becomes an immediate roadblock to scale. With this funding, we will engineer a

version of both ASSISTments-Coach and Teacher that is compatible with any of the commonly used Learning Management systems (i.e., Microsoft's LMS365, Power Schools's Schoology, Infinite Campus's Campus Learning) and that can be inserted in single-sign-on platforms like Clever or Classlink, for ease of use with other key systems the district uses.

Strategy 2 - Enhance ASSISTments-Coach to provide Teaching Lab and School coaches crucial data to support implementation of HQIM for Algebra I Readiness

ASSISTments-Coach has already been developed and deployed to schools (Appx. J.2), providing views of data on teacher curriculum implementation and student progress and achievement on assignments. In collaboration with Teaching Lab, we have identified **three additional data views** that will improve the ability of both Teaching Lab and School Coaches to use ASSISTments-Coach, that will make it easier for coaches to provide effective support to teachers with their curriculum.

The first is the ability to view the proportion of problems assigned by a teacher that are grade-level versus remedial. This will support discussions during coaching around the importance of providing all students the opportunity to engage with grade-level material, and strategies for scaffolding and tailoring such content based on formative data on student learning (see Appx. J.4.A). **The second** is the ability to view student progress on standards over time (see Appx. J.4.B & C). This allows Coaches to monitor student progress and provide just-in-time support when students need it most and focus on standards that connect to equations and expressions, which research demonstrates are slightly more important for Algebra I readiness (Klute, 2021). **The third** is a new student work analysis view. Appx. J.4.D shows images of designs for this new feature, which will streamline the ability of coaches to analyze student work. Coaches will use this view as a tool in PL and coaching sessions to unpack student

misconceptions and plan teaching strategies that leverage student work to engage students in mathematical discussions that build their understanding.

Strategy 3 - Develop a professional learning and coaching scope and sequence within TL's existing framework that integrates the use of ASSISTments data

During the first two years of the grant, Teaching Lab will engage in deep design work with pilot school partners and the ASSISTments Foundation (TAF) to modify their PL and coaching to leverage data from ASSISTments to enhance their approach. Exhibit 3 shows three examples of modifications Teaching Lab will make (See detailed images in Appx. J.4.A, B, C & D), given the available data from ASSISTments-Teacher and ASSISTments-Coach, to further improve teachers' pedagogical practices and student learning.

Existing Teaching Lab Routine	The Routine Enhanced with ASSISTments Data
Student assignments are collected, manually scored, and analyzed twice a year for some partners. Teaching Lab determines whether the task is on grade level and if so, manually scores the student proficiency on that task. Results are shared with those partners at the end of the year.	ASSISTments-Coach continuously captures and visualizes grade-level performance on assigned problems, thereby informing the effectiveness of shifts in teachers' instructional practices on student learning in real-time. Previously uncaptured data will also be shown in the form of insight into what was actually assigned in class, at grade level or not at grade level (Appx. J.4.A-C)
In Teaching Lab's "coach the coach" model, Teaching Lab coaches accompany school coaches to observe teachers' instructional practices, design individualized teacher improvement plans, and monitor growth.	Teaching Lab coaches and school coaches can pair instructional practice data through classroom walkthroughs with ongoing student performance from ASSISTments-Coach (Appx. J.2). This will allow for a much more comprehensive approach to monitoring and guiding teacher improvement as it specifically relates to students' Algebra readiness (Appx J.3.D).
Teachers are asked to bring student work samples and/or other evidence (e.g., field notes on implementing an evidence-based instructional practice) for reflection and discussion in group PL.	Teaching Lab will continue to do evidence-based reflection but work samples will now also be available to the coaches using the ASSISTments-Coach's new student work analysis view before the sessions (Appx. J.4.D).

Exhibit 3: Examples of how the routines that currently exist in the Teaching Lab Model will be enhanced by ASSISTments Data.

B.2 Adequacy of the Management Plan

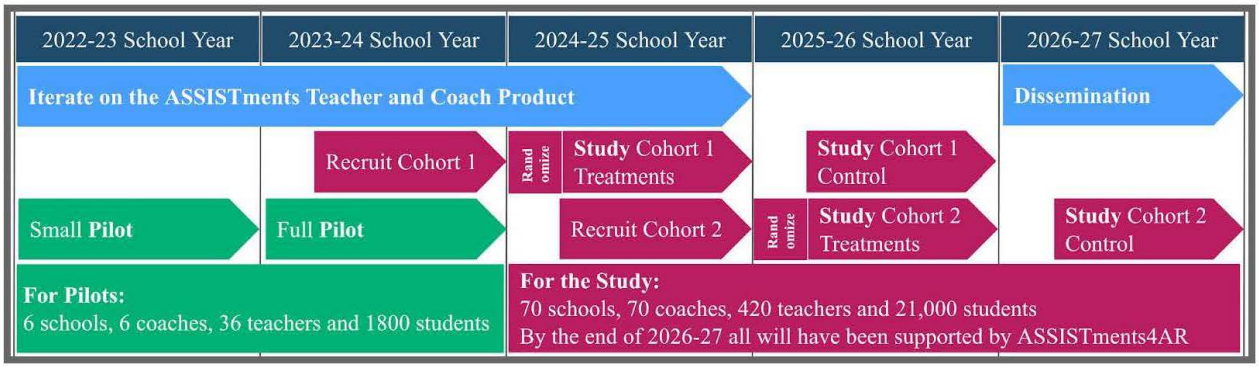

Exhibit 4: Timeline

Milestones	SY 1	SY 2	SY 3	SY 4	SY 5	Who is responsible?	Other
Objective 1. Enhance ASSISTments-Teacher and ASSISTments-Coach to be optimized for scale with HQIM							
Add Auto Grading answer types	x	x				**TAF,** TL	PS, SC, ST
Integration with multiple LMS	x	x				**TAF**	PS
Iteratively Improve Coach Data Views	x	x				**TAF,** TL	PS, SC, ST
Objective 2. Modify Teaching Lab's evidence-based PL and coaching scope and sequence to leverage ASSISTments							
Adapt PL Bootcamp and Cycles	x	x				**TL,** TAF	PS, SC, ST
Iteratively Improve Coaching	x	x				**TL,** TAF	PS, SC
Assess Feedback	x	x				**TL,** TAF	PS, SS, SC, ST
Objective 3. Rigorously evaluate the impact of ASSISTments4AR							
Recruit 70 schools and randomize		x	x	x		**TL,** WE	SS
Implement with treatment schools			x	x	x	**TL,** TAF, WE	SS, SC,ST
Prepare Data Collection		x				**WE,** TL, TAF	
Measure Implementation Fidelity			x	x	x	**WE,** TL, TAF	SS, SC, ST
Measure Student Outcomes			x	x	x	**WE,** TAF	SS, SC, ST
Analyze Cost-Effectiveness				x	x	**WE,** TL, TAF	SS
Objective 4. Disseminate findings and support scale of ASSISTmetns4AR							
Disseminate Learnings				x	x	**TAF,** WE, TL	PS, SS
Develop Business Plan				x	x	**TAF, TL**	
Key Institutions and Teams: lead in bold, TAF-The ASSISTments Foundation; WE-WestEd; TL-Teaching Lab; PS-Pilot School; SS-Study Schools SC-School Coach; ST-School Teachers							

Exhibit 5: Milestones, Timeline, and Responsibilities

B.3. 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 experience in technology development, professional learning, math education, Algebra Readiness, and rigorous and mixed-method research and evaluation. Exhibit 6 details key personnel's role and experience: a complete list of all personnel is in Appx. J.7.

Experience and Primary Responsibilities
The ASSISTments Foundation Personnel
[redacted], *Executive Director,* has a strong background in teaching and coaching middle school math, including seven years of teaching in middle schools in various settings. [redacted] earned a Master's Degree in Math Education working with [redacted], author of one of the best-selling books on instruction. Starting in 2005, [redacted] collaborated with [redacted] on ASSISTments. She has served as senior personnel on dozens of NSF, IES, EIR, Gates, Hewlett, Schmidt, 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 the sole license to all the intellectual property of the ASSISTments platform. For this project, she will: • Oversee the project and coordinates with all partners. • Manage TAF's role in the project. • Oversee product development. • Write the annual report and disseminate the findings.
Others leaders: [redacted] - VP of Product and Programs and [redacted] - VP of Product*.
Teaching Lab Personnel
[redacted], *CEO*, has over 12 years of experience leading teams and educational systems and is an expert in transformational educational change, implementation science, and evidence-based teacher professional learning. [redacted] joined Teaching Lab from the Overdeck Family Foundation, where she created and managed the Exceptional Educators Portfolio, a set of investments focused on improving teaching and learning nationwide. [redacted] also held leadership roles at the New York City Department of Education (NYCDOE) under the Deputy Chancellor for Teaching and Learning. In her roles, she managed strategy for a 600-person division, implemented a continuous improvement strategy across 1000 schools, and oversaw policy for the city's teacher development and evaluation system serving over 70,000 teachers. Before her work at the NYCDOE, she was a high school science teacher and founder of a social justice student leadership program in Washington, DC, and Oakland, CA. For this project, she will: • Provide strategic direction and guidance for the development of in-depth and evidence-based content- and curriculum-aligned PL and coaching, • Serve as a point of escalation for the project

Other leaders: [redacted], Chief Program Officer, [redacted], **Senior Director,** Learning & Research
Evaluation Key Personnel WestEd
[redacted], *Evaluation Director*, is a Senior Research Associate with the Learning and Technology group 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 on the design and development of intelligent tutoring systems and carried out analysis of system data. She is a WWC-certified reviewer. • Direct the evaluation and the impact study. • Lead implementation fidelity measures and analysis.
Other leaders: [redacted] - Lead Methodologist*
*For more information see Appx. J.7

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

B.4 Dissemination for Further Development and Replication

For both Teaching Lab and the ASSISTments Foundation, this collaboration connects directly with our strategic priorities, and is a clear win-win. ASSISTments provides an invaluable enhancement to Teaching Lab's core model, increasing their already strong value proposition to schools and districts (last fiscal year alone they earned $4.4 million in revenue from districts and states). For the ASSISTments Foundation, this collaboration allows us to tap into Teaching Lab's market, raising the visibility of our tool in new states and districts.

This win-win means that both organizations are deeply dedicated to developing this program for scale, and both are extremely well-positioned to do this. Given the rapid growth of ASSISTments to a user base of primarily Illustrative Mathematics and EngageNY teachers (approximately 80% of our users) and middle school teachers (approximately 40%), we have a ready market for adoption of this program. Teaching Lab has partnered with over 50 school systems and six state education departments, and provided PL for 10,000 educators, reaching 1 million students since its founding in 2017. Their CEO, [redacted], is well-networked with

the key CCSSO states, curriculum providers, and nonprofits, all of whom would be eager to learn about this kind of curriculum-specific programming.

Given Teaching Lab and TAF's shared commitment to growing this program beyond the grant, we are building into the plan elements that support longer-term sustainability and scale. This includes developing recruitment materials, signing a shared MOU that outlines roles and responsibilities, and establishing a business model for joint sales, with a commitment to offering the program at a scalable price point.

C. QUALITY OF THE PROJECT DESIGN

C.1 Conceptual Framework Underlying the Proposed Research

As articulated in our logic model's key activities (see snapshot below in Exhibit 7 and detailed version in Appendix G), the conceptual framework for this model centers on three evidence-based components: (1) Teacher-use of ASSISTments to provide students just-in-time targeted support with grade-level standards, (2) Evidence-based teacher PL needed to successfully implement their HQIM and (3) Capacity-building for school coaches to more effectively support teachers, enhanced with ASSISTments-Coach. At the foundation is the use of HQIM (component "zero"), a criterion for partnership.

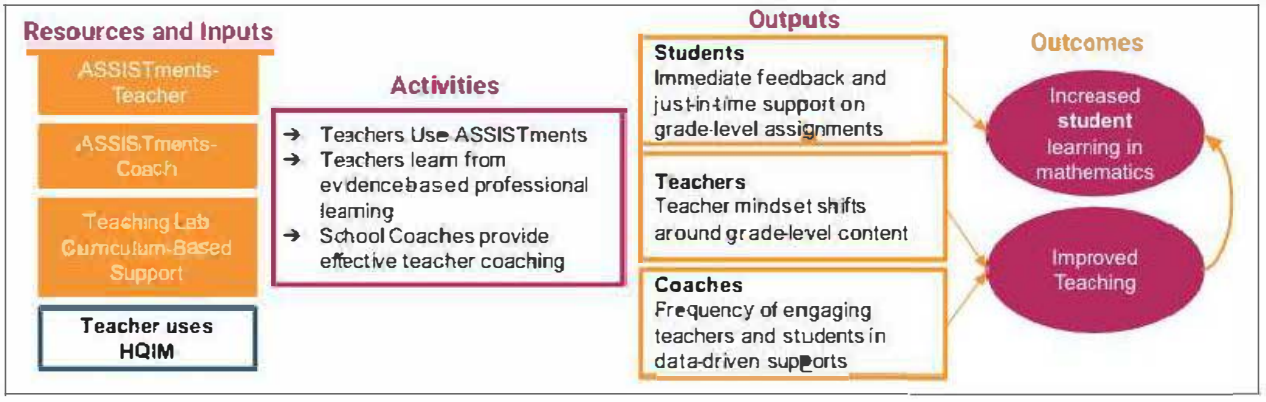

Exhibit 7: Snapshot of complete logic model found in Appx. G

The below table, Exhibit 8, highlights the key evidence in support of each component:

Conceptual Framework for ASSISTments for Algebra Readiness
0 Teacher use of HQIM
Adopting (and using) a quality grade-level curriculum is one of the most cost-effective opportunities for improving student learning at scale (Chingos & Whitehurst, 2012, Boser et al., 2015 see Figure 1). Student learning improves between 10 and 25 percentage points after schools adopt specific, high-quality curricula (Steiner, 2017).
1 Teacher-use of ASSISTments
Students who consistently completed assignments in ASSISTments performed significantly better than those whose teachers followed regular practices and did not use ASSISTments, on a standardized end-of-year math achievement assessment (effect size g = 0.22), representing an additional two-thirds of a year of gain in math learning. (Roschelle et al., 2016)
2 Evidence-Based Teacher Professional Learning
2a. *PL for teachers.* Teachers are the most important school-based factor for improving students' short and long term outcomes (Chetty et al., 2014; Goldhaber, 2021). **2b.** *Curriculum-aligned PL.* Deep curriculum and lesson study allows teachers to grow their content and pedagogical knowledge over time and spread that knowledge to their colleagues (Guskey & Yoon, 2009). The use of HQIM alone can lead to improvement in student learning (Agodini et al., 2010; Chingos & Whitehurst; 2012; Bhatt et al., 2013). The impact is magnified when HQIM is paired with PL (Jackson & Makarin, 2016). **2c.** *Teacher-led communities.* Teachers deserve to feel motivated and supported by their peers to learn and grow. In teacher-led communities, educators are more likely to buy into their own development and work collaboratively with their colleagues to improve instruction (Leana, 2011) **2d.** *Cycles of Inquiry.* Repeated cycles of inquiry afford teachers the time and space to reflect, incorporate new learning into practice, and verify changes to instruction using analysis of student work (Jensen et al., 2016)
3 Effective Teacher Coaching
Coaching teachers impacts student achievement when the support is focused on teacher-student interactions and student needs (Allen et al., 2011, 2015), content and pedagogical content knowledge (Heller et al., 2012a, 2012b); effective teacher practices and analysis of practice (Schmidt et al., 2017; Taylor et al., 2017); and the use of formative student learning data (Heller et al., 2012a; Schmidt et al., 2017).

Exhibit 8: Conceptual Framework of ASSISTments4AR

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

The four objectives of this project will guide us through implementing, iterating, evaluating and scaling ASSISTmetns4AR. The four activities in **Objective 1** will increase the scalability of using ASSISTments with HQIM. Through **Objective 2**, ASSISTments data will be thoughtfully integrated into Teaching Lab's PL and Coaching, with feedback from pilot partners on how and where it adds the most value. **Objective 3** supports our desire to assess the efficacy of ASSISTmetns4AR while **Objective 4** supports our dissemination and sustainable scale. All objectives, activities, outcomes, and measures are shown in Exhibit 9.

Activities	Outcomes	Measures
Objective 1. Enhance ASSISTments-Teacher and ASSISTments-Coach to be optimized for scale with HQIM		
1.1 Add additional auto-grading answer types to ensure rich data aligned to curricular content	500 of the existing problems that are not auto-score will be auto-scored.	Count of the problems in the curriculum that are auto-scorable.
1.2 Implement a system to support all LMS's	ASSISTments will be able to be adopted by schools with any LMS.	All participating schools will be able to access ASSISTments regardless of the schools adopted LMS
1.3 Iteratively improve new data views in ASSISTments-Coach to support HQIM and Algebra I Readiness	Three new data views that support Teaching Lab and School Coaches with coaching teachers effectively	Through surveys and interviews we will collect information from the TL and school coaches that they are satisfied with their ability to monitor Algebra I Readiness in real-time
Objective 2. Modify Teaching Lab's evidence-based PL and coaching scope and sequence to leverage ASSISTments		
2.1 Adapt existing PL courses (Bootcamp and Cycles of Inquiry) and coaching model to center the use of ASSISTments data and Algebra readiness	Evidence-based, data-driven PL and coaching model focused on Algebra Readiness	PL courses and coaching materials, including instructional design in Teaching Lab's Learning Management System
2.2 Pilot PL and coaching scope and sequence and make iterative improvements	-36 teachers engage in PL courses across 6 schools over two years -6 school coaches across 6 schools receive coaching support from Teaching Lab coaches over two years	Teaching Lab attendance and course completion systems
2.3 Assess participant feedback on the data-driven, PL and coaching	NPS > 40 80% of participants agree/strongly agree that they are satisfied with the overall quality of the PL courses and coaching	Teaching Lab participant feedback surveys
Objective 3. Rigorously evaluate the impact of ASSISTments4AR		
3.1 Recruit 70 schools and randomly assign to treatment and control	Samples of treatment and control schools with baseline equivalence in key school-level characteristics.	List of schools and their assigned condition. Report on the baseline equivalence of schools.
3.2 Implement ASSISTments4AR with Treatment Schools	Teachers will have attended the PL and school coaches would have received coaching	Attendance at the PL, Records of Coaching visits, ASSISTments assignment logs
3.3 Prepare data collection instruments and administration procedures and collect data	Improved materials and procedures for all aspects of the project.	Records from team meetings; annual report on revisions made to materials and procedures.
3.4 Measure and analyze implementation fidelity and explore the relationship between	Fidelity and compliance levels analyzed and their relationship with intervention effect analyzed; factors	ASSISTments log data, coaches logs, PD records. Implementation fidelity report from WestEd

implementation and its effects on learning.	that hinder or facilitate implementation documented.	according to the implementation compliance thresholds.
3.5 Assess the impact of ASSISTments4AR on student learning outcomes.	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.
3.6 Collect cost data and analyze cost-effectiveness.	Findings on cost and cost-effectiveness ratio.	Memo of findings regarding the cost and cost-effectiveness of ASSISTments4AR.
Objective 4. Disseminate findings and Support Scale of ASSISTmetns4AR		
4.1 Broadly disseminate learnings of how best to support students in achieving algebra readiness.	Support schools in implementing ASSISTmetns4AR improving teaching and learning of the skills needed for success in Algebra 1.	Number of presentations at conferences, blog posts, webinars, and publications.
4.2 Develop a business plan and finalize a model that is scalable to districts with high-needs students.	Develop a business plan and a pricing model for the ASSISTmetns4AR	Sell ASSISTmetns4AR to at least 3 school districts

Exhibit 9: Objectives, Activities, Outcomes, and Measures

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

With this project, and always, we are committed to advancing educational equity, raising the achievement of all students, and eliminating the racial predictability and disproportionality of which student groups occupy the highest and lowest achievement categories. Within the field we see a consistent challenge: Teacher PL does not consistently lead to educational equity. For this reason, ASSISTments4AR will not shy away from addressing equity in math head on. As examples, in math Bootcamp, teachers will create a vision for an equitable math classroom, nd examine their own biases when it comes to demonstrating high expectations for all students. During cycles of inquiry, this vision will be reinforced as teachers plan to support diverse learners in the context of specific lessons from their curriculum. ASSISTments data will keep teachers and coaches informed in real-time on student progress and needs. We know one of educators' top concerns is reaching the needs of students who have fallen behind in math due to school closures, and this program will empower them to do so (Kurtz, 2020).

D. PROJECT EVALUATION

The evaluation will consist of two components. The **formative component** aims at augmenting TAF and TLs development and continuous improvement of ASSISTments4AR and will be led by an expert in mathematics education and formative evaluation methods at TAF. It is guided by research questions around both the process and progress of the program's development. Process-related questions include: *Are the newly developed PL materials and product enhancements achieving their objectives?* Progress-related questions include: *Are project milestones being met as intended? Are the iterations of development improving quality and ease of serving students effectively?* To address these questions, the expert will interview collaborators across institutions, teacher participants and co-developers, conduct a comprehensive review of project materials and iterative development plans, and offer formative feedback during SY1 and SY2. The **summative evaluation component** will include a rigorous randomized controlled trial (RCT), led by an independent team of highly experienced researchers from WestEd. The RCT study is designed to 1) yield impact estimates of ASSISTments4AR on instructional practices and learning outcomes for a diverse population, 2) offer evidence of generalizability and cost-effectiveness, 3) examine implementation in details, and 4) provide guidance to the adoption and replication of ASSISTments4AR in other settings.

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

WestEd will conduct an independent impact study to address four research questions that are aligned with the ASSISTments4AR logic model (see Appx. G), about the program's impact on 6th- and 7th-grade students' math achievement, especially their readiness for Algebra I (RQ1), the differential effects of ASSISTments4AR on different types of students (RQ2), its impact on

coaches and teachers' pedagogical knowledge for teaching (PCK) and mathematical quality of instruction (MQI) (RQ3), and implementation fidelity and quality (RQ4).

Research Questions	Outcome	Data / Measures
(*RQ1) Outcomes.* What is the effect of ASSISTments4AR on the math achievement of 6th- and 7th-grade students compared to the business-as-usual condition?	Students' math achievement and algebra readiness	State 6th- and 7th-grade assessments scores; MDTP Algebra I Readiness Test; student and school demographics
(RQ2) Moderators. Do the effects of ASSISTments4AR vary for students with different prior achievement, and for students of different socioeconomic status, race/ethnicity, or with other policy-relevant characteristics?		
(RQ3) Impacts on pedagogical content knowledge for teaching and quality of instruction. (a) What is the impact of ASSISTments4AR on teachers' PCK? (b) What is the impact of ASSISTments4AR on teachers' MQI, particularly for algebra readiness?	Teacher's PCK and MQI	PCK assessment, Video observations, SEC survey
(RQ4) Implementation fidelity and effects of implementation on student learning. Do participating schools implement ASSISTments4AR as intended by the developer? How much usage occurs? To what extent is each feature used? What are the effects of implementation fidelity and dosage on learning? What are the factors that hinder or facilitate implementation?	Implementation fidelity and quality	Computer log data of usage; observation protocols; teacher surveys and logs; coach logs and interviews.

Exhibit 10. Summative Research Questions, Outcomes, & Measures.

Study design. The study will use a school-level, clustered randomized controlled experimental design, to meet What Works Clearinghouse standards without reservations (WWC, 2020). The RQs will be addressed with data collected from two consecutive cohorts of schools during SY 2024-25 (Cohort 1) and SY 2025-26 (Cohort 2). In each cohort, schools that will be randomly assigned to an intervention group or to a comparison group, blocked by districts. Schools in the intervention group will implement the ASSISTments4AR program with all the 6th and 7th-grade math teachers for one school year. In the comparison group, 6th and 7th-grade teachers will continue their current practices (business-as-usual). After the experiment is completed, control schools will also have access to ASSISTments4AR (delayed-treatment). Student rosters will be collected prior to the start of each school year. The team will track both overall and differential school- and student-level attrition and exclude joiners from the evaluation sample (see Appx. J.8 for Timeline of Impact Evaluation).

Sample and power. The team will recruit 70 public schools in districts that have adopted HQIM for mathematics. We will identify possible stratification variables (e.g., % of students who achieved proficiency, a % of low-income students) to improve precision and sample balance across groups and develop a plan for recruiting a representative sample to increase the generalizability of the results of the study, with a focus on involving high-need, low-performing schools and schools with lower socioeconomic status student bodies. A power analysis was conducted using PowerUp! (Doug & Maynard, 2013) for random assignment at the school level. Assuming an average of two math teachers, 50 students per teacher, would be recruited at each school in 6th- and 7th-grade, the proposed study has sufficient power to detect effects of 0.179 standard deviations on student achievement (see Appx. J.9 for justification).

Data Analysis Plan. To estimate the impact of ASSISTments4AR on student achievements, we will conduct an intent-to-treat (ITT) analysis using a three-level model with students tested in classrooms and in schools..Continuous scaled scores from different states will be transformed into z-score units for comparability. Analyses will test the overall impact of the intervention for 6^{th}- and 7^{th}-grade students, as well as the interventions' differential effects on students and schools with different characteristics. To improve the precision of the impact estimate, each model will adjust for blocking (by cohort and district), a baseline measure, other characteristics at the student, class and school levels (see the data analysis plan for all RQs in Appx. J.10).

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

The proposed study is designed to generate insightful guidance for successful replicating and scaling of ASSISTments4ARbroadly in four aspects. First, we will **strategically sample** and recruit schools representing diverse settings and conduct interviews with school administrators to

document the context of implementation and supportive structures and environment, such as district priorities, adoption status of technology and HQIM, school support and cultures for data use, peer collaboration in teacher-led community.

Secondly, the evaluation included **differential impact analysis** (RQ2) to assess to what extent the impact of ASSISTments4AR is moderated by the characteristics of students (e.g., ethnicity, eligibility for free or reduced-price lunch) or schools (e.g., grade levels served, locale, demographic composition) and to identify for which settings or populations the intervention is particularly effective. The prior efficacy study (Roschelle et al., 2006) found that ASSISTments helped close achievement gaps. WestEd will further investigate this during the impact study.

Thirdly, the study builds on **analysis of implementation data from multiple sources** (RQ4). The essential features that should be followed by future efforts to replicate or extend this intervention include 1) teachers' participation in coaching and PL; 2) consistency in using HQIM and assigning grade-level problems; 3) reviewing data in reports and adjust instruction accordingly; 4) coaches reviewing data on teachers' implementation of HQIM and students' grade-level progress, and 5) coaches engaging teachers in data-driven supports. Accordingly, we will collect and examine implementation data on all five aspects (see Appx. J.11 for details and Appx. J.12 for data collection timeline). We will focus on the **fidelity of the scaling-up approaches** and identify facilitators and barriers to implementation.

Lastly, policymakers and educational administrators require information on the cost of the resources needed to adopt, implement with fidelity, and sustain the interventions. To provide information about whether ASSISTments4AR is more cost effective than business-as-usual practices, WestEd will conduct a **cost analysis** and calculate a **cost-effectiveness** ratio. Costs will

be gathered systematically using the "ingredients method" as described in Levin and McEwan (2001). Once all the ingredients are identified, we will use *CostOut*® the CBCSE Cost Tool Kit (Hollands et al., 2015-2019) to facilitate the estimation of costs and cost-effectiveness (see Appx. J.13 for details of cost analysis plan).

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

Student Outcomes. The **primary student learning outcome measure** will include 6th- or 7th-grade statewide mathematics assessment scores, using the relevant accountability measures of math achievement within each state. WestEd will secure agreements with participating districts to obtain the state math test scores in 5th (as covariate), 6th (as covariate and outcome) and 7th (as outcome) grades, as well as demographics. Additionally, we propose to use the Algebra Readiness Test developed by the Mathematics Diagnostic Testing Project (MDTP) as a **supplemental measure** to gauge students' readiness for foundational topics necessary for success in an Algebra I course at the end of the study. The MDTP test has been reviewed for reliability (CSU/UC, 1995) and validity (Anthony, 2005; CSU/UC, 2012; Huang et al., 2014; Snipes et al., 2015; KR-20=0.87) and is highly predictive of success in Algebra I (Huang et al., 2014). The test will be administered online to all students towards the end of the study year.

Teacher Outcomes. We will use a math knowledge for teaching assessment to measure teachers' PCK (e.g., see Ball et al., 2008) (RQ3a), which emphasizes the application of conceptual understanding in teaching contexts, as well as the ability to make sense of student work and thinking (Ball et al., 2008; Hill et al., 2005). The assessment will include items that have been developed and validated by the Learning Mathematics for Teaching (LMT) Project

(Hill et al., 2004). To minimize the burden, we will carefully select items that are aligned with the standards related to algebra readiness, aiming for the test-taking time to be 45 minutes.

We will measure teachers' quality of instruction (RQ3b) through (a) video observations and (b) the Survey Of Instructional Practices (Teacher Survey for Grades K–12 Mathematics), a part of the *Surveys of the Enacted Curriculum* (SEC) data tools (WCER, 2012). A sample of 20 intervention and 20 comparison teachers will be asked to video-record two consecutive classes in September (baseline) and two more consecutive class sessions (outcome) in March. Drawing from technology, teacher training protocols, and incentives used in prior studies that resulted in high-quality video recordings, WestEd will support teachers in video-recording these sessions and securely uploading them for analysis. The purpose of these video recordings is to understand teachers' use of the recommended practices emphasized in the Teaching Lab learning sequence (e.g., citing relevant evidence, productive struggle, student talk). Trained researchers at WestEd will use the Mathematical Quality of Instruction (MQI) observational rubric (Hill et al., 2012) to code the video recordings. Summary ratings will be assigned for overall MQI and each of the five domains (common core-aligned student practices, working with students and mathematics, richness of mathematics, errors and impression, and classroom work is connected to mathematics) for analyses of the subsample of instructional practices.

The SEC survey has demonstrated itself to be a valid and reliable instrument for measuring key aspects of instructional practices (Porter et al. 1993). The overall reliability coefficient for the instrument ranges from 0.72 to 0.87. All teachers will take the survey online in summer (as baseline) and again at the end of the year. To measure teacher background characteristics and service contrast, WestEd will adapt the survey to add questions such as years of teaching math,

certification type(s), and math background and the types, focus, and duration of other PL teachers participated in during the study. In addition, both intervention and comparison teachers will be asked to fill out brief instructional logs in four weeks on lesson topics covered, grade-level assignments made, and data and/or pedagogical approaches used.

D.4. Components, Outcomes, and Measurable Threshold for Implementation.

The design of the evaluation is informed by the key components and major teacher and student learning outcomes as illustrated in the logic model. The key components of ASSISTments4AR include (a) bootcamps, coaching, and teacher PL community to support **teachers** in learning about ASSISTments4AR and addressing their questions and concerns in a collaborative manner, (b) providing pre-built HQIM and data reports to make it easy for teachers to select grade-level practices and to make informed instructional adjustments, (c) supporting school **coaches** in oversight of teachers' implementation of HQIM.

The ASSISTments system records of student, teacher, and coach actions will serve as a primary source of data in assessing the extent to which ASSISTments4AR is implemented with fidelity. WestEd will collect data on attendance of bootcamps, PL, and coaching sessions. School coaches will complete a short 5-min online log each time they deliver coaching to teachers. The log will ask about the duration, content covered, practice assignments made in ASSISTments4AR that were referenced. We will analyze the logs together with the system record data to determine the dosage. In addition, we plan to interview 20-percent of school coaches, asking them to reflect on teachers' interaction with peers in the community, supports or barriers that facilitate or hinder engagement of the coaching, how they integrate data from the ASSISTments4AR platform into their coaching practices.

WestEd will measure the fidelity of implementation, using the five-step procedure proposed by Nelson et al., (2012). Referring to the framework by Weiss, Bloom, & Brock (2014), we will develop metrics and set associated thresholds of fidelity for each metric. There will be metrics for three types of **indicators of implementation compliance**: one on the extent to which teachers and school coaches complete PL and participate in community activities, one based on intensity and continuity of delivery of data-driven coaching sessions, and one on the extent to which teachers make grade-level assignments and use the ASSISTments reports and school coaches use teacher implementation oversight tools. 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: Teachers are expected to participate in the three-day virtual bootcamps and three cycles of inquiry accumulating at least 71 hours of PL; school coaches are expected to review teacher implementation data at least once per week and deliver four coaching sessions per month, in addition to in-person modeling and/or co-teaching once a week per teacher; teachers are expected to assign grade-level HQIM problems two to three times per week and review reports for at least half of the assignments. During the study year, WestEd will calculate the metrics at the individual level monthly and share the results with the team to inform continuous improvement.

References

Agodini, R, Harris, B., Thomas, M., Murphy, R., & Gallagher, L, (2010) Achievement Effects of Four Early Elementary School Math Curricula: Findings for First and Second Graders (NCEE 2011-4001), Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education (2010). https://files.eric.ed.gov/fulltext/ED512551.pdf

Allen, J.aP., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An Interaction-Based Approach to Enhancing Secondary School Instruction and Student Achievement. *Science*, 333(6045), 1034–1037. Retrieved from:https://eric.ed.gov/?id=ED556046

Allen, J.aP., Hafen, C. A., Gregory, A. C., Mikami, A. Y., & Pianta, R. (2015). Enhancing secondary school instruction and student achievement: Replication and extension of the My Teaching Partner-Secondary intervention. *Journal of Research on Educational Effectiveness, 8(4)*, 475–489. Retrieved from:https://eric.ed.gov/?id=EJ1078792

Anthony, M. (2005). Consequential-related validity evidence for MDTP tests. Data submitted in response to California Community Colleges Assessment Standards for renewal of placement test instruments. Retrieved from Mathematics Diagnostic Testing Project website: http://mdtp.ucsd.edu

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.https://www.nber.org/papers/w27555

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher E ducation, 59(5)*, 389–407. doi:10.1177/0022487108324554

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* , 1165–1188.

Bhatt, R., Koedel, C., & Lehmann, D.(2013) Is Curriculum Quality Uniform? Evidence from Florida, *Economics of Education Review 34* (June 2013): 107–121.

Blazar, D., Heller, B., Kane, T., Polikoff, M., Staiger, D., Carrell, S.,...& Kurlaender, M. (2019). Learning by the Book: Comparing math achievement growth by textbook in six Common Core states. Research Report. Cambridge, MA: Center for Education Policy Research, Harvard University. https://cepr.harvard.edu/files/cepr/files/cepr-curriculum-report_learning-by-the-book.pdf

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Evaluation Review, 8* (2), 225–246.

Bloom, H. S., Richburg-Hayes, L., & Black, A. R. (2007). Using covariates to improve precision for studies that randomize schools to evaluate educational interventions.*Educational Evaluation and Policy Analysis, 29* (1), 30-59.

Boser, U., Chingos, M. & Straus, C. (2015, October). The Hidden Value of Curriculum Reform Do States and Districts Receive the Most Bang for Their Curriculum Buck? Center for American Progress. https://cdn.americanprogress.org/wp-content/uploads/2015/10/06111518/CurriculumMatters-report.pdf

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

Carpenter, J. (2021) Early Texas standardized test scores show big declines in math, steady

reading rates in high school. June 15, 2021 6:03 p.m. *Houston Chronicle*. Retrieved from https://www.houstonchronicle.com/news/houston-texas/education/article/Early-Texas-standardized-test-scores-show-big-16249920.php

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

CDT (2021) Key Views Toward EdTech, School Data and Student Privacy. Retrieved from https://cdt.org/wp-content/uploads/2021/11/CDT-Teacher-Parent-Student-Survey-Fall-2021-Final.pdf

Chetty, R., Friedman, J.N., & Rockoff, J.E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), pp. 2633-2679. https://www.nber.org/papers/w19424

Chingos, M. & Whitehurst, G. (2012, April 10). Choosing Blindly: Instructional Materials, Teacher Effectiveness, and the Common Core. The Brookings Institute. https://www.brookings.edu/research/choosing-blindly-instructional-materials-teacher-effectiveness-and-the-common-core/

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.

CSU/UC: California State University/University of California Mathematics Diagnostic Testing Project (CSU/UC) (1995). Mathematics Diagnostic Testing Project user manual. San Diego, CA: University of California. Retrieved February 27, 2015, from http://mdtp.ucsd.edu/

CSU/UC: California State University/University of California Mathematics Diagnostic Testing Project (CSU/UC) (2012). *California Community Colleges Validation Study Without College Names For Posting on Web Site*. Retrieved March 20, 2020 from https://www.venturacollege.edu/sites/default/files/files/college-information/accreditation/2016-documentation/standard-iic/ii.c.7.7_approval_of_mdtp_validity_studies_2012.pdf

Doan, S, Grant, D, Henry, D., Kaufman, J., Lawrence, R., Tuma, A., Setodji,C., Stelitano, L., Woo, A., & Young, C., (2021) American Instructional Resources Surveys: 2021 Technical Documentation and Survey Results. Santa Monica, CA: RAND Corporation, 2021. https://www.rand.org/pubs/research_reports/RRA134-10.html

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

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2021). COVID-19 and education: The lingering effects of unfinished learning. McKinsey & Company. Available at: https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-education-the-lingering-effects-of-unfinished-learning

EdGlossary. (2013). Opportunity Gap.The Glossary of Education Reform. Created by the Great Schools Partnership. https://www.edglossary.org/opportunity-gap/

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

EdReports (2020) State of the Market 2020: The Use of Aligned Materials. Retrieved from https://cdn.edreports.org/media/2021/03/EdReports-2020-State-of-the-Market-Use-of-Aligned-Materials_FINAL_2-1.pdf

Education Trust, The (2018) Checking In: Are Math Assignments Measuring Up? A report by Dysarz, K. https://edtrust.org/the-equity-line/power-classroom-assignments/ https://edtrust.org/wp-content/uploads/2014/09/CheckingIn_MATH-ANALYSIS_FINAL_5.pdf

Ejiwale, J. (2013). Barriers to successful implementation of STEM education. *Journal of Education and Learning. Vol.7* (2) pp. 63-74. https://doi.org/10.11591/edulearn.v7i2.220

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

Evidence for ESSA. (2020). ASSISTments shown to be one of only 3 interventions with strong evidence for middle school math. https://www.evidenceforessa.org/programs/math?field_evidence_rating=%5B4%5D&field_grades=%5B12%5D.

Evidence for ESSA. (2022a). There are 35 reading professional development programs with strong evidence. https://www.evidenceforessa.org/programs/reading?field_evidence_rating=%5B4%5D&field_features=%5B30%5D

Evidence for ESSA. (2022b). There are 17 math professional development programs with strong evidence https://www.evidenceforessa.org/programs/math?field_evidence_rating=%5B4%5D&field_features=%5B30%5D

Evidentally. (2022). An Evidentally Demographic Profile: Teaching Lab. March 2022. The company was funded to do a report on Teaching Lab's user base. Retrieved from https://drive.google.com/file/d/12T75EFahPOk1GBTutRdHBV12353BTXJC/view?usp=sharing

Gates Foundation (2022) Balance the Equation: A Grand Challenge for Algebra 1. See page 2. Retrieved from https://usprogram.gatesfoundation.org/-/media/usp/usp-resources/k12-resources/balancetheequationgrandchallengealgebraonefactsheet.pdf

Goldhaber, D. (2021). In Schools, Teacher Quality Matters Most. Education Next, 21(3). Available at: https://www.educationnext.org/in-schools-teacher-quality-matters-most-coleman/

Guskey, T. R., & Yoon, K. S. (2009). What works in professional development. Phi delta kappan, 90 (7), 495-500.

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlations for planning group-randomized experiments in education. *Educational Evaluation and Policy Analysis*, 29 , 60–87.

Heller, J. I., Daehler, K. R., Wong, N., Shinohara, M., & Miratrix, L. W. (2012a). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. Journal of Research in Science Teaching, 49(3), 333–362. Retrieved from: https://eric.ed.gov/?id=EJ989348

Heller, J. H. (2012b). Effects of Making Sense of SCIENCETM professional development on the achievement of middle school students, including English language learners (NCEE 2012-4002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from: https://eric.ed.gov/?id=ED530414

Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M.aA., Beisiegel, M., Humez, A., Litke, E., & Lynch, K. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment, 17*(2–3), 88–106.

Hill, H.C., Lynch, K., Gonzalez, K.E., & Pollard C. (2020). Professional development that improves STEM outcomes. *Phi Delta Kappan, 101*(5), 50-56. https://doi.org/10.1177/0031721720903829

Hill, H. C., Rowan, B., & Loewenberg, D. (2005). Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal 42*, no. 2: 371-406.

Hill, H.C., Schilling, S.G., & Ball, D.L. (2004) Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal 105*, 11-30.)

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

Huang, C.-W., Snipes, J., & Finkelstein, N. (2014). Using assessment data to guide math course placement of California middle school students (REL 2014–040). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from http://ies.ed.gov/ncee/edlabs. http://eric.ed.gov/?id=ED546776

Jackson, C. K., & Makarin, A. (2016). Can Online Off-The-Shelf Lessons Improve Student Outcomes? Evidence from A Field Experiment. National Bureau of Economic Research, (22398). https://www.nber.org/papers/w22398

Jensen, B., Sonnemann, J., et al. (2016). Beyond PD: Teacher Professional Learning in High-Performing Systems. Washington, DC: National Center on Education and the Economy. https://www.ncee.org/wp-content/uploads/2015/08/BeyondPDWeb.pdf

Kelly, K., Heffernan, N., Heffernan, C., Goldman, S., Pellegrino, J., & Soffer-Goldstein, D. (2014). Improving student learning in math through web-based homework review. In Liljedahl, P., Nicol, C., Oesterle, S., & Allan, D. (Eds.). (2014). *Proceedings of the Joint Meeting of PME 38 and PME-NA 36 (Vol. 3)*. Vancouver, Canada: PME. pp. 417-424. https://files.eric.ed.gov/fulltext/ED599834.pdf

Klute, M. (2021) Ready or Not: Using Existing Data as Clues for Algebra I Readiness. A report from the IES REL of Appalachia. https://ies.ed.gov/ncee/edlabs/regions/appalachia/blogs/blog43_using-data-for-algebra-I-readinesss.asp

Koedinger, K., McLaughlin, E. & Heffernan, N. (2010). A Quasi-Experimental Evaluation of an On-line Formative Assessment and Tutoring System. *Journal of Educational Computing Research. Baywood Publishing. 4, 489-510.* *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.210.5186*

*Kraft, M.A., Blaz*ar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. Review of Educational Research, 88(4), 547–588. https://doi.org/10.3102/0034654318759268

Kraft, M.aA. & 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

Kraft M., & Papay J. (2014) Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Effectiveness and Policy Analysis.* 2014;36 (4) :476-500. Retrieved from https://scholar.harvard.edu/files/mkraft/files/kraft_papay_-_prof_env_teacher_development_eepa_full.pdf

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

Kurtz, H.(2020) Nation Survey Tracks Impact of Coronavirus on Schools: 10 Key Findings. *EducationWeek.* April 10th, 2020. Retrieved from https://www.edweek.org/teaching-learning/national-survey-tracks-impact-of-coronavirus-on-schools-10-key-findings/2020/04

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

Leana, C. (2011). The Missing Link in School Reform. Stanford Social Innovation Review, Fall 2011, 30-35.

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.

Lynch, K., Hill, H.C., Gonzalez, K.E, & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. Educational Evaluation and Policy Analysis, 41(3), 260–293. https://doi.org/10.3102/0162373719849044

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

McAleavy, T., Hall-Chen, A., Horrocks, S., & Riggall, A. (2018). Technology-Supported Professional Development for Teachers: Lessons from Developing Countries. Education Development Trust. https://www.educationdevelopmenttrust.com/our-research-and-insights/research/technology-supported-professional-development-for-

Mendicino, M., Razzaq, L., & Heffernan, N. T. (2009). A comparison of traditional homework to computer supported homework. *Journal of Research on Technology in Education (JRTE). 41(3)*, 331–358. https://files.eric.ed.gov/fulltext/EJ835243.pdf

Miller, A. F. & Partelow, L. (2019, September 20). Successful implementation of high-quality instructional materials: 5 case studies. Center for American Progress.

https://www.americanprogress.org/article/successful-implementation-high-quality-instructional-materials/

Miller, G. I, Zheng, Y., Means, B. & Van Brunt, J. (2013). Next Generation Leaming Challenges Wave II Final Evaluation Report. SRI Project Report # P20778. Paid for by the Gates Foundation. Retrieved on May 20, 2014, from https://drive.google.com/file/d/0B2X0QD6q79ZJU19Kd2JuVTN0VWhTYVRhX254QV85Njdqc1Vj/edit?usp=sharing

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.

Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, J. L., & Schneider, S. A.(1993). Reform up close: An analysis of high school mathematics and science classroomsa Madison, WI: Wisconsin Center for Educational Research. https://eric.ed.gov/?id=ED364429

[redacted]

Schmidt, R. A., Wang, H., Cassidy, L., & Laguarda, K. (2017). A comprehensive model of teacher induction: Implementation and impact on teachers and students evaluation of the New Teacher Center's i3 validation grant, final report. Menlo Park, CA: SRI International.

Schochet, P. Z. (2008). Statistical power for random assignment evaluations of education programs. Journal of Educational and Behavioral Statistics, 33 (1), 62–87.

Singh, R., Saleem, M., Pradhan, P., Heffernan, C., Heffernan, N., Razzaq, L. Dailey, M. O'Connor, C. & Mulchay, C. (2011). Feedback during Web-Based Homework: The Role of Hints. In Biswas et al. (Eds) *Proceedings of the Artificial Intelligence in Education Conference 2011*, 328–336. https://link.springer.com/chapter/10.1007/978-3-642-21869-9_43

Star, J. R., Caronongan, P., Foegen, A., Furgeson, J., Keating, B., Larson, M. R., Lyskawa, J., McCallum, W. G., Porath, J., & Zbiek, R. M. (2015). Teaching strategies for improving algebra knowledge in middle and high school students (NCEE 2014-4333). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://whatworks.ed.gov.

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/
Steiner, D. (2017, March). Curriculum research: What we know and where we need to go. Standard Works. https://standardswork.org/wp-content/uploads/2017/03/sw-curriculum-research-report-fnl.pdf
Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Sage Publications, Inc.
Taylor, J., Roth, K., Wilson, Stuhlsatz, M., & Tipton, E. (2017). The Effect of an Analysis-of-Practice, Videocase-Based, Teacher Professional Development Program on Elementary Students' Science Achievement. Journal of Research on Educational Effectiveness, 10(2), 241-271. Retrieved from: https://eric.ed.gov/?id=EJ1135795
Teaching Lab. (2020). *Impact Report.* Retrieved from: https://static1.squarespace.com/static/5ef2087220c7256a8632676b/t/6234b00f4b73635ba88d0053/1647620120707/Teaching+Lab+-+2020+Impact+Report.pdf
TNTP. (2018). The Opportunity Myth: What Students Can Show Us About How School Is Letting Them Down—and How to Fix It. The New Teachers Project. https://tntp.org/assets/documents/TNTP_The-Opportunity-Myth_Web.pdf
TNTP (2021) Accelerate, Don't Remediate: New Evidence from Elementary Math Classrooms. The New Teacher Project. May, 2022. https://tntp.org/assets/documents/TNTP_Accelerate_Dont_Remediate_FINAL.pdf
United State Department of Education (2018, Nov.) A Leak in the STEM Pipeline: Taking Algebra Early. Retrieved from https://www2.ed.gov/datastory/stem/algebra/index.html#data-story-title
Walker, L. (2017) A Case for High-Quality Instructional Materials. Achieve The Core. Student Achievement Partners Inc.. Retrieved from https://achievethecore.org/peersandpedagogy/a-case-for-high-quality-instructional-materials/
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). (2019). WWC Review of Roschelle et al (2016) on ASSISTments and finds it meets "without" reservation and qualifies for Tier 1. https://ies.ed.gov/ncee/wwc/Study/86375
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
White, I. R., & Thompson, S. G. (2005). Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine*, *24*(7), 993–1007. https://doi.org/10.1002/sim.1981
Wiener, R. & Pimentel, S. (2017, April). Practice what you teach: Connecting curriculum & professional learning in schools. The Aspen Institute: Education & Society Program. https://assets.aspeninstitute.org/content/uploads/2017/04/Practice-What-You-Teach.pdf
Wisconsin Center for Education Research (WCER) (2012). SURVEYS OF ENACTED CURRICULUM. Survey Of Instructional Practices. Teacher Survey. Grades K-12. Mathematics. Retrieved August 2018 from

https://secure.wceruw.org/seconline/Reference/SECmathSurvey2012.pdf

U.S. Department of Education. (2018, Nov). A Leak in the STEM Pipeline: Taking Algebra Early. https://www2.ed.gov/datastory/stem/algebra/index.html#data-story-title