THE EARLY MATH INITIATIVE: SCALING AN INNOVATIVE INTERVENTION
TO PREPARE CHILDREN FOR ELEMENTARY SCHOOL MATHEMATICS

PROJECT NARRATIVE

TABLE OF CONTENTS

Significance

All citizens need a broad range of basic mathematical skills and understanding to make informed decisions in their jobs, households, and communities. Careers in the 21st century require an increasing level of proficiency in mathematics (Glenn Commission, 2000; U.S. Dept. of Labor, Bureau of Labor Statistics, 2000). American students' math performance, however, is well below what is desired and needed for success in the modern world (National Mathematics Advisory Panel, 2008; Provasnik et al., 2012). A series of national and international assessments of mathematics achievement has revealed an overall level of proficiency in American students well below their peers in several other countries and below what is desired and needed (Kilpatrick, Swafford, & Findell, 2001; Mullis et al., 1997; Mullis et al., 2000; NCES, 2016). This cross-national gap in American students' math achievement is the result, at least in part, of the "mile-wide and inch-deep" mathematics curriculum used in American schools. In contrast, several nations with higher achievement use mathematics curriculum that is narrower and deeper (i.e., mathematically central or foundational) at each grade level (e.g., NCES, 2008, 2016). In response to these concerns, the adoption of world-class math standards by American schools has been recommended (National Council of Teachers of Mathematics, 2008; National Mathematics Advisory Panel, 2008). Most states or school districts within states have responded to these concerns by revising their existing standards (e.g., Arizona Department of Education, 2016) or by adopting or modifying the Common Core State Standards for Mathematics (CCSSM; National Governors Association Center for Best Practices, 2010). These reforms are intended to help American schools close the national gap in mathematics achievement.

A major challenge educators face in implementing these standards, however, arises immediately in grade K where these standards begin. Elementary schools are expected to align instruction with the higher learning expectations built into the new standards, but teachers still face the unsolved problem that many children enter school unprepared for the math curriculum in grade K (National Research Council, 2009). Gaps in early math knowledge are especially

pronounced for children from low-income backgrounds (National Center for Education Statistics, 2011; Reardon, 2011, with socioeconomic- (SES-) related gaps being larger than gaps related to other demographic characteristics, including race/ethnicity (Duncan & Magnuson, 2011). Over the past three decades the black-white achievement gap for math and literacy has decreased but the income-related gap has grown (Reardon, 2013), Thus, there is an SES gap within the national gap for mathematics achievement. Low-SES children comprise a large segment of our student population, with approximately 1 in 4 American children under age 6 living in poor families (Jiang, Granja, & Koball, 2017). At the end of pre-kindergarten (pre-K), low-SES children are almost one year behind their middle- class peers in math knowledge, a large difference at such an early age (Starkey & Klein, 2008). Furthermore, the math gap in kindergarten between children in the lowest and highest quintiles of SES is estimated at 1.3 standard deviations (ECLS-K math measure; Duncan & Magnuson, 2011; Duncan & Sojourner, 2013). Left unaddressed, an abundance of research shows that this early math gap will persist and increase over time (Anunola et al., 2004; Entwisle & Alexander, 1989, 1990; Jordan, Huttenlocher, & Levine, 1992; Morgan, Farkas, & Wu, 2009; Rampey, Dion, & Donahue, 2009; Rathbun & West, 2004). Large student gaps make implementation of world class mathematics standards an ongoing and expensive challenge at grade K and beyond, affecting the quality and/or the amount of time and costs associated with instructing students who lag far behind their peers, including increased child needs assessments, tier 2 or tier 3 instruction, and grade retention.

The proposed project is designed to improve low-income children's readiness for elementary school mathematics through use of an innovative intervention, *Pre-K Mathematics*. The project will also build on children's mathematical growth in pre-K by strengthening the mathematics instruction children receive in K. Our *Pre-K Mathematics* intervention is listed on the What Works Clearinghouse (WWC) as effective. It has been evaluated in gold-standard RCTs, all finding significant, positive evidence of effectiveness, with effect sizes (ES) for child

outcomes meeting the WWC's standard for substantively important effects (ES≥0.25; What Works Clearinghouse, 2014). In our recently completed i3 validation project, the SES gap in mathematical knowledge was closed. Children who received the intervention, followed by standards-based math instruction in K, improved to a level equal to the national average at the end of pre-K and at the end of K (see **Absolute Priority 1**). Also, our professional development has a positive impact on pre-K teachers' math instructional practices, with treatment teachers increasing the amount of time spent overall supporting math (ES= 1.26) and using a set of effective math practices (ES=1.70). This intervention will help the nation's public preschools and schools raise mathematics achievement to the level needed to implement world class mathematics standards.

Approach to the Priorities

The EIR grant program and its predecessor (i3) has a multi-tier goal structure, such that successful intervention projects can advance from an early (development) phase to a mid-(validation) phase, or from a mid- (validation) phase to an expansion (scale-up) phase. We propose to advance our work on a fully developed intervention, which has been successfully validated at mid phase, to conduct an expansion phase project. The proposed project will examine, at the national level, the effectiveness, sustainability and scalability of the pre-k intervention. It will also examine the degree to which child gains in pre-k can be maintained and built upon when their K teachers use a strong mathematics instructional system (described below). This project addresses two absolute priorities and two invitational priorities.

Absolute priority 1. Strong Evidence. The proposed project is exceptional, given the extent of prior evidence amassed for our innovative intervention and the plan to further evaluate the intervention at the next, expansion phase of EIR, in a gold-standard experimental design using a sampling plan developed using both purposive and random sampling methods. The *Pre-K Mathematics* intervention has been evaluated in five gold-standard RCTs (Thomas, Cook, Klein, Starkey, & DeFlorio, in press), four of which are described in the **Effectiveness Form**).

All found significant, positive evidence of effectiveness, with effect sizes for child outcomes ranging from 0.30-0.83 (Hedges g).

It is noteworthy that the intervention has been found to be effective in studies conducted in several states, and most recently, when implemented at the state level of scale in California. In that study, treatment children, who received the *Pre-K Mathematics* intervention in pre-K followed by K math instruction that was aligned to rigorous state standards, were assessed on the ECLS-K: 2011 Mathematics measure, which was developed by the U. S. Department of Education and normed on a national sample. The mean normed score of treatment children at the end of K was 50.57, which slightly exceeds the mean national norm, 50.00. Thus, mathematical knowledge in these children was improved to a level equivalent to the national average, which is an important finding.

Absolute priority 3. Field-Initiated Interventions - Promoting STEM Education. Our intervention is a dedicated pre-K mathematics intervention. High mathematical competence is needed for entry into college, graduate school, and many professional programs, and mathematical knowledge is used in STEM occupations (e.g., science and engineering) and in STEM-related occupations (e.g., accounting and finance).

Invitational priority 2. Early Learning and Cognitive Development. Under this priority, the EIR grant program has invited early childhood projects. Note that EIR's responses to frequently asked questions (FAQ) includes the following: "Early learning programs can be considered to serve children from birth to age three, in preschools, and in kindergarten through third grade." Also, "Be especially mindful if you are proposing to serve Pre-K early learners or to address postsecondary preparation. It's okay, for example, to focus on the transition between pre-K and elementary education, or to focus on the transition from high school to college as long as you can indicate how the project impacts K-12 students at some point during the project, thus aligning with the ESEA." Consistent with these program requirements, mathematical cognition is a domain of cognitive development (e.g., Ginsburg, Klein, & Starkey, 1998). This project will provide an intervention in pre-K and follow-up support in K – a K mathematics instructional

system, and assessment in pre-K, K, and 1. Providing high-need children with effective mathematics support from pre-K through K is exceptional and builds on our i3 validation project.

Invitational priority 1. Personalized Learning. The proposed project utilizes personalized learning. The *Pre-K Mathematics* intervention uses flexible groupings, progress monitoring, additional curriculum dosage when needed, and dyadic parent-child home math activities, which are described below. Follow-up support in K will enable K teachers to conduct math screening to identify at-risk children, and to provide personalized math instruction during regularly scheduled math lessons.

The Innovative Intervention and Strategy for Scaling It Up

'Theoretical Framework and Model of Causation

Educational interventions can fail either because they have an inadequate theoretical foundation or because they are implemented poorly. For that reason, we will first address theoretical considerations guiding our intervention. Later, we will detail our training, implementation procedures, implementation data collection, and quality control system.

The early development of mathematical cognition. The primary conceptual foundations of children's early mathematical knowledge include the cognitive domains of number and space. These domains are partly structured during infancy (Geary, 1994; Starkey & Cooper, 1980; Starkey, Spelke, & Gelman, 1990). The constraints imposed by this partial structuring enable children to attend to and assimilate mathematically relevant inputs from the environment (Gelman & Williams, 1998). Children first develop informal mathematical knowledge – knowledge that depends upon the presence or mental representation of sets of concrete objects (Piaget, 1952). This knowledge develops considerably during the first three years of life (Baroody, 2004; Starkey, 1992), and children often have several mathematical competencies when they enroll in preschool at age three (Bisanz, et al., 2005; Ginsburg, Klein, & Starkey, 1998). The extent of children's knowledge at the beginning of preschool, however, depends on their developmental history, especially in the first three years of life.

Informal mathematical knowledge continues to develop during the preschool years and beyond, and research has identified some of the developmental sequences that occur (e.g., Baroody, 2004; Ginsburg, et al., 1998; Sophian, 1996). The significance of informal mathematical knowledge is that it serves as a conceptual foundation for the acquisition of formal mathematical knowledge – the ability to use abstract numerical notation such as the written numerals (1, 2, 3, etc.) and arithmetic operation signs (+, -, etc.). The transition to formal mathematical knowledge begins at age 4-6 years, depending on children's culture and socioeconomic status (Starkey & Klein, 2003). Children with more extensive informal mathematical knowledge in preschool tend to acquire formal mathematical knowledge earlier and more extensively in early elementary school (e.g., Aunio & Niemivirta, 2010; Purpura, Baroody, & Lonigan, 2013). Low-SES preschool and kindergarten (K) children, relative to their middle-class peers, possess less extensive informal mathematical knowledge (Jordan, Huttenlocher, & Levine, l992, 1994; Starkey & Klein, 1992, 2008; Starkey et al., 2004; West, Denton, & Germino-Hausken, 2000). These early differences are conceptually broad—encompassing informal knowledge of number, arithmetic, patterns, space/geometry, and measurement (Starkey et al., 2004)—as well as persistent, and widen over time (Morgan, Farkas, & Wu, 2011; Jordan, Glutting, & Ramineni, 2010; Rathbun & West, 2004).

Theory of learning environments embodied in the intervention. Children's early mathematical knowledge is constrained by a developmental niche (Super & Harkness, 1996), comprised primarily of the home and school learning environments. The mathematical support provided in children's niches partly determine the foundation of informal mathematical knowledge they develop. Mathematical knowledge develops primarily in, or as a consequence of, social activity settings (Lantolf & Poehner, 2014; Lantolf, Thorne, & Poehner, 2015; Vygotsky, 1978) – specifically, settings in which children actively participating in concrete math activities with teachers or parents who scaffold their learning. Therefore, math instruction is most effective when teachers possess (a) knowledge of math content, (b) knowledge of milestones in early mathematical development, and (c) knowledge of how curriculum activities

can be sequenced to coincide with known sequences in early mathematical development.

Model of causation. Measures aligned with components of this model are given in the **Project Design** and **Project Evaluation**. The active ingredients in the intervention are modeled as the mathematics content from the *Pre-K Mathematics* curricular intervention, which is aligned with world-class mathematics standards, including CCSSM clusters for kindergarten. Intensive and frequent PD will be the primary means through which teachers become able to deliver the curriculum with both fidelity and understanding (cf., Shulman, 2000). The in-depth, domain-specific PD support that teachers will receive – math focused workshops, on-site training aligned with the mathematics curricula, and continuous improvement assistance - will ensure that they (a) learn the essential mathematics content comprising the scope and sequence of the math curricula, (b) learn to implement with fidelity, including delivery of recommended curriculum dosage, (c) are able to support student engagement and learning of mathematics through explicit, teacher-guided instruction, and (d) develop and use pedagogical content knowledge.

We expect that teaching essential mathematics content through effective delivery techniques in school classroom settings will change the nature of teaching and learning opportunities for students. Thus, we predict that the mathematics experiences of students will be different in treatment classrooms than in control classrooms, and we expect that the frequency and topography of the instructional interactions between teachers and children will be different in treatment classrooms relative to control classrooms. At the pre-K level, for example, treatment children will spend more time than control children engaged in developmentally sensitive, teacher-scaffolded small-group mathematics activities. At the K level, treatment children will spend more time than control children engaged in mathematics activities that are closely aligned with state standards. The impact of the will be tested in moderation analyses. Potential moderators include child, teacher, program, and location variables. We predict that implementation of the pre-K intervention as intended will have a positive and direct causal effect
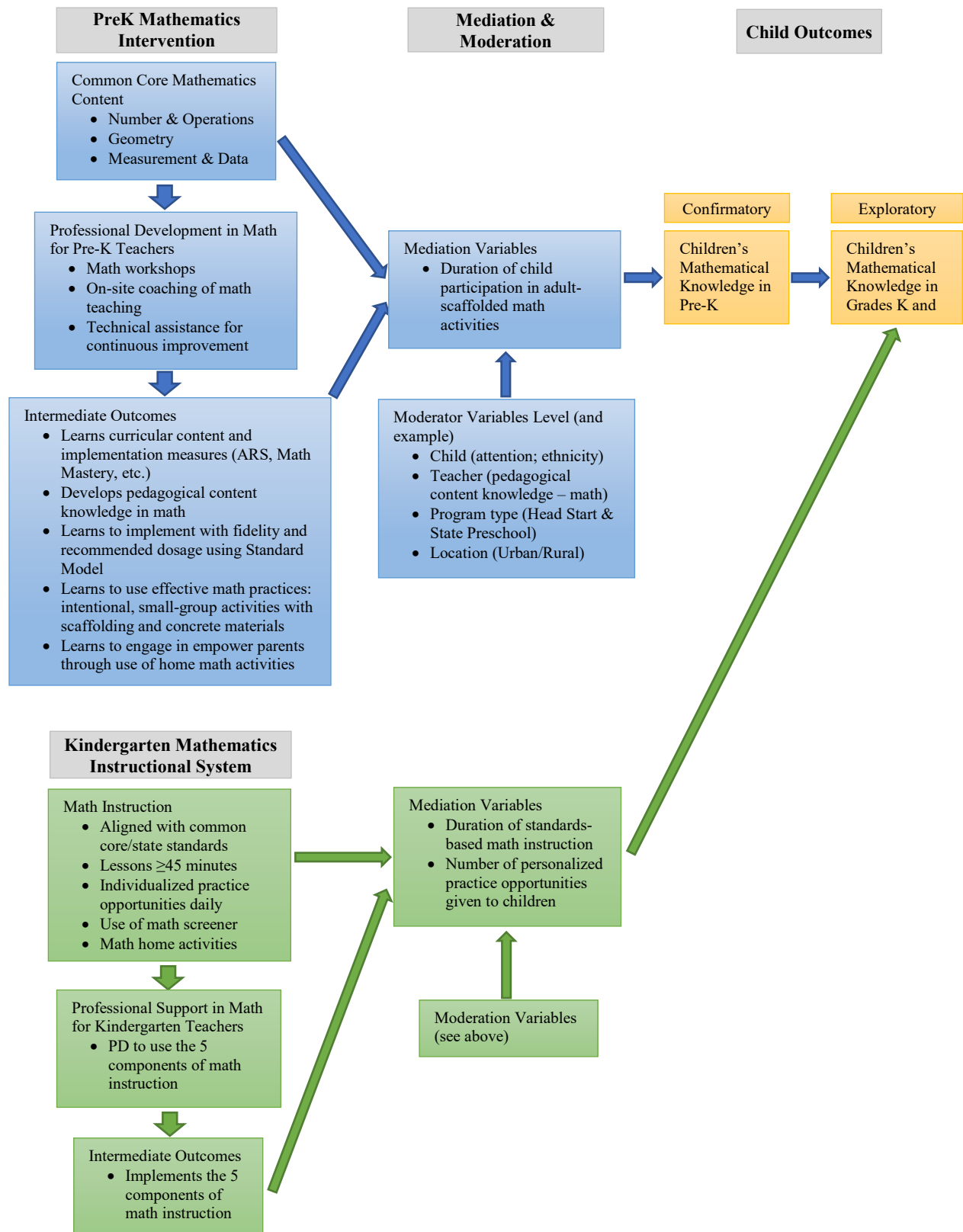
Figure 1. Model of Causation

on students' mathematical knowledge. The strengthened math program in K will maintain, or improve on, the gains children made in pre-K.

Components of the Innovative Intervention and Follow-Up Support in K

   The *Pre-K Mathematics* intervention. *Pre-K Mathematics* (Klein & Starkey, 2004) is innovative, in part because pre-k programs had not previously supported children's mathematics learning intentionally or systematically. Development and evaluation the intervention began approximately 20 years ago (Starkey & Klein, 2000). It underwent multiple evaluations, and the training and implementation models associated with it have been improved for the past decade (Klein, et al., 2008; Starkey, Flynn, & Klein, 2015; Starkey, Klein, & Wakeley, 2004). *Pre-K Mathematics* includes math activities that target the pre-K classroom and home learning environments of young children (**Appendix G**). The set of classroom math activities provide conceptually broad support for the development of children's informal mathematical knowledge. The intervention consists of small-group math activities with concrete manipulatives and a math learning center for the pre-kindergarten classroom. The mathematical content of activities is based on developmental research about the nature and extent of early mathematical knowledge. Units and activities within *Pre-K Mathematics* prepare children for each of the clusters of standards included in the CCSSM (**Appendix G**). They are also explicitly linked to NCTM Focal Points (NCTM, 2008). We have found that embedding pedagogical content knowledge into the intervention's math activities helps pre-K teachers learn to use it effectively. This is because many pre-K teachers have little or no training in mathematics teaching and learning. Downward (less challenging) extensions of the math activities are provided for children who are not ready for a given activity, and upward (more challenging) extensions are included for children who complete an activity easily. Common child errors or misunderstandings and suggested scaffolding to address these challenges are also provided. Assessment sheets that accompany each math activity enable the teacher to record individual children's learning over the course of the year **(Appendix G)**. A progress monitoring instrument is used by teachers to track children's

exposure to math content (curriculum dosage) and mastery of this content **(Appendix G)**. Children who encounter difficulty with particular key content are grouped periodically for extra dosage and scaffolding.

Teachers also send Spanish or English versions of math activities home to parents. Teachers receive training in reaching out to parents to conduct these activities with their children and send a Parent Feedback Form (home curriculum dosage) to parents to track parents' use of these activities **(Appendix G)**. Teachers also learn to mathematically enrich the classroom learning environment by (1) adding a math learning center or by making improvements if one is already in place, and (2) by systematically providing and tracking children's use of math software in a technology-based student center.

Follow-up Support: Kindergarten Mathematics Instructional System. Effects of math and other early childhood interventions persist, but diminish somewhat, as children matriculate through elementary school (Bailey, Duncan, Odgers, & Wu, 2016; Clements, Sarama, Wolfe, & Spitler, 2013; Starkey, Klein, & DeFlorio, 2014). This is not surprising, given that low-SES children's impoverished circumstances persist beyond their time in pre-K. Bailey et al. (2016) have recommended scaled-up early childhood interventions followed by use of "more advanced curricula" in elementary school. This was the successful approach that was taken in our i3 validation project and that will be taken again in this EIR expansion project. Support in K will intensify implementation of the district's math curriculum. As in our i3 project, when state and district policy allow, K teachers will choose between adoption of an evidence-based tier 1 curriculum, *Early Learning in Mathematics (ELM)* (Chard et al., 2008**)** or modification of their district's curriculum using a set of design features used in *ELM*. These features include curriculum alignment to the CCSSM and any non-CCSSM state standards, lessons scheduled to provide a minimum of 45 minutes of daily math instruction, individual practice opportunities that are informed by progress monitoring and provided daily, and home math activities provided weekly.

We will examine alignment of all K teachers math curriculum to standards and anticipate that some will need assistance. Although states require school districts to use standards-aligned mathematics content, a recent national survey of school leaders found that the instructional materials some teachers are using are not aligned with key aspects of state standards (Kaufman & Tsai, 2018). The survey authors pointed to a need for support for school administrators and teachers' in the selection and use of aligned materials. We will provide this assistance to K teachers through the regional training centers.

Support for personalized learning will be improved through use of a math screener at the beginning of the school year. Teachers will be trained to use the *Screener for Early Number Sense* (SENS) to identify children in their classroom who are at-risk in mathematics (Jordan & Klein, 2018; **Appendix G**). The SENS is a math screener for pre-K to grade 1 that has been developed for use by teachers in classroom settings. It has been developed and field tested through an IES instrument grant. To ensure that its number content is relevant to what is being taught in the classroom and to maximize its usefulness for educators and researchers, the SENS is closely aligned with the CCSSM in grades K and 1 as well as the Numbers, Relations, and Operations Core in pre-K and K endorsed by the Committee on Early Childhood Mathematics of the National Research Council (NRC, 2009). At the beginning of the school year, the SENS will identify students who are at-risk in their mathematical development and will likely require additional personalized learning opportunities throughout the year.

Scaling Framework

An important next step for intervention research is to determine whether interventions found to be effective in initial efficacy studies are scalable. There is no universally accepted framework in education for scaling up promising interventions and evaluating their effectiveness at a broader scale (Schneider & McDonald, 2007). The proposed project is primarily based on scaling frameworks by McDonald, Keesler, Kauffman, & Schneider (2006), Coburn (2003), an i3 white paper (DeWire, McKithen, & Carey, 2017), and our previous work on scaling and sustainability, including IES- and i3-funded projects described below.

The scaling framework of McDonald, Keesler, Kauffman, & Schneider (2006) emphasizes the increased number of settings and populations that often accompany scaling up. A program of scale-up research can progress through a series of stages. Stage 1 is comprised of initial efficacy studies that are conducted to determine whether an intervention (e.g., our *Pre-K Mathematics* intervention) improves an intended outcome (children's early mathematical knowledge) for a targeted population (low SES 4-year-olds) in specific settings (public preschool classrooms). Each study, however, is conducted with a sample having a particular demographic composition, in programs with particular characteristics, and in geographic areas that are home to particular types of communities. Stage 2 of scale-up research involves the accumulation of evidence about the effectiveness of a promising intervention across additional classes of persons and settings. Stage 3 scale-up research attempts to understand variables identified in stage 2 that impact the effectiveness or sustainability of the intervention (see Derzon, 2018, for a similar position). Thus, the focus shifts from whether to why effectiveness of the intervention varies. Our proposed EIR expansion project has stage 2 and stage 3 features. As describe in the **Evaluation Plan**, we will be scaling into some new populations and settings that were not included in our prior projects. As described in the **Training and Implementation** section, we also plan to collect implementation and contextual data that will make it possible to identify variables that impact scalability of the intervention.

Coburn's (2003) framework focuses on four interrelated dimensions of scale – spread, depth, sustainability, and ownership – all of which are an integral part of our approach for expanding implementation of the *Pre-K Mathematic* intervention into new populations and settings without diminishing its effectiveness. In our scaling context, the dimension of *spread* concerns expansion of the intervention within and across LEAs, which will become increasingly diverse in the populations it serves and the settings in which it is situated. We agree with Coburn that spread must occur at a deep level – the dimension of *depth*. The active ingredients of our intervention (its math content), its intermediate (i.e., teacher) outcomes (e.g., development and use of pedagogical content knowledge, use of effective math practices), and its principal

mediator (child participation in adult-scaffolded math activities) must spread into new classrooms for the intervention's effectiveness to be maintained at scale. We will utilize a professional support model that has been successfully used at a statewide level to ensure that spread occurs at this necessary level of depth (see **Model of Causation; Training and Implementation**). Coburn's other dimensions of scale, *sustainability* and *ownership*, are intertwined. We have addressed these dimensions of scale in a recent IES-funded continuous improvement project (Starkey et al., 2015). This project tested a continuous improvement process that was intended to improve the sustainability of the *Pre-K Mathematics* intervention and the deep change in classroom practice that the intervention entails. The continuous improvement process was found to be sustainable, as described in detail below. Briefly, LEA teams (early childhood teachers and administrators) engaged in a guided process of scaling in the intervention (Cohen & Ball, 2007). This process involved making local adaptations that LEA teams deemed necessary for sustainability (e.g., providing training and setting responsibilities for instructional assistants, and agreed-upon administrative supports and procedures), but with our guidance, ensuring that essential features of the intervention (fidelity and curriculum dosage) were preserved. Scaling-in positively impacted both sustainability and true ownership of the intervention (Starkey et al., 2015).

Unmet Demand for Effective Math Support in Public Preschool Programs

The adoption of curriculum guidelines and standards that include mathematics for public preschool programs represent a demand by federal and state policy makers and early education administrators for effective math support. Local programs are expected to utilize curriculum and instructional practices that prepare children from low-SES backgrounds for elementary school mathematics (e.g., Florida Department of Education, 2017; Health and Human Services, Administration for Children and Families, Office of Head Start, 2015).

For many years, public preschool programs did not have rigorous evidence of whether they were meeting this demand. This began to change in the 2000s, with the advent of funding for rigorous evaluations by the Administration for Children and Families (ACF) and the Institute

of Education Sciences (IES). ACF's Head Start Impact Study was an RCT at the national level of the effectiveness of Head Start on multiple child outcomes, including mathematics (U.S. Dept. of Health and Human Services, Administration for Children and Families, 2010). Head Start was found to be ineffective in the domain of mathematics for 4-year-olds. IES's Preschool Curriculum Evaluation Research (PCER) program funded rigorous evaluations (RCTs) of early childhood curricula, including the Creative Curriculum, which is the most widely used curriculum in public preschool programs. Both RCTs that evaluated the effectiveness of the Creative Curriculum found it to be ineffective in the domain of mathematics for 4-year-olds (PCER Consortium, 2008).

Improving low SES children's math outcomes, however, has proved very difficult for programs to achieve. One strategy that programs have tried is to increase the percentage of teachers with BA degrees in Head Start programs. Head Start teachers, in general, have less education than K-12 teachers. A policy-driven approach, the Head Start Act of 2007, is intended to improve the effectiveness of Head Start teachers by increasing teacher education requirements. Research, however, has cast doubt on the sufficiency of this approach to produce effective Pre-K teachers in academic areas such as emergent literacy and math. Early and colleagues (2007) used seven major studies of early care and education to predict classroom quality and children's academic outcomes based on teachers' education level. Findings were largely null or contradictory, indicating that policies focused solely on increasing teachers' education are not sufficient for improving young children's growth in academic domains.

A strategy that state preschool programs have tried is to use quality rating improvement systems (QRIS) to improve program elements, such as teacher-child interactions (Pianta, La Paro, & Hamre, 2008) the general program environment (Harms, Clifford, & Cryer, 2014), and administrator qualifications, in state Pre-K programs. Quality Rating and Improvement Systems (QRIS) are a nationwide effort to better understand the quality of early education that is required to produce desired child outcomes (QRIS National Learning Network, 2015). Forty-nine states

currently implement or plan to implement QRIS, and 20 states have participated in the $550,000,000 Race to the Top-Early Learning Challenge (RTT-ELC) program to improve the quality of early learning and development programs and thereby improve developmental and school readiness outcomes of young children with high needs. Use of a QRIS makes it possible to rate the quality of state funded local education agencies (LEAs), and centers or individual classrooms within an LEA, on a common scale within a state. The rating system is intended to spur improvements in program quality, with the goal ultimately of improving child development and school readiness outcomes. RTT-ELC grantees are required to conduct QRIS validation studies. These and other QRIS studies have revealed a problem regarding the ability of QRISs, as currently structured, to predict child outcomes. These studies have been reviewed (Karoly, 2014) and the review has been recently updated (American Institutes for Research & RAND Corporation, 2016). QRIS validation studies, however, have obtained little evidence that ratings predict (i.e., are correlated with) child development and school readiness outcomes, and provide no evidence that the improvements made are causally related to improved child outcomes. It is noteworthy that none of the QRIS efforts included evidence-based curricula among the improvements that were tested.

Neither the strategy of increasing the percentage of Head Start teachers with BA degrees nor the QRIS strategy have been found in rigorous evaluations to improve children's math outcomes. Thus, there is a widespread and unmet demand for effective math support in public preschool programs. *Pre-K Mathematics* has the potential to meet this demand, given the extensive evidence of its effectiveness.

We next identify barriers preventing widespread use of *Pre-K Mathematics* and the expansion strategies that will be employed to address these barriers.

Barriers that Prevent Scaling Up the *Pre-K Mathematics* Intervention and Strategies for Removing Them

Barrier 1. Pre-K teachers have insufficient professional preparation to teach mathematics effectively. Pre-K teachers do not presently have the professional preparation

needed to implement the *Pre-K Mathematics* intervention effectively. There is abundant evidence that, at present, the quality and quantity of mathematics teaching at the preschool level is inadequate. Preschool teachers report that they infrequently engage preschool children in math-related activities (Thornton, Crim, & Hawkins, 2009) and, indeed, research has found that preschool children spend only 8-11 minutes per school day engaged in math activities (Connor, Morrison, & Slominski, 2006; Starkey et al., 2014). Much of this time is spent repeatedly on a few limited topics such as counting, to the exclusion of other mathematical concepts and skills (Copley, 2010).

The limited amount and quality of math instruction in preschool classrooms likely result from teachers' inadequate preparation in and knowledge of informal (concrete object-based) mathematics and children's early mathematical learning and development (e.g., Isenberg, 2000). A survey of colleges providing preschool educator training found that only 21% of programs offered an entire course focused on academic content, comprised of math, language and literacy and other content, with the remainder of programs offering no coursework or part of one course on academic content (Lobman, Ryan, & McLaughlin, 2005). Further, in a joint position paper, the National Association for the Education of Young Children and the National Council of Teachers of Mathematics specifically attributed the weak quality of preschool mathematics instruction to the general lack of good preschool teacher preparation in mathematics (National Association for the Education of Young Children, 2005).

Solution 1.A. Establish a national network of regional training centers. Despite calls for stronger pre-K mathematics instruction, a national training capacity in this content area has not been developed. We will develop this capacity by constructing regional training centers, modeled after the statewide center that was established in California in our current i3 validation project. The reach of this center has been expanded into neighboring states and is now providing LEAs in Arizona, California, and Nevada with early math support. This support includes trainers' institutes, multi-day math workshops, on-site coaching, continuous improvement support to sustain effective math instruction, support for family math events, and conference

presentations. This center provides trainers and teachers with training in early math, including three early math interventions, (1) *Pre-Pre-K Mathematics*, a tier 1 curricular intervention for 3-year-olds, (2) *Pre-K Mathematics* for 4-year-olds, (3) *Pre-K Mathematics Tutorial* for use with very low performing 4-year-olds, as well as tier 1 curriculum support for grade K. All of the early math interventions have been found to be effective in IES efficacy projects (e.g., Barnes, et al., 2016; Klein et al., in press). Regional centers serving three broad regions, the Mid-Atlantic, Mid-West, and South, will be established, the existing center serving the West will be expanded, and all four will be networked through the national office in California. This overall effort is our Early Math Initiative, and by the end of the project these centers will be scaled to serve 23 states and the District of Columbia, with the schedule for further expansion into other regions to be dictated by demand.

Solution 1.B. Provide pre-k teachers with PD and coaching that will enable them to teach mathematics effectively. The regional training centers will provide Pre-K teachers intensive in-service PD as described below in the **Project Design**. We have found that the 44 hours of workshops and coaching support that teachers receive is effective. The PD and the experience of implementing the *Pre-K Mathematics* intervention effect a deep change in teachers' mathematics teaching. We have found in multiple projects that this deep change in teachers mediates the causal influence of the *Pre-K Mathematics* intervention. The Early Math Initiative will make it possible to cause this deep but necessary change as the intervention is scaled up.

Barrier 2. There are no ready-made manipulatives for *Pre-K Mathematics* activities. The *Pre-K Mathematics* intervention requires that teachers use manipulatives when they present math activities to children. We initially attempted having LEAs produce sets of manipulatives, but programs made production errors and gave feedback that producing materials was an inefficient use of program staff time.

*Solution 2A. Provide commercially produced materials kits to teachers***.** We have secured an agreement with a commercial materials producer to produce and market materials kits

to accompany *Pre-K Mathematics* activities **(**see *Lakeshore* letter in **Appendix C**). An initial sale of 1,000 sets is a condition of starting production of these materials. The Early Math Initiative will enable production of these materials to begin. Future production will depend on orders from Pre-K programs.

Barrier 3. Start-Up Costs and Curriculum Roll Outs Require Long-Term Planning. Head Start grantees and state preschool programs receive discretionary funding for professional development. A barrier our intervention faces is that Pre-K programs need information on costs and roll-out options for training their staff to support and implement *Pre-K Mathematics* effectively. Without this information, programs cannot engage in financial planning needed to commit to a rollout of *Pre-K Mathematics*.

Solution 3.A. Provide public preschool programs with pricing and multi-year roll-out options. Our i3 validation project was sufficiently large to enable us to determine that *Pre-K Mathematics* start-up costs were $4,200 per classroom and $176 per student. No annual fees are charged after start-up. The only necessary recurring costs incurred in future years are for home activities, which families keep: $19 per student. Some programs conduct a family math night at the beginning of the school year. Costs vary widely, depending on whether teacher and parent volunteers or paid staff run the event and whether pot-luck or catered food is provided; the only hard costs are $2.50 per family for a door prize, with every family winning a set of Unifix Cubes. Programs can also incur the cost of training new teachers if teacher turnover occurs. Thus, if no teacher turnover occurs, recurring costs are very low. A member of our national office will create and price roll-out options and make this information available to public preschool programs through our regional centers.

Solution 3.B. Establish partnerships with programs to identify and leverage state and federal funding opportunities for professional development. The project director will be responsible for knowing and tracking public preschool policy and funding for professional development. The Program Liaison for our national office, with broad administrative experience in California's state preschool program, will work proactively with Head Start grantees and state

preschool programs to select, budget for, and schedule a rollout option. National office staff will also collaborate with SEAs and local programs on federal, state, or private funding opportunities for PD (e.g., Indian Education grants and ED/OII/SEED grants).

Solution 3.C. Lower costs by utilizing economies of scale and technology. Trainings will be combined across multiple LEAs when feasible. This will lower costs and make it financially feasible to conduct trainings in some rural areas. Progress monitoring and class curriculum dosage monitoring will be moved from a paper-based tool to a programmed spreadsheet to reduce PD costs. The rationale for the shift to a digital system context comes from a continuous improvement project (Starkey et al., 2015). Teachers, who had one or more years of experience using the paper-based progress monitoring tool while implementing the *Pre-K Mathematics* intervention, expressed the need for a way to monitor children's progress that requires less time and paperwork. In addition, some program administrators expressed a need to store and transfer implementation data electronically. Third, continuous improvement meetings to foster sustainability are conducted at the end of the school year. When possible, these meetings will be conducted remotely using video conferencing.

Plan for Disseminating Project Products and Findings at the Local and National Level

Project findings and products will be actively disseminated using multiple strategies. Our approach to dissemination will expand out the dissemination procedures that staff in our i3 validation project are currently using in California and neighboring states.

Dissemination strategy 1. We currently use direct contact to disseminate project findings and products to public preschool programs. Lists of local programs were obtained from regional associations and funding agencies. Direct contact is made with local Head Start and state preschool programs through an e-mail survey, with information provided in the e-mail and through a link to WestEd's website. This is followed by phone calls to interested programs in which funding, training, and roll-out options are discussed. During the proposed EIR Expansion project, this strategy will be implemented in the regions served by the regional training centers.

We will subsequently expand dissemination into other regions, adding training centers in accordance with demand from the field.

      Dissemination strategy 2. A dissemination strategy targeting national and state stakeholders will be used after the main confirmatory findings have been obtained. WestEd newsletters will be sent to stakeholders. Also, empirical presentations will be made at national practitioner, research, and policy-oriented conferences, such as conferences held by the National Council of Teachers of Mathematics, the National Head Start Association, the Society for Research on Educational Effectiveness, and the Association for Public Policy Analysis & Management.

Project Design and Management Plan

Principal Project Goals, Objectives, and Expected Outcomes

      The principal goals of this project are (1) to remove barriers that impede scaling up the innovative *Pre-K Mathematics* intervention to the national level, (1) to improve low-SES children's readiness for elementary school mathematics, and (3) to obtain scientific evidence on whether and why the effectiveness the intervention varies among populations or settings. Goals, project objectives, and outcomes are described below and in Table 1. Management Plan: Project Goals, Objectives, and Strategies in **Appendix G.**

      Goal 1. Remove barriers that impede scaling up the *Pre-K Mathematics* intervention to the national level. Our three objectives for accomplishing this goal are to remove three barriers that have been identified. Objective 1.1 is to remove the first barrier - insufficient professional preparation of pre-K teachers to teach mathematics effectively. Pre-k teachers could not implement *Pre-K Mathematics* effectively if they were simply given the teachers manual. They require professional development (PD) to implement the intervention as intended, with sufficient fidelity, curriculum dosage, and record keeping, to implement effectively and to sustain an effective implementation. The national capacity for early math training, however, is insufficient to provide this needed PD to the nation's pre-K teachers. Therefore, the strategies for

meeting this are: (1) the project director, national trainer, and experienced trainers in our training network will train regional trainers in the regions in which the project will be conducted, (2) teams of regional trainers, with quality control monitoring (including shadowing and record inspection) by national staff, will conduct trainers' institutes for local trainers in each region, (3) teams of local and regional trainers will conduct pre-K teacher workshops, (4) local trainers, with quality control monitoring (including shadowing and record inspection) by regional staff will conduct on-site coaching, and (5) regional trainers will provide sustainability (continuous improvement assistance) to LEA teams (local trainers, pre-K administrators, and teachers). The numerous milestones measuring the success of these strategies are step-by-step completion of the training targets in 1-5 above (e.g., completion of training and certification of regional trainers in regions 1 and 2). To ensure quality control in the development of regional centers, we will document key training activities using a Trainers' Fidelity of Implementation Matrix (**Appendix G**). Data sources will include the following: attendance sheets for trainers' institutes; certification records for three curriculum activities with a lead trainer and children in the classroom; records showing interrater reliability was established through a fidelity certification process and three co-fidelity visits during each school year; and notes from biweekly trainers' check-ins. A general timeline for completion of training and implementation objectives is given in Appendix G. A successful outcome for objective 1.1 will be creation of sufficient training capacity in each of four regions to provide on-demand training for programs in the region.

Measurement of the quality of training provided to regional and local trainers will be accomplished using a certification system. Measurement of the quality of teacher training will be accomplished through use of a Fidelity of Implementation Matrix used in our i3 project. This matrix is a record of the amount of PD provided to each teacher, a fidelity score, and classroom and home math activity dosage records, for each teacher (**Appendix G**).

In this proposal, scaling barriers pertain to the *Pre-K Mathematics* intervention. Capacity building, however, is also needed to provide follow-up support for kindergarten teachers. The national staff will provide training to regional trainers, who will then provide assistance to K

teachers grouped by LEA. The national staff will provide quality control monitoring (including shadowing and record inspection) of regional trainers. As for pre-K, a task-by-task spreadsheet will be used to manage and track execution of strategies and accomplishment of capacity building. A successful outcome will be creation of sufficient capacity in each of four regions for providing teachers with training and assistance for use of the K Math Support System. Outcomes will be measured using a training matrix as described above. As done in pre-K, a certification system will be used for trainers and a Fidelity of Implementation Matrix will be used to record the amount of PD and assistance provided, fidelity (e.g., duration of observed math lessons, personalized practice opportunities, etc.).

The second barrier is the unavailability of readymade manipulatives for *Pre-K Mathematics* activities. Objective 1.2 is for manipulative kits to be produced. The measurable outcome will be receipt of the number of units ordered. The third barrier is insufficient advance planning for start-up costs and curriculum rollout. Objective 1.3 is to for local programs to use long-term planning to determine how and when they can cover the costs of training their teachers to teach math effectively. The measure is the number of programs that have a specific plan for obtaining this training. Strategies to accomplish this objective are scaling solutions 3.A, 3.B, and 3.C described above. The measurable outcome will be a count of programs that are able to determine when and how to provide math training for their teachers.

Goal 2. Improve low-income children's readiness for elementary school mathematics. The *Pre-K Mathematics* intervention is intended to improve children's mathematical knowledge and prepare them for K mathematics curricula aligned to rigorous state standards including the CCSSM. Objective 2.1 for meeting this goal is for pre-K teachers in the four regions to implement the *Pre-K Mathematics* intervention effectively, defined as with adequate fidelity ($\geq$ .90) and delivering at least the minimum recommended classroom dosage ($\geq$ 75%) and home dosage (usage, $\geq$75%). Objective 2.2 is for K teachers to implement a K mathematics support system effectively, defined by an implementation checklist (all 5 components used – see Model of Causation), tier 1 math fidelity fidelity ($\geq$ .90), and classroom

dosage (≥ 75%) and home dosage (usage, ≥75%). See **Implementation Measures** for information on how implementation data are collected, and **Appendix G** for instruments used to measure implementation. Objectives 2.2 and 2.3 are improvement in children's mathematical knowledge in pre-K and K, respectively (see **Evaluation Plan**). It is predicted that children who receive the *Pre-K Mathematics* intervention, relative to children who do not, will experience significantly greater growth in mathematical knowledge across the pre-K year; this growth will meet the What Works Clearinghouse's most rigorous standards, including the standard for substantively important effects (effect size ≥ 0.25; measure: ECLS-B math). It is also predicted that treatment children who receive the intervention, relative to control children who do not, will score at an average or above level in math on the ECLS-B in pre-K. An exploratory question is whether children who receive the Pre-K intervention followed by strengthened math support in grade K, relative to children who do not, will have significantly higher math scores at the end of grade K and will score at an average or above level in math on ECLS-K:2011 in K. Another exploratory question is whether fewer treatment children than control children will be at-risk on the SENS math instrument at the end of K. The Model of Causation (see **Figure 1**) depicts the causal relation between the active ingredients in the intervention and child outcomes.

   Goal 3. Obtain scientific evidence on differential impact among sub-groups and variables related to differences that are found. Our evaluation team includes a purposive and random sampling plan that was developed by a methodologist who specializes in statistical sampling. The sampling plan and coordinated recruitment plan used in this project will produce a sample that is representative of the national population in Head Start and categorical (family income-restricted) state preschool programs. Objective 3.1 is to recruit a child sample whose composition is no more than 5% discrepant from what was called for in the sampling plan (see **Evaluation Plan)**. Objective 3.2 is improvement in children's mathematical knowledge by sub-groups of children or settings in pre-K and K. The objective under this goal is to examine growth in mathematical knowledge by children in one population, setting, or population X setting sub-group relative to children in a second population, setting, or population X setting sub-

group (see ECLS-B, ECLS-K, and **Data Analysis Plan**). Objective 3.3 is to identify variables related to any differential impact that is Analyses will be conducted to identify variables (e.g., children's entry level math scores, differences in trainers' training quality, teachers' implementation quality, or parents' home curriculum dosage) that are related. The variables of interest are those examined by implementation measures (e.g., fidelity and classroom dosage; classroom observations of practice: minutes of math support on EMCO, see Evaluation Plan). These analyses will help us determine whether scaling barriers were not eliminated evenly across populations or settings or whether additional barriers exist that had not been previously identified.

Plan for Achieving Goals and Objectives Through Systems Management

The process of meeting the three principal goals of the project will be accomplished and managed through a set of project objectives with specific strategies, milestones, tasks, and expected outcomes, and responsible agents and timelines for accomplishing tasks. Based on this information and in collaboration with our Federal Project Officer, we will develop and review monthly a detailed management plan. We will employ a management-by-objectives system to systematically track project activities. This system will be crucial for regularly and systematically communicating expectations, problems or barriers, as well as ensuring quality control in conducting activities and producing desired outcomes.

Overall project preparation, monitoring, and quality improvement. One of the most critical challenges with large, complex projects is maintaining effective partnerships with LEAs and other members of the overall project team including the evaluation team. It is essential to establish procedures for monitoring and quality improvement early in the project. We will spend the first two years of the project (1) negotiating and executing a detailed MOU with each LEA partner in cohort 1 (in year 1) and in cohort 2 (in year 2), and (2) developing and implementing a communications system and organizational chart for the project team. In reviewing the MOU details with LEA representatives, the LEA will identify a primary project lead from the agency.

Monthly check-ins with the LEA project lead, and other relevant stakeholders as needed, will be included as part of the MOU. In our experience with other large-scale projects, these regular check-ins with LEAs are crucial for ensuring quality control in conducting activities and identifying potential problems or local barriers in advance.

Further, the WestEd team has extensive experience working with multiple collaborators within large-scale projects. To streamline communication, various WestEd team members will take lead responsibilities associated with various project activities and any external agencies responsible for those activities (see **C.3. Ensuring Feedback and Quality Control** for details). In addition to meeting regularly with external team members, WestEd will hold monthly internal team meetings to systematically communicate expectations and ensure quality control in conducting activities across the major project objectives (e.g., implementation vs. evaluation).

Training and implementation procedures for the *Pre-K Mathematics* intervention. For teaching that uses the *Pre-K Mathematics* intervention to be effective and sustainable, it is important that (1) Pre-K teachers receive sufficient PD to implement *Pre-K Mathematics* with fidelity and at recommended dosage levels, and (2) that each LEA use a locally sustainable model of implementation. The National Trainer will work closely with the regional training centers and report regularly to the WestEd team to ensure both of these conditions are met.

Professional development: teacher workshops, on-site coaching, and continuous improvement. *Pre-K Mathematics* workshops will be conducted for teachers in each local program. Teachers will be required to attend two three-day workshops, one in summer and one in winter during their first full year of implementation. These workshops are designed to enable pre-K teachers to receive professional training to become effective early math teachers. Through the workshops, teachers will receive professional training in (1) the *Pre-K Mathematics* intervention, including lead trainers' demonstrations of activities and explanations of math content in these activities, discussion of *Pre-K Mathematics* practices found to be effective, and hands-on practice by teachers in groups using the classroom activities, (2) collection and use of implementation data, including authentic assessment of children's performance on each math

activity (see ARS, **Appendix G**), progress monitoring, home dosage records, and preparation of pre-K math reports on each child at the end of pre-K, (3) efficient use of review days for children who encountered difficulty with some activities or were absent, (4) the home math activities and parent outreach procedures, including staging of a family math event, (5) mathematical enrichment of the classroom learning environment periodically throughout the school year, (6) systematic, intentional use of computer math software, with an instructional assistant trained to mind the ecology of the computer area (e.g., enforcing turn taking and ensuring that all children use it weekly, (7) early mathematical development, including research demonstrating an SES gap in early math development and the positive effects of early intervention on early mathematics achievement, (8) early math milestones and standards (e.g., Head Start's Mathematics Development Domain; Common Core State Standards for Mathematics for kindergarten), (9) using pedagogical content knowledge to decide when and how to make developmental adjustments during math instruction, (9) small-group and classroom management techniques, (10) supporting EL children, (11) personalizing children's learning, (12) the purpose of having local trainers conduct classroom visits during implementation of *Pre-K Mathematics*, and (13) use of a continuous improvement process to sustain an effective implementation of *Pre-K Mathematics* long-term.

After each workshop, local trainers will provide teachers with on-site coaching support and formative feedback as they implement *Pre-K Mathematics* activities in their classrooms. This support, which will be provided during a minimum of eight classroom visits per classroom across the school year, will help teachers (1) learn to implement all aspects of *Pre-K Mathematics* with fidelity, as scheduled, at the recommended level of dosage, using progress monitoring, and (2) mathematically enrich the classroom learning environment with a math learning center and a technology-based independent student center(see Fidelity of Intervention: *Pre-K Mathematics*, **Appendix G**). Additional training and technical assistance (TA) will be provided by the local trainer as needed during these visits.

Continuous improvement. Finally, regional and local trainers will lead sustainability efforts via a continuous improvement process described above. Specifically, project leaders will conduct monthly check-ins with administrators to develop administrative procedures to monitor and support implementation and to identify and resolve local implementation challenges. Procedures developed for each LEA will be recorded. This process will help sustain an effective implementation of *Pre-K Mathematics* and provide a process for addressing future implementation challenges that may arise.

Monitoring and quality control of implementation in pre-K. To ensure quality control for each LEA, data will be collected by trainers during workshops and coaching visits to document implementation, formative evaluation, and progress monitoring by the teachers (see Fidelity of Intervention: *Pre-K Mathematics*, Math Mastery, and PK Fidelity of Implementation Matrix, **Appendix G**). Data on program-level training activities will include (1) information about participation in Trainers' Institute (source: sign-in sheets at 4-day workshop, documentation by Regional Trainer of 2-day field PD and 3 co-fidelity visits), and (2) evidence of Continuous Improvement Technical Assistance (source: notes from initial meeting, monthly check-ins, family outreach event, and end of year meeting). Data on classroom-level training activities will include: (1) information on teacher workshops (source: sign-in sheets completed at workshops), and (2) the frequency and quality of on-site facilitation provided to teachers (source: trainers logs and field notes; co-fidelity visit data collected from regional trainers). At each bi-weekly training visit to teachers' classrooms, local trainers will check and record implementation of each component of the intervention expected from teachers, including classroom math activities and distribution of home math activities. Any type of formative feedback given will be recorded in a box on their fidelity form. Data on teachers' implementation activities will include (1) teachers' records of dates that specific math activities were implemented, and which children participated (source: digitized progress monitoring and online tracking system), and (2) parent feedback forms that report use of each home math activity by individual families (see sample Parent Feedback Form, **Appendix G**).

Training and implementation procedures for the Kindergarten Mathematics Instructional System. It is important that school administrators are prepared to support K teachers math instruction, and that K teachers will implement their district curriculum through a standards-aligned scope and sequence. The National Trainer will work closely with the regional training centers and report regularly to the WestEd team to ensure these outcomes are met as scheduled.

Math needs assessment survey (MNAS). In the spring prior to the year of the kindergarten intervention, relevant district and school administrators, as well as kindergarten teachers will be surveyed to determine their existing practices for mathematics instruction. Topics of survey will include: adopted math curriculum, its alignment to the state standards, and teachers' opinions about the curriculum; amount of time spent on math instruction; procedures in place for enhancing math learning in the home environment; instructional practices in place for personalized learning; use of a math screening instrument to identify at-risk students; practices in place to monitor student progress; and how technology is being used to support math practice.

Professional development component: K teacher workshops and on-site coaching. Kindergarten teachers who are dissatisfied with their current math curriculum will be offered the option of learning to use *Early Learning in Mathematics*, if allowed by their state and districts. They will be provided intensive in summer and winter workshops and on-site training and implementation monitoring during the year, and they will implement the K math intervention according to a weekly curriculum plan. All teachers, regardless of the math curriculum they use, will receive two-days of professional development, one just prior to the school year and one about halfway through the school year. Outcomes for the initial one-day workshop include: (1) creating a scope and sequence for the first half of the year that includes mathematical activities and practices aligned to the state standards, (2) developing a general daily schedule for 45 minutes of math instruction that includes individualized practice opportunities, (3) establishing a system for consistent and effective progress monitoring either as part of their adopted math

curriculum or as a general system, and (4) planning for sending home and tracking parental use of home math activities on a weekly basis.

Regional and local trainers will guide teachers through their district's mathematics curricula to identify curriculum activities that align with their state standards. Every effort will be made to include school leaders in collaborative working groups with teachers to create the scope and sequence for the first half of the year. In addition to selecting appropriate activities, special attention will be given to maximizing effective practices by ordering activities in a development sequence supported by research.

Additionally, regional and local trainers will provide direct training and guided practice on the *SENS*, along with instruction on identifying fundamental K math activities from their curriculum that could be used to intensify instruction or to provide extra practice for children at-risk based on the SENS results. Collaborative working groups will also be utilized to establish the progress monitoring system with instruction on how to effectively adjust instruction based on progress-monitoring data. Finally, teachers will be trained on how to enhance math learning in the home environment, so that it works in conjunction with the classroom math curriculum and how to use a technology-based student independent center to increase opportunities for math practice.

The second one-day workshop will focus on developing a scope and sequence for the second half of the year, and taking deeper dives into working with progress monitoring data, strengthening the home-school connection, and maximizing personalized learning opportunities, including digital activities. Teachers will prepare K math reports linked to individual K standards in CCSSM on each child at the end of K.

In-class support. After each workshop, regional local trainers will provide teachers with on-site coaching support and formative feedback as they implement the standards-aligned scope and sequence of their district-adopted curriculum in their classrooms. This support, which will be provided during eight classroom visits per classroom across the school year, will help teachers (1) learn to implement standards-aligned math activities for 45 minutes a day, use the SENS and

progress monitoring, and (2) mathematically enrich the classroom learning environment with a math learning center and a technology-based independent student center. Additional training and technical assistance (TA) will be provided by the local trainer as needed during these visits.

Monitoring and quality control of implementation in kindergarten. For each LEA, data will also be collected by trainers during workshops and coaching visits to document use of all five components of the Kindergarten Professional Support System using a K Fidelity of Implementation Matrix (**Appendix G**). Data on program-level training activities will include evidence of Technical Assistance to school leaders on curriculum alignment and data-driven decision making (source: notes from initial meeting, monthly check-ins, family outreach event, and end of year meeting). Data on classroom-level training activities will include: (1) information on teacher workshops (source: sign-in sheets completed at workshops), and (2) the frequency and quality of on-site facilitation provided to teachers (source: trainers logs and field notes; co-fidelity visit data collected from regional trainers). At each monthly (eight visits total) training visit to teachers' classrooms, local PD specialists will check and record implementation of each component of the activities expected from teachers, including use of curriculum activities aligned with the state standards, 45 minutes of math instruction, use of home activities, and utilizing strategies for personalized learning (e.g., independent student practice, digital activities). Any type of formative feedback given will be recorded on their fidelity form. Data on teachers' activities will include (1) teachers' records of dates that specific math activities were implemented, and which children participated (source: progress monitoring system), and (2) parent feedback forms that report use of each home math activity by individual families like those used in pre-K

Management and quality control for participant retention and data collection. Management of the evaluation includes (1) recruitment of a sample whose composition conforms to the sampling plan with a discrepancy of 5% or less, (2) longitudinal tracking of the sample from pre-K to grade 1, (3) using administration protocols and manuals for each instrument (developed during our i3 validation project), (4) hiring, training, and certifying data collection

staff in each region for each data collection waive, (5) collecting all planned data during designated data collection windows, and (6) scoring, organizing it for analysis, and analyzing it throughout the project to ensure timely reporting of project findings. The independent evaluation (see **Evaluation Plan**) will be directed by Prof. Tom Cook and conducted by the same team of researchers who evaluated on our i3 validation study, each specializing in a specific aspect of the evaluation: SRI International for data collection, ETS for ECLS instrument scoring, and Dr. Jaime Thomas who will assist Prof. Cook in data analysis (see details below in C.4). Dr. Elizabeth Tipton, a sampling statistician, is joining the team to design the national sampling plan, oversee sampling, and collaborate on data analysis and report preparation. WestEd will track the sample by employing the same procedures used to retain the i3 validation study sample sufficiently to comply with conservative WWC attrition standards, a sample that likewise included a transition from the pre-K to kindergarten year, and a follow-up assessment at grade 1. WestEd will also update the administration protocols and manuals for the assessments used in prior studies (e.g., ECLS-B, ECLS-K), and create new ones as needed. SRI International, the data collector for the i3 validation study, will hire and train assessors, and conduct the data collection and cost-effectiveness analyses. Finally, Prof Cook will oversee the evaluation, and in collaboration with Dr. Thomas, will draft evaluation reports, conduct data analyses, and participate in leadership team meetings.

Participant retention. Based on our collective experience, this team is already familiar with potential challenges that can arise in all aspects of large-scale research studies conducted in field settings, including LEAs. We have proactive systems in place for handling them. To address LEA and teacher retention, the WestEd team will draw on experience from our IES Continuous Improvement Education Grant (Starkey et al., 2015) to keep LEAs engaged over multiple years of the study and to help them develop and use a sustainable model of implementation based on their respective local needs, rather that dropping the intervention due to perceived implementation or staffing challenges. Through classroom coaching visits and use of a continuous improvement process, which will include monthly check-ins with program directors,

biannual site visits by senior researchers, and reviewing of progress monitoring data, LEAs will be assisted with scaling in their implementation and, in doing so, will usually develop a sense of ownership of the intervention. Also, since we have successfully implemented the *Pre-K Mathematics* intervention in multiple projects, it has a highly developed and feasible implementation model, including teacher workshops, in-class coaching, and trainers' institutes, and easily used instruments for measuring implementation. This planning and experience will facilitate retention of teacher participants, ensure high quality implementation, and facilitate sustainability.

Pre-k children will be screened during the consenting process to ensure that they meet the district residency requirements to attend a participating elementary school for K. During the screening process, children's addresses will be checked to ensure that they match a participating project neighborhood school, and parents will be asked to state their intention for K enrollment to confirm their child's attendance at a matching participating school. Once the study begins, researchers will create a secure database to track children's classroom placements throughout the school year. For the continued accuracy of the database throughout the school year, researchers will check in with teachers by phone once per month to briefly review the list of participating students to confirm that all are still enrolled and will follow-up with families of children who have moved or transferred to confirm the new placement.

Data collection in the field. Further, the collective experience between WestEd and SRI will be used to ensure the successful collection of implementation data and child assessments. Specifically, to safeguard the quality and accuracy of implementation data, WestEd will proactively monitor the data collection through coaches' surveys, biweekly check-ins with coaches, and data checks by research assistants. Further, SRI will employ careful procedures to ensure that child data are collected accurately and as scheduled (see **D. Independent Project Evaluation**). For example, quality control checks of data will be made, both in the field and redundantly at headquarters, early in each wave of data collection and for each data collector.

Finally, various WestEd team members will take lead responsibilities associated with various project activities and meet monthly with the relevant external agencies responsible for those activities to ensure that project objectives are being met on schedule. We describe the team members and their respective roles and responsibilities below.

Qualifications of the Project Team

We are experienced and extremely well-positioned to conduct an expansion phase project at the national level. Our team at WestEd, including senior staff, pre-K and K trainers, and research assistants, and the external evaluation team collaborated on our successfully executed i3 validation project (Starkey, 2012).

National office and responsibilities. *Dr. Prentice Starkey* of WestEd (see CV, Appendix B), will serve as the Project Director (PD) and will oversee the entire project, including monitoring the activities of the National Office. Dr. Starkey, along with Principal
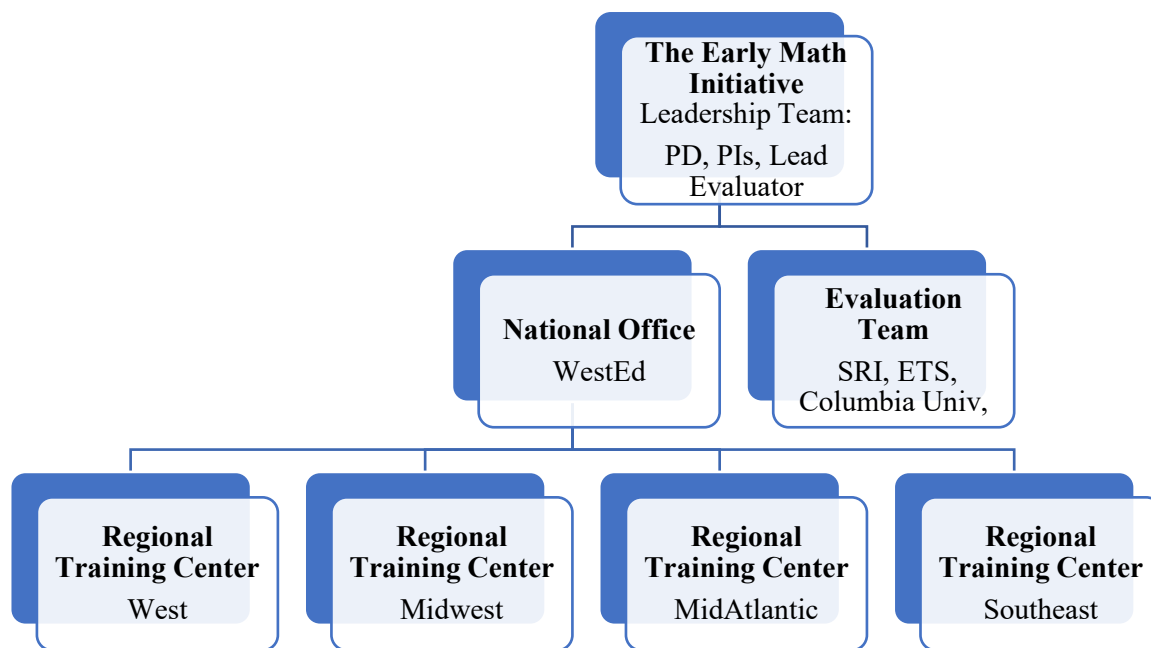
Investigator (PI) Dr. Alice Klein (see below), has successfully directed or co-directed two multi-state intervention projects (Klein, et al., 2008; Starkey et al., 2014), an i3 validation study on a statewide scale, and several other IES, NSF, and NIH funded projects involving intervention and data collection in preschool and elementary school settings in multiple countries and states. Co-

PI ***Dr. Kylie Flynn*** of WestEd (see CV) will serve as the National Trainer for the proposed

project. Dr. Flynn recently served as the Statewide Trainer for WestEd's i3 validation study

(Starkey, 2012) and has over 12 years of experience working on large-scale federally funded

efficacy studies, overseeing the implementation and fidelity of early educational interventions.

Dr. Starkey and ***Dr. Flynn*** will lead the development of the regional training centers and

monitor the overall implementation of the *Pre-K Mathematics* intervention, the continuous

improvement process, and the Professional Support System at kindergarten. They will also meet

regularly with ***Dr. Tipton*** in the first two years of the study to ensure that the children from our

LEA partners reflect the national sampling plan. Finally, ***Dr. Patricia Krizek***, will serve as the

Program Liaison in the national office. Dr. Krizek has over 30 years as an Early Education

Administrator and has collaborated with the WestEd team during the i3 validation study and the

continuous improvement study. ***Dr. Alice Klein*** of WestEd (see CV) will serve as PI and will

coordinate and monitor the agencies comprising the external evaluation team (see below). Dr.

Klein has led or co-led several federally funded projects with Dr. Starkey referenced above. Dr.

Klein also served for several years as PI for Elementary School Mathematics on the What Works

Clearinghouse. She will serve as Evaluation Team Coordinator and will oversee evaluation

activities and coordinate with the various agencies associated with the evaluation team.

Specifically, she will be responsible for regular meetings with SRI International in preparation of

and during each wave of data collection. She will also work closely with ETS on the scoring of

the ECLS data and with ***Dr. Thomas Cook*** and Dr. Jaime Thomas regarding data analysis.

      Independent Evaluation Team. ***Prof. Thomas Cook*** of George Washington University

(see CV) will lead the evaluation. He will be responsible for the study design, and in

collaboration with Dr. Jaime Thomas of Mathematica, will plan and conduct data analysis and

reporting of impact findings. Dr. Cook has served on several large federally funded evaluation

projects, including The Head Start Impact Study, Reading First, Advisory Committee on Head

Start, and others. Dr. Cook is an internationally esteemed quantitative methodologist and co-

author of the standard text on experimental and quasi-experimental designs. ***Dr. Elizabeth***

Tipton of Columbia University (see CV) will serve as the sampling statistician. Dr. Tipton's research focuses on the design and analysis of field experiments, with a particular focus on issues of external validity and generalizability in experiments and the use of (cluster) robust variance estimation. She has received Early Career Awards and funding from IES and NSF. *Michelle Najarian* of the *Educational Testing Service* (see CV) will score ECLS-B and ECLS-K math prior to data analyses by the project's statistician. She worked in this capacity in our i3 validation study.

      SRI International is one of the nation's premier research and evaluation organizations. The PI for the evaluation, *Dr. Erika Gaylor*, has 15 years of experience conducting research and evaluation studies in early childhood and pre-K to third grade. Her expertise includes working with policymakers, funders, and program leaders to identify desired outcomes, strategies to achieve them, and ways to collect and use high-quality data to monitor progress. SRI has conducted classroom observations and direct assessments in Head Start and state preschool classrooms for numerous national and statewide studies. For example, SRI has collected such data in multiple states as part of three i3 grants, a Race-to-the-Top Early Learning evaluation, a Preschool Development Grant evaluation, and an evaluation of Next Generation Math for PBS KIDs. Further, SRI has served for four years, as the external data collector for WestEd on its i3 validation grant, *Early Math Project*, during which SRI staff used the same measurement tools included in the proposed study. SRI will use its well-honed data collection expertise to ensure high response rates and high-quality data.

      WestEd (with the Far West and Southwest Regional Labs beginning operations in 1966) has over 50 years of experience working with LEAs and schools across the country, and a strong track record of increasing student achievement, improving academic outcomes and closing achievement gaps through its work with LEAs and schools. WestEd has extensive experience managing large, complex educational projects.

      Additional considerations. The team of PIs have strong merit based on two important factors. First, we have a strong track record of securing research and development funds through

i3, IES and other federal and state entities. We have developed and evaluated interventions through IES and other funding sources. Through this work, we have developed a replicable methodology for implementing rigorous development and research activities in actual field settings (i.e., schools and classrooms). Perhaps most important is the fact that across several projects, we have (1) developed interventions, (2) secured funding for efficacy and scale-up projects, (3) completed methodologically and pragmatically successful trials, and (4) reported our main child outcome findings in prestigious peer-reviewed journals, such as *Science, Cognition, and Early Childhood Research Quarterly*. We believe this to be a major strength in our application and worthy of additional consideration for funding.

Project Resources After the Grant's Funding Period

Multi-year financial and operating model and plan. The start-up cost of implementing the *Pre-K Mathematics* intervention in our i3 validation project was $4,200 per classroom. The start-up and recurring costs of *Pre-K Mathematics* over a 5-year period shown in the table below. Start-up costs include workshops and coaching visits for lead teachers. The only recurring costs are for production of home activities (primarily photocopying), which are given to families to keep.

Table 1. Start-Up and Recurring Costs of Pre-K Mathematics Over 5 Years of Implementation

	Year 1	Year 2	Year 3	Year 4	Year 5
Cost/Classroom	$4,200	$456	$456	$456	$456
# Children Served	24	24	24	24	24
Cost/Child/Year	$175	$19	$19	$19	$19

In the EIR Expansion project, efforts will be made to reduce these costs through the use of technology and economies of scale. Lower costs, however, will be offset somewhat by continuous improvement training, described above, which helps programs scale in (adapt) their implementation of *Pre-K Mathematics* without sacrificing its effectiveness. This makes effective implementation more sustainable because implementation is adapted in light of local program

needs. It also gives the local program a sense of ownership of the intervention. Start-up costs will be tracked through project year 3 to determine whether and how much costs were lowered.

Operation of regional training centers after the grant's funding period. Minimum staffing of a regional training center is a team of two trainers working from home offices. Using i3 project costs of $4,200 per classroom, 2 trainers would need approximately 10 work days to prepare for, travel to, and conduct complete trainings (multi-day workshops, on-site coaching visits, and a continuous improvement support) for teachers in 16 classrooms. This revenue generated by one training is calculated as follows: $4,200 X 16 classrooms = $67,200. The training team would need to conduct 24 complete trainings per year ($67,000 X 24 = $1,612,800) to generate sufficient revenue for the regional training team to be fully employed at WestEd and cover overhead, travel, training material expenses, as well as costs for program recruitment. The number of training teams deployed in a region will depend on the demand for training.

In years 4 and 5, we will begin recruitment of programs for fee-for-service math trainings, providing information about pricing for different rollout options. If programs have PD or administrative staff (e.g., Head Start site supervisors) who can be trained to coach, the start-up costs for *Pre-K Mathematics* will be reduced by approximately 40%, even given the costs of a trainers' institute. In year 5, trainings will commence and continue thereafter as dictated by demand. We will also collect data on costs of providing follow-up support to kindergarten teachers through year 4. It will then be possible to determine the amount of training that would be required in a given region for this service to be financially feasible as well.

Cost effectiveness of the innovative intervention. The WestEd team, in collaboration with SRI, will estimate cost effectiveness using the ingredients method (Levin and McEwan, 2001). We will identify program inputs through document review and implementation interviews and will then identify and cost out a master list of program ingredients using data collected, including school records, and the "Cost Out" tool (Hollands, et. al., 2015). Through analysis that compares the cost of the program to the achieved student outcomes (math achievement, reduction in need for tier 2 or other special support for math difficulties) on a per unit basis, this

evaluation can provide policy-makers with useful information upon which to make decisions regarding the allocation of resources, which in turn will affect program sustainability.

Commitment from partners. WestEd will continue to employ regional training staff commensurate with demand beyond the project period. Since WestEd is a non-profit, training will be provided at cost. Furthermore, the PD and Co-PI own the intellectual property rights to the *Pre-K Mathematics* intervention. They will continue to provide the intervention to public preschool programs at the cost of producing training and implementation manuals for teachers. Lakeshore will continue producing curriculum kits commensurate with demand beyond the project period.

Stakeholder support. State, regional, and local agencies of public preschool programs have shown great enthusiasm for the Early Math Initiative and this expansion-phase proposal. As shown in Appendix G, within a short one-month timeframe, we have obtained letters of support from 16 agencies that represent nearly 200,000 high needs preschool children. In particular, we were able to establish support in our targeted regional locales at the regional, state, and local levels from both Head Start and state preschool programs. This bodes extremely well for our capacity to establish regional training centers in these targeted areas. Specifically, in our current West region, we have support from the Region IX Head Start Association, along with local support from the Navajo Head Start in AZ, the Neighborhood House Association in San Diego, and the Fresno County Office of Education with a combined student population of more than 64,000 preschoolers. In the Mid-West region, we obtained support from three state Head Start Associations (MI, IL, and MN), along with three local programs in MN and MI, serving a combined population of approximately 40,000 preschoolers. In the Mid-Atlantic region, we received backing from the VA Head Start Association and public school districts within two VA counties, which have more than 9,000 preschoolers combined. Finally, in the Southeast, we have our largest supporter from the Region IV Head Start Association which represents more than 82,000 preschoolers. We also have support from the Miami-Dade public schools, the fifth largest

district in the country. This demonstrates broad recognition of the need for expansion of this project and our ability to recruit within the parameters of our sampling plan.

<div align="center">**Independent Project Evaluation**</div>

Meeting the What Works Clearinghouse Standards

Overview of evaluation plan. The evaluation of the proposed project will use an experimental design (RCT) and rigorous methods to ensure that the evidence of effectiveness obtained from this project meets the WWC standards without reservations. Random assignment to condition will be conducted by the independent statistician. Procedures will be implemented to prevent treatment diffusion and other potential threats to internal validity. In addition, attrition from the Intent to Treat sample of children will be carefully monitored. Multiple tracking procedures will be used to locate and assess as many children as possible on the principal outcome measures at each wave of data collection to minimize both overall and differential attrition. The overarching goal of the evaluation is to examine the impact of the *Pre-K Mathematics* intervention on measures of child math outcomes as well as on teachers' math practices. At the child level, the evaluation will measure growth in mathematical knowledge at the end of pre-K, K, and Grade 1. At the teacher level, the evaluation will examine the impact of the intervention on teachers' math practices in pre-K and K, and on their knowledge about how to teach math to pre-K children. The research questions related to the impact of the *Pre-K Mathematics* intervention are presented below in conjunction with the **Data Analysis Plan** for answering these questions.

Experimental design. We will evaluate the national impact of *Pre-K Mathematics* using a cluster randomized controlled trial (RCT) in which the pre-K classroom is the unit of random assignment. Pre-K classrooms will be located at randomly selected preschool centers or school sites in four regions of the country - West, Midwest, South, and Mid-Atlantic (see **Sampling Plan** below). Within each pre-K classroom, children will be consented for the study at the beginning of their pre-K year, and then followed longitudinally into their K and Grade 1

classrooms. All 4-year-old children who will be age-eligible to attend K the following year, who are from low-income families, and for whom parental consent to participate is obtained will be selected for the study sample, up to 20 per classroom. Special needs children will be identified insofar as feasible and included in the sample if otherwise eligible. Based on a power analysis to detect effects at the end of pre-K, K and Grade 1 and making assumptions about attrition at each of these time points (see **Statistical Power Analysis** below), the RCT study will include 210 classrooms and 3,150 children.

After consenting the child sample, pre-K classrooms at each preschool site will be randomly assigned to either the treatment condition or control condition. For all sites with an even number of pre-K classrooms, half of the classrooms will be randomly assigned to the treatment condition and half to the control condition. When a site has a single pre-K classroom, we will combine it with a nearby site into one synthetic "school" with an even number of pre-K classrooms, and then randomly assign half of the pre-K classrooms to the treatment condition and half to the control condition. This was the procedure used in the i3 Validation study throughout California, and inspection of the pretest balance revealed no imbalance among the children who remained in the design at the end of pre-K, K, or Grade 1.

The project will be conducted using a two-cohort experimental design with an equal number of preschool classrooms and children recruited from two different regions of the country for each cohort (i.e., a total of 105 classrooms, 1,575 children per cohort). Thus, implementation of the *Pre-K Mathematics* intervention will begin with the first cohort of classrooms and children in Year 2 of the project, and with the second cohort of classrooms and children in Year 3. We will follow the same recruitment and classroom random assignment procedures in both cohorts. The use of a two-cohort design to rollout the intervention on a national scale is largely pragmatic. It will enable the project to reduce costs and increase implementation quality by allocating training and professional development resources more densely within two regions at a time rather than simultaneously spreading out resources less densely across all four regions.

Some attrition is anticipated over three years of the study due to family circumstances such as relocation. However, in our recent i3 Validation study, overall attrition was 4% during the pre-K year and only 11% at the end of K (see **Power Analysis** below). Although one might expect slightly higher attrition rates in the proposed national study, multiple tracking procedures will be utilized to maintain contact with participants and keep overall and differential attrition at levels that meet the WWC group design standards without reservation. Tracking procedures include obtaining extensive parent and relative contact information at the beginning of the study and updating this information by contacting parents twice per year over each year of the project.

Sampling plan.

Populations of interest. At-risk 4-year-olds typically attend preschool in one of two program contexts: (1) Head Start programs or (2) state pre-K programs. State Pre-K programs are operated by school districts (on site within elementary schools), or in stand-alone programs run by private entities. We consider the national populations of both Head Start and State Pre-K schools in this study. To define the target population and develop a sampling plan for this study, we turned to two data sources. Data on the Head Start population of centers was made available from the Head Start Program Information Report from 2016-17. For the state pre-K population, we turned to the Common Core of Data collected annually by NCES (and used the most recent data available, 2015-16). These databases allowed us to obtain student demographics (race, gender, language) and to determine urbanicity.

Overview of the sampling plan. The sampling plan for this study builds upon both purposive and random sampling methods that have been used in the evaluation of educational interventions (Tipton et al, 2014; Tipton, 2014; Tipton and Peck, 2017; Tipton, Yeager, Schneider, & Iachan, 2018). These methods share the primary goal of representing well a target population while accounting for on-the-ground constraints and contexts that occur in recruitment. The plan we have created will allow for two stages of recruitment: (1) a first stage based on the currently recruited Head Start and state pre-K programs within four regions of the country (see **Stakeholder support** above), and (2) a second stage that is more strategic. The overall goal is

for the final sample to be similar in composition to the target population on the key demographics defined for this study: 5 categories of race/ethnicity (White, African American, Hispanic, Native American and Asian) and 3 categories of urbanicity (Urban, Suburban, and Rural). In addition, the sampling plan is designed to have adequate power to estimate both average treatment effects (ATE) and subgroup impacts.

Benefits of this approach. There are several benefits to this approach which utilizes a combination of both random and purposive sampling methods. In the ideal, of course, the entire sample would be selected probabilistically. However, since recruitment is time-intensive and costly, this approach reduces the proportion of the sample that needs to be recruited probabilistically. By randomly sampling a portion of the sample, however, we can increase the diversity of the sample, both in terms of the demographics a priori specified and those not specified. Finally, by keeping records on any "non-response" that occurs in recruitment, we will be able to better understand the types of schools that may be interested in adopting the *Pre-K Mathematics* intervention after the study is completed.

Effective Strategies and Guidelines for Replication

Data will be collected on aspects of implementation that should be followed by future efforts to replicate or extend this intervention research. The essential features of implementation include (1) the curriculum plan teachers follow, (2) the level of fidelity at which teachers implement the intervention, (3) the curriculum dosage levels delivered to children by teachers and parents, (4) use of progress monitoring (Math Mastery instrument), and (5) pre-K and K math reports linked to CCSSM. As described above (see **Data on program implementation, formative evaluation, and progress monitoring),** high quality data will be collected directly on each of the above essential features of implementation through periodic classroom observations. Local trainers will also use these data formatively during implementation to monitor the quality of implementation. For example, record-keeping systems used as part of implementation will make it apparent to a trainer that a particular teacher has begun to fall behind in the curriculum plan; trainers will have been trained to assist teachers by providing feedback, discussing why the

curriculum is being implemented slowly, and working with the teacher to solve this implementation challenge.

Performance Data and Outcomes

See **Measures and Data Collection Timeline** for the Evaluation (**Table 3, Appendix G).**

　　Child math outcomes. The principal math outcome measures for this study are the Early Childhood Longitudinal Study, Birth Cohort **(ECLS-B)** mathematics assessments for pre-K, and the Early Childhood Longitudinal Study, Kindergarten Class 2010-2011 **(ECLS-K:2011)** mathematics assessments for Kindergarten and First Grade (USDE, 1998-99; Najarian, et al., 2010). The ECLS-B will be administered in fall and spring of pre-K, and the ECLS-K:2011 will be administered in spring of K and Grade 1. The ECLS instruments are adaptive, IRT-based measures of mathematical knowledge, and they are scaled psychometrically from pre-K to First Grade to permit analyses of growth across grade levels. Items on the ECLS-B and ECLS-K:2011 measure mathematical knowledge in the content areas of number sense, operations, geometry and spatial sense, measurement, data analysis, and patterns. The ECLS measures provide scale and theta scores for conducting project analyses as well as normed scores (based on the national ECLS sample) for establishing performance targets. The reliability (internal consistency) reported for the ECLS-B preschool math assessment is .89 and reliability for the ECLS-K math assessment is .92. The ECLS instruments have been translated into Spanish, and Spanish-speaking, EL children will be assessed in Spanish and English using bilingual assessors and conceptual scoring.

　　A second math outcome measure, the Child Math Assessment–Abbreviated **(CMA-A)**, will be used to measure growth in mathematical knowledge over the pre-K year. The CMA-A measures a broad range of early mathematical knowledge and skills using manipulative materials appropriate for preschool children. It was adapted specifically for use in the PCER national evaluation of preschool curricula (PCER, 2008) from a longer, researcher-developed instrument (Klein & Starkey, 2012). The CMA-A is comprised of 5 tasks (Object Counting, Addition and Subtraction with Objects, Construction of Equivalent Sets, Shape Recognition, and Pattern

Duplication), and it is more closely aligned with the content of the *Pre-K Mathematics* intervention than the ECLS-B. Thus, it will serve as a secondary measure of the impact of the intervention on child math outcomes.

The Screener for Early Number Sense **(SENS)** – Kindergarten form measures children's knowledge of number, number relations, and number operations (Jordan, Klein, & Huang, 2018). The SENS is an IRT-based measure consisting of 30 items that are scored as correct/incorrect and yield a total score as well as percentiles. As a screening instrument, the SENS can be used to identify those children who are at-risk for mathematical learning difficulties in early elementary school. The SENS will be administered to all children in the spring of K to identify their risk status.

Child behavior questionnaire. Pre-k treatment and control teachers will complete the Child Behavior Questionnaire **(CBQ)** (Rothbart et. al., 2001, adapted for teachers by Eisenberg et al., 2004), for each child in the study sample during the winter of pre-k. The CBQ measures a child's focused attention, inhibition of behaviors, and impulsivity as rated by their classroom teacher. The CBQ provides a measure of the child's ability to attend to instructional activities in the classroom and, for this reason, it is relevant to educational interventions. Specifically, the CBQ subscales of attention focusing and inhibition have been found to be significantly related to children's math outcomes at the end of pre-K and K (Barnes, Klein et. al., 2016).

Classroom observation measures. The Early Mathematics Classroom Observation **(EMCO)** measures (1) the nature of pre-K/K teachers' mathematics practices (e.g., intentional vs. incidental support for math; use of scaffolding; use of whole group, small group, and individualized instruction), (2) the amount of time each type of grouping is used (e.g., minutes of whole group math per child), (3) the conceptual breadth of the math practices, and (4) pedagogical techniques (e.g., teacher initiation of activities, scaffolding, or extending activities upward or downward developmentally) used by teachers (**Appendix H**). The Generalized Fidelity can measure fidelity of implementation of a generic (non-specific) math curriculum, so it

can be administered in both treatment and control classrooms and can detect treatment diffusion in control classrooms if it occurs.

Pre-K pedagogical content knowledge questionnaire. Pre-K treatment and control teachers' will complete a Pedagogical Content Knowledge Questionnaire in mathematics in the fall and spring of the pre-K year. This questionnaire measures three dimensions of teachers' knowledge of mathematics: Informal and Formal Math Content, Mathematical Learning and Development Sequences, and Instructional Strategies. Items are scored as correct/incorrect, and a total correct score will be used in the analyses.

Data collection procedures. Child assessments on the ECLS measures will be conducted at four time points during the study: fall and spring of pre-K, spring of K, and spring of Grade 1. In addition, the CMA-A will be administered as a second math outcome measure in fall and spring of pre-K. Children will be assessed individually in a quiet location outside of the classroom at their school. Classroom observations of teachers' math practices will be conducted in all treatment and control classrooms during the pre-K intervention year (EMCO measure) as well as during the K math Enhancement? Support? Year (KMCO measure). Assessors and classroom observers will be blind to the condition assigned to classrooms and children.

Managing data collection. Dr. Erika Gaylor, the PI at SRI International, will have lead responsibility for carrying out the data collection plan, supervising the cost effectiveness study, and coordinating the efforts among key individuals to ensure that all tasks are implemented in an efficient and organized manner. Dr. Gaylor will work with the Operations Director at SRI to supervise the Field Managers and field assessment teams in each region where the project is being implemented to ensure high-quality data collections that are completed on time.

Scoring ECLS measures. The principal child outcome measures are the ECLS-B and ECLS-K:2011 mathematics assessments. The scored data files for the ECLS-B and ECLS-K:2011 math assessments will be sent to the Educational Testing Service (ETS) for IRT scoring and psychometric analyses (see Letter from ETS). ETS will score these data and provide scale and theta scores to the statisticians on the Evaluation team to conduct all the project analyses.

Key Project Components, Mediators, Outcomes, and Implementation Threshold

The independent evaluation will examine whether the early math intervention as implemented under realistic conditions on a national scale produces significant and positive gains in mathematics achievement for high-need students in the treatment group as compared to those in the control group. A set of confirmatory and exploratory research questions that address overall impacts of the math intervention on child outcomes, moderation of treatment effects, and impact of the math intervention on teacher practices and knowledge are detailed below.

Research Questions

RQ 1: Pre-K Effects (Confirmatory): How does the pre-kindergarten math intervention affect children's mathematics achievement (as measured by the ECLS-B) at the end of Pre-Kindergarten?

RQ 2: Kindergarten Effects (Exploratory): How does the cumulative effect of receiving a pre-kindergarten math intervention followed by a strengthened system of math instruction at kindergarten affect children's' mathematics achievement (as measured by the ECLS-K:2011) at the end of Kindergarten?

RQ 3: Grade 1 Effects (Exploratory): How does the cumulative effect of receiving a pre-kindergarten math intervention followed by a strengthened system of math instruction at kindergarten affect children's mathematics achievement (as measured by the ECLS-K:2011) at the end of Grade 1?

RQ 4: Moderation of Treatment Effects: Based on our proposed **model of causation** (see above), several variables will be examined as potential moderators of treatment effects. Specifically, are math intervention effects moderated by variables related to the child (ethnicity), the teacher (pedagogical content knowledge), program type (Head Start, public Pre-K), or geographical context (urbanicity)?

RQ 5: Effects on Teacher Math Practices at Pre-K and Kindergarten: Does implementation of the pre-kindergarten math intervention affect the duration and breadth of math-related activities that teachers engage in with children in the pre-K classroom? Does implementation of a

strengthened system of math instruction at kindergarten affect the duration and quality of math instruction that teachers provide in the K classroom?

RQ 6: Effects on Teacher Math Knowledge: Will teachers who implement the pre-kindergarten math intervention acquire more pedagogical content knowledge about math and how to teach math to young children (as measured by the Pedagogical Content Knowledge Questionnaire) at the end of pre-K?

Measurable threshold for acceptable implementation. Treatment-on-the-treated (TOT) analyses of pre-K outcomes will include treatment children who received at least 75% of their classroom math dosage over the school year, as recorded on teachers' Assessment Record Sheets in treatment classrooms (Appendix G) and as measured by attendance records in control classrooms. TOT analyses of K outcomes will include who received 75% of the classroom math dosage both in pre-K and in K.

Statistical Power Analysis

In this section, we describe the statistical power for each of the principal research questions and for each type of outcome.

Overall impacts at the child level. The evaluation will have ample statistical power to detect effects of *Pre-K Mathematics* on children at the end of pre-K, K, and Grade 1. We propose to include 210 classrooms and 3,150 children in the evaluation study. With these sample sizes, the evaluation will be able to detect overall impacts at the national level that are much smaller than we expect, given the effect sizes from previous evaluations of *Pre-K Mathematics*. However, the larger sample sizes are required for adequate power when assessing subgroup effects, such as different race/ethnic groups of children.

We calculated minimum detectable effect sizes (MDES) for the intended sample of 210 classrooms and 3,150 children under a variety of different assumptions about attrition (see **Appendix G** for the detailed formula and parameter values used to calculate MDES values). Without any attrition, the MDES is 0.110 standard deviations. With an attrition rate of 15 % the sample size would fall to180 classrooms and 2,700 children, but the MDES would still only be

0.113. To put these estimates into perspective, the most recent i3 Validation study of *Pre-K Mathematics* in California found an effect size of 0.298 at the end of Pre-K and the attrition rate was zero at the classroom-level and 5 percent at the child-level. Thus, reasonably anticipated attrition rates pose no impediment to achieving a national impact whose power exceeds the usually expected 0.20 or 0.25 level.

Subgroup impacts at the child level. The evaluation will examine several moderator variables, of which the most important are the five racial/ethnic groups of students: African American, Asian American, Latino, Native American, and White. We also propose to estimate effects for three urbanicity categories: urban, suburban, and town/rural, but power will be greater for each urbanicity category than for the smaller race/ethnic categories. Under the same assumptions that we used to calculate MDES values for overall impacts at the child level (see **Appendix G** for formula and parameter values), we calculated MDES values for individual subgroups as a function of how large they are relative to the intended total sample. For equally sized race/ethnic subgroups (20 percent of the full sample), MDESs are an acceptable 0.251.

Impacts on classroom-level teacher measures. Given the anticipation of minimal classroom attrition during the pre-K year, the proposed sample size of 210 classrooms is large enough to detect moderate-sized effects of *Pre-K Mathematics* (0.275) on classroom-level measures such as teacher math practices and teacher math knowledge (see Appendix G for detailed formula and parameter values).

Data Analysis Plan

We now describe the analytic models and measures we propose to use to answer the evaluation study's research questions. The detailed analytic models, including formulas and parameter values, for each set of analyses below are included in **Appendix G.**

Effects on children's math achievement. Research Questions 1-3 address the overall effects of *Pre-K Mathematics* on children's math achievement at the end of Pre-K, K, and Grade 1. In fall of pre-K, we will administer the ECLS-B mathematics assessment as the pretest baseline measure. In addition to the ECLS-B pretest, we will include the following covariates in

our model: Age, Gender, Race/Ethnicity, Language, and Socioeconomic status. The ECLS-B will be administered again in the spring of pre-K as the posttest outcome measure. The ECLS-K:2011 Kindergarten mathematics assessment will be administered as the outcome measure in spring of K, and the ECLS-K:2011 First Grade mathematics assessment will be administered as the outcome measure in spring of Grade 1. Using these outcome measures and covariates, we propose to answer the research questions about math effects on children nationally by using a two-level hierarchical linear model where children are nested within clusters defined by the pre-K classrooms that constitute the unit of random assignment.

Moderation of treatment effects. Research Question 4 addresses the moderation of treatment effects. We propose to use the same math outcome measures (ECLS-B, ECLS-K:2011) to examine how changes in math achievement vary with some potential moderators of treatment effects. These are child-level variables such as race/ethnicity or English language learner (ELL) status, classroom-level variables such as the pre-K teacher's pedagogical knowledge, program-level variables such as program type (Head Start or public pre-K), and geographical context (urban, suburban or town/rural). For each set of variables, we propose to conduct separate subgroup analyses to examine the moderation of treatment effects. For example, to conduct an analysis of race/ethnicity as a moderator of treatment effects, we would estimate effects for children within each race/ethnicity category separately, then conduct statistical tests to determine whether effects differed by children's race/ethnicity. To conduct these analyses, we propose to use the same equation for the two-level hierarchical linear model (see **Appendix G**), modified as needed for these subgroup analyses. For example, when examining subgroups defined by race/ethnicity, we will omit the race/ethnicity indicator from the estimating equation since it will not vary within one racial/ethnic subgroup.

Given the number of subgroups that to be examined in the moderation analyses, we can anticipate the danger of misinterpreting false positive results. To guard against this, we will adjust the nominal alpha rate using multiple comparisons adjustments such as the Bonferroni correction or Benjamini-Hochberg procedure (McDonald, 2014).

Effects on teacher math practices and knowledge. Research Questions 5 and 6
address the effect of the intervention on teachers' math practices and pedagogical content knowledge. The proposed study also seeks to examine proximal effects on teacher math practices and knowledge. In pre-K, we propose to assess teachers' math practices using the Early Mathematics Classroom Observation (EMCO). In kindergarten, we will use the Kindergarten Mathematics Classroom Observation (KMCO). To assess teachers' math knowledge, we propose to use the Pedagogical Content Knowledge questionnaire. A linear model will be used to estimate these classroom-level outcomes. These analyses will use data on some teacher-level covariates measured at baseline (for pre-K teachers) and in the fall of the kindergarten year (for kindergarten teachers). They will include: Baseline pedagogical content knowledge, Age, Gender, Race/ethnicity, Education level, and Teaching experience.

Mediation. It is hypothesized that the amount of time teachers and children participated together in small-group math activities, as measured by the EMCO, will mediate the causal influences of the *Pre-K Mathematics* intervention. The design requires a multilevel model, in which case the product of the effects is the indicator of indirect effects. In this multilevel model, the outcome is measured at level one of the model, and the experimental variable and the mediator are measured at level two. Only one model is needed to model both effects. The within-subjects portion of the model includes the child math outcome at end of pre-K, the child's pretest score on the outcome measure, and the child's age at posttest.

References

American Institutes for Research & RAND Corporation. (2016). *Independent Evaluation of California's Race to the Top–Early Learning Challenge Quality Rating and Improvement System: Year 2 Technical Report – REVISED*. Unpublished manuscript. Retrieved from: https://www.cde.ca.gov/sp/cd/rt/documents/rttelcqrisevalreport.pdf

Arizona Department of Education (2016). *Arizona Mathematics Standards: Kindergarten*. Retrieved from: https://cms.azed.gov

Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences, 20,* 427–435. doi:10.1016/j.lindif.2010.06.003.

Aunola, K., Leskinen, E., Lerkkanen, M. K., & Nurmi, J. E. (2004). Developmental dynamics of math performance from preschool to Grade 2. *Journal of educational psychology*, *96*(4), 699. http://dx.doi.org/10.1037/0022-0663.96.4.699

Bailey, D., Duncan, G.J., Odgers, C. & Wu, W. (2016). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness, 10*(1), 7-39. https://doi.org/10.1080/19345747.2016.1232459

Barnes, M. A., Klein, A., Swank, P., Starkey, P., McCandliss, B., Flynn, K., ... & Roberts, G. (2016). Effects of tutorial interventions in mathematics and attention for low-performing preschool children. *Journal of Research on Educational Effectiveness*, *9*(4), 577-606. http://dx.doi.org/10.1080/19345747.2016.1191575

Baroody, A. J. (2004). The developmental bases for early childhood number and operations standards. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education*, (pp. 173-219). Mahwah, NJ: Lawrence Erlbaum Associates.

Bisanz, J., Sherman, J.L., Rasmussen, C., & Ho, E. (2005). Development of arithmetic skills and knowledge in preschool children. In J.I.D. Campbell (Ed.) *Handbook of mathematical cognition* (pp.143-162). New York: Psychology Press.

Chard, D. J., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K., & Smolkowski, K. (2008). Preventing early mathematics difficulties: The feasibility of a rigorous kindergarten mathematics curriculum. *Learning Disability Quarterly*, *31*(1), 11-20. https://doi.org/10.2307/30035522

Clements, D. H., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies: Persistence of effects in the third year. *American Educational Research Journal, 50*(4), 812-850. doi: 10.3102/0002831212469270

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational researcher*, *32*(6), 3-12. https://doi.org/10.3102/0013189X032006003

Cohen, D. K., & Ball, D. L. (2007). Educational innovation and the problem of scale. In B. Schneider & S. McDonald (Eds.), *Scale up in education: Ideas in principle*, (vol. 1, pp. 19-36). Lanham, MD: Rowman & Littlefield.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's

emergent literacy growth. *Journal of Educational Psychology, 98*(4), 665-689.

doi:10.1037/0022-0663.98.4.665

Copley, J. V. (2010). *The young child and mathematics*. Washington, D.C.: National Association

for the Education of Young Children.

Derzon, J.H. (2018). Challenges, opportunities, and methods for large-scale evaluations.

Evaluation & the Health Professions, 41(2), 321-345.

https://doi.org/10.1177/0163278718762292

DeWire, T., McKithen, C., & Carey, R. (2017). *Scaling up Evidence-Based Practices: Strategies*

from Investing in Innovation (i3). Washington, DC: U.S. Department of Education.

Retrieved from: https://i3community.ed.gov/insights-discoveries/2207

Duncan, G.J., & Magnuson I.K. (2011). *The long reach of early childhood poverty.* Washington:

Brookings.

Duncan, G. J., & Sojourner, A. J. (2013). Can intensive early childhood intervention programs

eliminate income-based cognitive and achievement gaps?. *Journal of Human*

Resources, *48*(4), 945-968. doi: 10.3368/jhr.48.4.945

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., ... & Henry, G.

T. (2007). Teachers' education, classroom quality, and young children's academic skills:

Results from seven studies of preschool programs. *Child development*, *78*(2), 558-580.

https://doi.org/10.1111/j.1467-8624.2007.01014.x

Eisenberg, N., Spinrad, T.L., Fabes, R. A., Reiser, M., Cumberland, A., Shepard, S.
A.,Thompson, M. (2004). The relations of effortful control and impulsivity to children's
resiliency and adjustment. *Child Development, 75*, 25-46. https://doi.org/10.1111/j.1467-
8624.2004.00652.x

Entwisle, D. R., & Alexander, K. L. (1989). Early schooling as a "critical period" phenomenon.
In K. Namboodiri & R. G. Corwin (Eds.), *Research in the Sociology of Education and
Socialization*, (pp. 27–55). Greenwich, CT: JAI

Entwisle, D. R., & Alexander, K. L. (1990). Beginning school competence: Minority and
majority comparisons. *Child Development, 61,* 454-471. https://doi.org/10.1111/j.1467-
8624.1990.tb02792.x

Florida Department of Education (2017). *The Florida early learning and developmental
standards – Birth to kindergarten.* Retrieved from
.http://flbt5.floridaearlylearning.com/standards.html#d=V&a=three_year_olds,four_year_
olds

Geary, D.C. (1994). *Children's mathematical development*. Washington, D.C.: American
Psychological Association.

Gelman, R., & Williams, E.M. (1998). Enabling constraints for cognitive development and
learning: Domain specificity and epigenesis. In W. Damon (Series Ed.), D. Kuhn, & R.S.
Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2: Cognition, perception and
language* (5th ed., pp. 575-630). New York: Wiley.

Ginsburg, H. P., Klein, A., & Starkey, P. (1998). The development of children's mathematical

    thinking: Connecting research with practice. In W. Damon (Series Ed.) & I. E. Sigel & K.

    A. Renninger (Vol. Eds.), *Handbook of child psychology: Vol. 4: Child psychology in*

    practice (5th ed., pp. 401-476). New York: Wiley.

Glenn Commission. (2000). *Before it's too late: A report to the nation form the National*

    Commission on Mathematics and Science Teaching for the 21st Century. Washington,

    D.C.: U.S. Department of Education. Retrieved from: https://eric.ed.gov/?id=ED441705

Harms, T., Clifford, R. M., & Cryer, D. (2014). *Early childhood environment rating scale*.

    Teachers College Press.

Health and Human Services, Administration for Children and Families, & Office of Head Start.

    (2015). *Head Start early learning outcomes Framework: Ages birth to five*. Washington,

    DC: Administration for Children and Families Office of Head Start. Retrieved

    from: https://eclkc.ohs.acf.hhs.gov/hslc/hs/sr/approach/pdf/ohs-framework.pdf

Hollands, F. M., Hanisch-Cerda, B., Levin, H. M., Belfield, C. R., Menon, A., Shand, R., … &

    Cheng, H. (2015). *CostOut: the CBCSE cost tool kit*. Center for Benefit-Cost Studies of

    Education, Teachers College, Columbia University. Retrieved from:

    www.cbcsecosttoolkit.org

Isenberg, J. P. (2000). The state of the art in early childhood professional preparation. In D.

    Horm-Wingerd & M. Hyson (Eds.), *New teachers for a new century: the future of early*

    childhood professional preparation (pp. 17-58). Washington, DC: U.S. Department of

    Education

Jiang, Y., Granja, M., Koball, H. (2017). *Basic Facts about Low-Income Children: Children under 6 Years, 2015*. New York: National Center for Children in Poverty, Columbia University Mailman School of Public Health..Retrieved from: http://nccp.org/about/publications_1226.html

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and individual differences*, *20*(2), 82-88. https://doi.org/10.1016/j.lindif.2009.07.004

Jordan, N. C., Huttenlocher, J., & Levine, S.C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology, 28,* 644-653. doi:10.1037/0012-1649.28.4.644

Jordan, N. C. & Klein, A. (2018). *Screener for Early Number Sense* [Measurement instrument]. Unpublished instrument.

Jordan, N., Klein, A., and Huang, C.W. (2018). *Refining and extending a number sense screener for identifying children at risk for mathematical difficulties in school.* Funded by IES/U.S. Department of Education, Grant R305A150545

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General*, *123*(3), 284-296. http://dx.doi.org/10.1037/0096-3445.123.3.284

Karoly, L. A. (2014). *Validation studies for early learning and care quality rating and improvement systems: A review of the literature*. Retrieved from https://www.rand.org/pubs/working_papers/WR1051.html

Kaufman, J. H., & Tsai, T. (2018). *School Supports for Teachers' Implementation of State Standards: Findings from the American School Leader Panel.* Santa Monica, CA: RAND Corporation. Retrieved from: https://www.rand.org/pubs/research_reports/RR2318.html.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *The strands of mathematical proficiency*. *Adding it up: Helping children learn mathematics.* Washington, DC: The National Academies Press.

Klein, A., & Starkey, P. (2004b). *Scott Foresman – Addison Wesley Mathematics: Pre-K.* Glenview, IL: Pearson Scott Foresman. [2nd edition of Pre-K Mathematics]

Klein, A., & Starkey, P. (2012). Psychometric study of the Child Math Assessment. Final report on Scaling Up the Implementation of a Pre-Kindergarten Mathematics Curriculum in Public Preschool Programs (IES Grant R305K05004).

Klein, A., Starkey, P., Clements, D., Sarama, J. & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness, 1,* 155-178. https://doi.org/10.1080/19345740802114533

Klein, A., Starkey, P., & DeFlorio, L. (in press). Improving the mathematical Knowledge of At-Risk Preschool Children: Two Approaches to Intensifying Early Math Intervention. In D. C. Geary & D. B. Berch (Eds). *Informal and formal mathematics interventions: Cognitive foundations* (Vol. 5). Academic Press.

Lantolf, J. P., & Poehner, M. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. New York: Routledge.

Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2015). Sociocultural theory and second language development. In B. Va nPatten and J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 207 – 226). New York: Routledge.

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

Lobman, C., Ryan, S., & McLaughlin, J. (2005). Reconstructing teacher education to prepare qualified preschool teachers: Lessons from New Jersey. *Early Childhood Research & Practice*, *7*(2), 1-15. Retrieved from: https://eric.ed.gov/

McDonald, J.H. 2014. *Handbook of Biological Statistics* (3rd ed.). Sparky House Publishing, Baltimore, Maryland.

McDonald, S. K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, *35*(3), 15-24. Retrieved from: http://journals.sagepub.com/doi/pdf/10.3102/0013189X035003015

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities, 42*, 306-321. doi: 10.1177/0022219408331037

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind?. *Journal of Learning Disabilities*, *44*(5), 472-488. doi: 10.1177/0022219411414010

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *Mathematics achievement in the primary school years: IEA's third international mathematics and science study (TIMSS).* Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R. A.,O'Connor,K. M., … & Smith, T. A. (2000*). TIMSS 1999 international mathematics report*. Boston, MA: The International Study Center, Boston College, Lynch School of Education

National Association for the Education of Young Children. (2005). *Early childhood Mathematics: Promoting good beginnings. A joint position statement of the National Association for the Education of Young Children (NAEYC) and the National Council of Teachers of Mathematics (NCTM)*. Retrieved from www.naeyc.org/about/positions.asp

National Center for Education Statistics [NCES] (2008). *Highlights from TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth- and Eighth-Grade Students in an International Context* (NCES Number 2009001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from: http://nces.ed.gov/pubsearch/getpubcats.asp?sid=073

National Center for Education Statistics [NCES] (2011). *The condition of education 2011* (NCES 2011-033). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved from: https://nces.ed.gov/pubs2011/2011033.pdf

National Center for Education Statistics [NCES] (2016). *Highlights from TIMSS and TIMSS Advanced 2015* (NCES Number 2017002). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from: https://nces.ed.gov/timss/timss15.asp

National Council of Teachers of Mathematics (2008). Curriculum focal points: pre-k to kindergarten. A quest for coherence. *Teaching Children Mathematics, 14*(6), 361-365.

National Governors Association Center for Best Practice, & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington, D.C.: National Governors Association Center for Best Practices, Council of Chief State School Retrieved from http://www.corestandards.org/Math/

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

National Research Council. (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academies Press.

Piaget, J. (1952). *The child's conception of number*. London: Routledge & Kegan Paul Ltd.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*. Paul H Brookes Publishing.

Preschool Curriculum Evaluation Research Consortium (2008). *Effects of Preschool Curriculum Programs on School Readiness* (NCER 2008-2009). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S., & Jenkins, F. (2012). *Highlights from TIMSS 2011: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context (NCES 2013-009).* National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology, 105*, 453–464. doi:10.1037/a0031753.

QRIS National Learning Network. (2015). *QRIS State Contacts & Map.* Retrieved from http://www.qrisnetwork.org/qris-state-contacts-map

Rampey, B. D., Dion, G. S., & Donahue, P. L. (2009). *NAEP 2008 Trends in Academic Progress* (NCES 2009–479). Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Rathbun, A., & West, J. (2004). *From kindergarten through third grade: Children's beginning school experience.* National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2--4--7#

Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. Duncan & R. Murnane (Eds.) *Whither Opportunity Rising Inequality, Schools, and Children's Life Chances,* (pp. 91-115). New York: Russel Sage Foundation.

Reardon, S.F. (2013) The widening income achievement gap. *Educational Leadership*, 70(8), 10-16. Retrieved from: http://www.ascd.org/publications/educational-leadership.aspx

Rothbart, M.K., Ahadi, S.A., Hershey, K., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development, 72,* 1287-1604. https://doi.org/10.1111/1467-8624.00355

Schneider, B., & McDonald, S. K. (2007). *Scale-up in education: Issues in practice* (Vol. 2). New York: Rowman and Littlefield

Shulman, L. S. (2000). Teacher development: Roles of domain expertise and pedagogical knowledge. *Journal of Applied Developmental Psychology, 21*(1), 129-135. https://doi.org/10.1016/S0193-3973(99)00057-X

Sophian, C. (1996). *Children's numbers*. Boulder, CO: Westview Press.

Starkey, P. (1992). The early development of numerical reasoning. *Cognition*, *43*(2), 93-126. https://doi.org/10.1016/0010-0277(92)90034-F

Starkey, P. (2012). *Validation of the effectiveness of an innovative early mathematics intervention for high-needs students.* Funded U.S. Dept. of Education, Office of Innovation and Improvement, Award number: U411B120053.

Starkey, P., & Cooper, R.G. (1980). Perception of numbers by human infants. *Science*, *210*, 1033-1035. doi: 10.1126/science.7434014

Starkey, P., Flynn, K.S., & Klein, A. (2015). *Continuous improvement of a what works clearinghouse rated early mathematics intervention*. Funded by IES/U.S. Dept. of Education, Grant R305H150093

Starkey, P., & Klein, A. (1992). Economic and cultural influences on early mathematical

    development. In F. L. Parker, R. Robinson, S. Sombrano, C. Piotrowski, J. Hagen, S.

    Randolph, et al. (Eds.), *New directions in child and family research: Shaping Head Start*

    in the 90s (pp. 440–443). New York: National Council of Jewish Women.

Starkey, P., & Klein, A. (2000). Fostering parental support for children's mathematical

    development: An intervention with Head Start families. *Early Education and*

    Development, 11, 659-680. https://doi.org/10.1207/s15566935eed1105_7

Starkey, P., & Klein, A. (2003). *Supporting young children's readiness for school mathematics*

    through a pre-kindergarten classroom and family math curriculum. Final Report:

    OERI/US Department of Education, Grant R307F60024.

Starkey, P., & Klein, A. (2008). Sociocultural influences on young children's mathematical

    knowledge. In O. N. Saracho & B. Spodek (Eds.), *Contemporary perspectives on*

    mathematics in early childhood education. (pp. 253-276). Charlotte, NC: Information

    Age Publishing.

Starkey, P., Klein, A., & DeFlorio, L. (2014). Promoting math readiness through a sustainable

    prekindergarten mathematics intervention. In: Boivin M, Bierman KL (Eds.), *Promoting*

    school readiness and early learning: the implications of developmental research for

    practice (pp. 187-210). New York, NY: The Guilford Press

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical

    knowledge through a pre-kindergarten mathematics intervention. *Early Childhood*

    Research Quarterly, 19(1), 99-120. https://doi.org/10.1016/j.ecresq.2004.01.002

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants.

Cognition, 36, 97-127. https://doi.org/10.1016/0010-0277(90)90001-Z

Super, C.M., & Harkness, S. (1996). The cultural structuring of child development. In J.W.

Berry, P.R. Dasen, & T.S. Saraswathi (Eds.), *Handbook of cross-cultural psychology*

Vol.2: Basic processes and human development (2nd ed., pp. 1-39). Boston: Allyn &

Bacon.

Thomas, J., Cook, T. D., Klein, A., Starkey, P., & DeFlorio, L. (in press). The sequential scale-

up of an evidence-based intervention: A case study. *Evaluation Research.*

Thornton, J. S., Crim, C. L., & Hawkins, J. (2009). The impact of an ongoing professional

development program on prekindergarten teachers' mathematics practices. *Journal of*

early childhood teacher Education, *30*(2), 150-161.

https://doi.org/10.1080/10901020902885745

Tipton, E. (2014). Stratified sampling using cluster analysis: A sample selection strategy for

improved generalizations from experiments. *Evaluation Review, 37*(2), 109-139. doi:

10.1177/0193841X13516324

Tipton, E., Hedges, L.V., Vaden-Kiernan, M., Borman, G.D., Sullivan, K. & Caverly, S. (2014).

Sample selection in randomized experiments: A new method using propensity score

stratified sampling. *Journal of Research on Educational Effectiveness, 7*(1), 114-135.

https://doi.org/10.1080/19345747.2013.831154

Tipton, E., & Peck, L. R. (2017). A design-based approach to improve external validity in

    welfare policy evaluations. *Evaluation review*, *41*(4), 326-356. doi:

    10.1177/0193841X16655656

Tipton, E., Yeager, D., Schneider, B., & Iachan, R. (2018). Designing probability samples to

    identify sources of treatment effect heterogeneity. Chapter in *Experimental Methods in*

    Survey Research: Techniques that Combine Random Sampling with Random Assignment

    (Editor: Lavrakas, P.J.) New York, NY: Wiley.

U.S. Department of Health and Human Services, Administration for Children and Families,

    (2010). *Head Start impact study. Final report.* Washington, DC: U.S. Department of

    Health and Human Services. Retrieved from:

    https://www.acf.hhs.gov/sites/default/files/opre/hs_impact_study_final.pdf

U.S. Department of Labor, Bureau of Labor Statistics. (2000). *The outlook for college graduates,*

    1998–2008, 2000. In Getting ready pays off!, US DOE, October 2000, and BLS,

    Occupational employment projections to 2008, in NAB, Workforce Economics 6(1).

Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

West, J., Denton, K., & Germino-Hausken, E. (2000). *America's kindergartners: Findings from*

    the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99, Fall 1998.

    NCES 2000-070. Washington, DC: National Center for Education Statistics

What Works Clearinghouse (2014). *Procedures and standards handbook version 3.0.*

Washington, DC: Institute of Education Sciences. Retrieved from

https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_

handbook.pdf