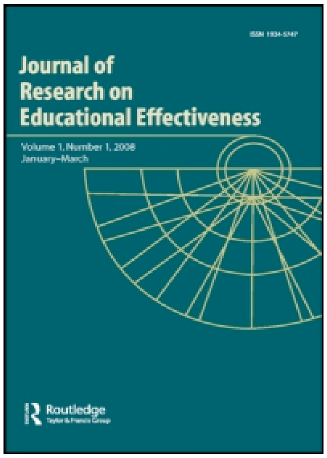

Journal of Research on Educational Effectiveness

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t787187010

Scaling Up the Implementation of a Pre-Kindergarten Mathematics Curriculum: Teaching for Understanding With Trajectories and Technologies

Julie Sarama [a]; Douglas H. Clements [a]; Prentice Starkey [b]; Alice Klein [b]; Ann Wakeley [c]
[a] University at Buffalo, State University of New York, Buffalo, New York, USA
[b] University of California, Berkeley, California, USA
[c] California Institute on Human Services, Sonoma State University Institute of Human Development, UC Berkeley, Berkeley, California, USA

To cite this Article: Sarama, Julie, Clements, Douglas H., Starkey, Prentice, Klein, Alice and Wakeley, Ann (2008) 'Scaling Up the Implementation of a Pre-Kindergarten Mathematics Curriculum: Teaching for Understanding With Trajectories and Technologies', Journal of Research on Educational Effectiveness, 1:2, 89 — 119

To link to this article: DOI: 10.1080/19345740801941332
URL: http://dx.doi.org/10.1080/19345740801941332

Journal of Research on Educational Effectiveness, 1: 89–119, 2008
Copyright © Taylor & Francis Group, LLC
ISSN: 1934-5747 print / 1934-5739 online
DOI: 10.1080/19345740801941332

Routledge
Taylor & Francis Group

INTERVENTION, EVALUATION, AND POLICY STUDIES

Scaling Up the Implementation of a Pre-Kindergarten Mathematics Curriculum: Teaching for Understanding With Trajectories and Technologies

Julie Sarama and Douglas H. Clements
University at Buffalo, State University of New York, Buffalo, New York, USA

Prentice Starkey and Alice Klein
University of California, Berkeley, California, USA

Ann Wakeley
California Institute on Human Services, Sonoma State University
Institute of Human Development, UC Berkeley, Berkeley, California, USA

Abstract: This study used a randomized field trial design to evaluate the efficacy of a research-based model for scaling up an intervention focused on preschool mathematics. Although the successes of research-based educational practices have been documented, equally well known is the paucity of successful efforts to bring these practices to scale. The same research corpus provides guidelines to scale up successful interventions. We designed an intervention model based on that research, including mathematics curricula with an emphasis on teaching for understanding following developmental guidelines, or learning trajectories, and using technology at multiple levels. We then implemented that model and evaluated the implementation with a limited scale up study. Within a design involving 25 classrooms serving children at risk for later school failure, we examined the impact of the model, using measures of fidelity of implementation, classroom observations of mathematics environment and teaching, and child outcomes. High levels of fidelity of implementation resulted in consistently higher scores in the intervention, compared to control, classes on the observation instrument and significantly and substantially greater gains in children's mathematics achievement in the intervention, compared to the control, children (effect size = .62).

Keywords: Preschool, mathematics, scale-up, randomized, curriculum

Address correspondence to Julie Sarama, Baldy Hall, University at Buffalo, State University of New York, Buffalo, NY 14260, USA. E-mail: jsarama@buffalo.edu

There may be a no more challenging issue than that of effectively scaling up an educational intervention with the diverse population that teaches pre-Kindergarten (Pre-K) and the diversity of program structures in the early childhood system in the United States. We conducted a limited scale up of the implementation of an integrated, research-based, Pre-K mathematics curriculum, with an emphasis on teaching for understanding following developmental guidelines, or learning trajectories, and using technology at multiple levels. Using a randomized field trials design, we examined the impact of professional development and fidelity of implementation on and child outcomes within an experimental design that included two of the largest types of Pre-K programs for low-income children, Head Start and state-funded Pre-K programs, in two geographically distant states, New York and California.

BACKGROUND AND NEED

Our work springs from the confluence of two educational needs, in mathematics education and in early childhood education. All citizens need a broad range of basic mathematical understanding, and careers require an increasing level of proficiency (Campbell & Silver, 1999; Glenn Commission, 2000; Kilpatrick, Swafford, & Findell, 2001). However, U.S. proficiency is well below what is desired (A. Ginsburg, Cooke, Leinwand, Noell, & Pollock, 2005; Kilpatrick et al., 2001; Mullis et al., 2000). Moreover, children who live in poverty and who are members of linguistic and ethnic minority groups demonstrate significantly lower levels of achievement (Bowman, Donovan, & Burns, 2001; Campbell & Silver, 1999; Denton & West, 2002; Mullis et al., 2000; Natriello, McDill, & Pallas, 1990; Secada, 1992; Starkey & Klein, 1992). These achievement differences have origins in the earliest years—low-income children have been found to possess less extensive mathematical knowledge than middle-income children of Pre-K and kindergarten age (Denton & West, 2002; H. P. Ginsburg & Russell, 1981; Griffin, Case, & Capodilupo, 1995; Jordan, Huttenlocher, & Levine, 1992; Klein & Starkey, 2004; Saxe, Guberman, & Gearhart, 1987). This gap encompasses several aspects of informal mathematical knowledge: numerical, arithmetic, spatial/geometric, patterning, and measurement knowledge (Klein & Starkey, 2004). The probable reason for this gap is that many children from low-income families receive less support for mathematical development in their home and school environments (Blevins-Knabe & Musun-Miller, 1996; Bryant, Burchinal, Lau, & Sparling, 1994; Farran, Silveri, & Culp, 1991; Holloway, Rambaud, Fuller, & Eggers-Pierola, 1995; Saxe et al., 1987; Starkey et al., 1999; Zill et al., 2001).

High-quality early mathematical interventions can help young children develop a foundation of informal mathematics knowledge (Clements, 1984; Klein & Starkey, 2004), including those living in poverty, and those with special needs (Campbell & Silver, 1999; Fuson, Smith, & Lo Cicero, 1997; Griffin, 2004;

Griffin et al., 1995; Klein & Starkey, 2004; Ramey & Ramey, 1998). Unfortunately, most American children are not in high-quality programs (Hinkle, 2000), and successful programs are difficult to replicate. We need widespread efforts that can be scaled up (Campbell & Silver, 1999). We conducted a rigorous experimental test of implementing a research-based, comprehensive mathematics intervention across the diverse settings of Pre-K education to evaluate the success of the model in a limited scale-up context and to begin to identify the critical variables related to the success of scaling up this implementation.

THEORETICAL FRAMEWORKS

Our TRIAD intervention (*T*echnology-enhanced, *R*esearch-based, *I*nstruction, *A*ssessment, and professional *D*evelopment) follows guidelines for scaling up based on research. The approach goes beyond adopting new curricula, the most common, but often unsuccessful, external intervention, instead employing the efficacious strategy of supporting "interactions among teachers and children around educational material" (Ball & Cohen, 1999, p. 3). This strategy creates opportunities for teachers to focus on mathematics, goals, and children's work and its improvement, which improves teachers' knowledge of subject matter, teaching, and learning (D. K. Cohen, 1996, p. 98) and increases child achievement (Ball & Cohen, 1999; Schoen, Cebulla, Finn, & Fi, 2003). The TRIAD model promotes specific roles for three categories of participants. We begin with a brief discussion of each.

Teachers

Research suggests that the most important feature of an effective educational environment is a knowledgeable and responsive adult (Bowman et al., 2001; Darling-Hammond, 1997; Ferguson, 1991) and that high-quality professional development is essential to build these competencies (Sarama & DiBiase, 2004; Schoen et al., 2003). Scaling up such professional development for early childhood teachers presents special challenges. Even graduates of 4-year early childhood programs with state licensure usually lack adequate preparation in mathematics, and those with less education have virtually no preparation (Clements, 2004; Sarama & DiBiase, 2004). Early childhood teachers are often uncomfortable with mathematics (Copley, 1999), have a narrow view of content and little knowledge of the development of these ideas and do not use research-based mathematics curricula (Copley, 2004; Sarama & DiBiase, 2004). To be effective, professional development must eschew "one-shot" interventions; weave together mathematics content, pedagogy, and knowledge of child development and family relationships (Baroody, 1998); and emphasize preparation to teach a specific curriculum (Schoen et al., 2003). Finally, a promising path to providing a core for developing teachers' understanding of learning, teaching, curriculum and assessment focuses on research-based models of children's thinking and learning (Bredekamp, 2004; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Hiebert, 1999). We put research-based *learning trajectories* at the core of our teacher/child/curriculum triad (Clements, 2002b; Clements & Sarama, 2002, 2004; Klein, Starkey, & Ramirez, 2002) because they help teachers focus on the "conceptual storyline" of reform curriculum (Heck, Weiss, Boyd, & Howard, 2002; Weiss, 2002).

Administrators

Administrative support is also important in scaling up interventions (Bodilly, 1998; Fullan, 1992; Heck et al., 2002; Kaser, Bourexis, Loucks-Horsley, & Raizen, 1999). For example, principal leadership is strongly correlated with levels of implementation (Berends, Kirby, Naftel, & McKelvey, 2001; Fullan, 1992; Jacobson & Battaglia, 2001).

Children, Parents, and Communities

Interventions that involve parents fully are more effective (Ramey & Ramey, 1998). Unfortunately, most parents also have a limited view of the breadth of mathematics appropriate for young children (Sarama & DiBiase, 2004). Low-income parents, compared to middle-income parents, believe that mathematics education is the responsibility of the preschool (Starkey et al., 1999). Specific interventions are warranted.

RESEARCH-BASED GUIDELINES FOR SCALING UP

The wider social context in which these groups are embedded, the American educational system, has a persistent "grammar of schooling" (Tyack & Tobin, 1992) composed of rigid cultural beliefs about proper teaching, learning, and knowledge. A historical analysis of innovations reveals that those challenging this grammar tend to be short-lived (Tyack & Tobin, 1992). Especially because we were attempting change close to the "core" of educational practice (Elmore, 1996a), we based our innovation on suggestions and cautions from the literature. TRIAD follows the 10 guidelines we abstracted.

1. *Involve, and promote communication among, key groups concerned with young children* (Huberman, 1992; Kaser et al., 1999), emphasizing a shared understanding of, and connections between, the project's goals, national and state standards, and greater societal need (Elmore, 1996a; Fullan, 2000; Sarama, Clements, & Henry, 1998). Promote clarity of these goals, of leadership, and of all participants' responsibilities and accountability

(Kaser et al., 1999). Begin creating (and eventually institutionalize) a support infrastructure (Kaser et al., 1999).

2. *Promote equity* through equitable recruitment and selection of participants, allocation of resources, and use of curriculum and instructional strategies that have demonstrated success with minorities and females (Kaser et al., 1999).
3. *Plan for the long term,* encouraging active participation with a "start small and build" strategy (Fullan, 1992).
4. *Focus on instructional change that promotes depth and quality of children's thinking* (Ball & Cohen, 1999; Fullan, 2000), *placing standards- and research-based learning trajectories at the core* of the teacher/child/curriculum triad to ensure that curriculum, materials, instructional strategies, and assessments are aligned with (a) national and state standards and a vision of high-quality mathematics education, (b) each other, and (c) "best practice" as determined by research and the wisdom of expert practice (Bodilly, 1998; Clements, 2002b; Clements & Sarama, 2004; Kaser et al., 1999).
5. *Build expectation and camaraderie to support a consensus around adaptation.* Reform strategies often perversely isolate teachers, either gathering them together in isolation or nurturing "special" teachers in a building (Elmore, 1996a). Instead, (a) promote "buy-in" in multiple ways, such as dealing with all participants as equal partners, and distributing resources to support the project (Berends et al., 2001); (b) establish and maintain cohort groups (Jacobson, Emihovich, Petrie, Helfrich, & Stevenson, 1998); (c) facilitate teachers visiting successful implementation sites and talking with other teachers there; and (d) build local leadership by involving principals and encouraging teachers to become teacher leaders (Fullan, 1992, 2000).
6. *Provide professional development that is multifaceted; extensive; ongoing; reflective; focused on common actions and problems of practice, and especially on children's thinking; grounded in particular curriculum material; and situated partially in the classroom.* Encourage sharing, risk taking, and learning from and with peers (Bodilly, 1998; D. K. Cohen, 1996; Fullan, 1992; Kaser et al., 1999). Develop teachers' knowledge and beliefs that the curriculum is appropriate and its goals are valued and attainable (Elmore, 1996a), keeping all professional development activities targeted toward those goals (Fullan, 2000).
7. *Give latitude for adaptation to teachers and schools, but maintain follow-through and integrity* (Fullan, 1992, 2000; Huberman, 1992). Emphasize the similarities of the curriculum with sound early childhood practice and what teachers already are doing. Do not allow dilution due to uncoordinated innovations (Fullan, 2000; Sarama et al., 1998).
8. *Maintain frequent, repeated communication and follow-through efforts emphasizing the purpose, expectations, and visions of the project* (Fullan, 1992; Kaser et al., 1999). Conduct formative and summative evaluations

connected to specific benchmarks for expected outcomes and methods for improvement (Huberman, 1992; Kaser et al., 1999).

9. *Give teachers continuous feedback from sources they trust that children are learning what they are taught and that these learnings* are *valued* (Bodilly, 1998; Elmore, 1996a; Sarama et al., 1998).
10. *Provide incentives for all participants, including intrinsic and extrinsic motivators linked to project work* (e.g., external expectations—from standards listed in the first guideline to pressures from administration; Berends et al., 2001; D. K. Cohen, 1996; Darling-Hammond, 1996; Elmore, 1996b; Jacobson & Battaglia, 2001; Mohrman & Lawler III, 1996).

TRIAD includes collaboration of key groups to establish and maintain (a) a Pre-K mathematics curriculum, with all components of the curriculum—a teacher's manual, demonstration videotapes, manipulatives, software, teaching strategies, assessments, and professional development—based on a common core of understanding the learning trajectories through which children develop mathematically; (b) professional development for teachers; (c) on-site support for teacher by facilitators during the school year; and (d) supportive roles and materials for parents. TRIAD's collaboration with key groups is dedicated to developing organizational structures that intensify and focus, rather than dissipate and scatter, teachers' motivation to engage in and maintain this challenging practice (Elmore, 1996a).

METHODS

Participants

Pre-K contexts included public preschool and Head Start classrooms in New York and California. Educational organizations that had expressed interest in participating in the research were invited to submit names of interested schools and teachers. From this list, 16 classrooms were randomly selected in New York and 10 in California; within each site, classrooms were publicly, randomly assigned to TRIAD or control groups.

These programs serve an ethnically diverse population of low-income families: African American and Latino families compose the largest group (60% in New York programs; 80% in California), with Asian American, Caucasian, and Middle-Eastern families composing the remainder. In the New York programs, 99% (Head Start) and 74% (state funded) of the children received reduced or free lunch; 100% did so in the California programs. From each classroom, 8 kindergarten-intending children were randomly selected; one control teacher moved immediately after the start of the school year, resulting in 25 classrooms (13 experimental, 12 control), including 25 teachers, and 209 children (average age = 4.3 years in October in both treatments; 45% male), that participated in all aspects of data collection. Average class size was 16.8 in the TRIAD classes and 17.5 in the control classes.

Intervention

Following the research-based guidelines previously listed, the TRIAD model emphasizes professional development and curriculum implementation.

Professional development. The TRIAD model provides multiple forms of support for teachers. The main two implemented in this study were the sequence of professional development sessions and in-class coaching.

After the start of the school year, teachers began participating in the professional development sessions. The initial sessions included a brief description of the study and an overview of the integrated curriculum and its goals. From that point forward, most sessions included a short introduction to the learning trajectory for a topic and hands-on experience and interactions with peers and staff around the curriculum's activities. We emphasized using materials as developers intended rather than on teachers adapting them, an approach supported by research (Weiss, 2002). These experiences included basic software skills (Ely, 1990).

Professional development sessions included 1 hr of distance education video meetings including all project staff and all teachers. In a typical session, one of the co-principle investigators presented upcoming mathematics content and research on the teaching and learning of that content, the group discussed questions and issues arising from the presentation, and teachers presented and discussed their own classroom experiences with the curriculum. This was followed by 2 hrs of hands-on experience with the curriculum for children, in which teachers rotated through practicing small group, computer, whole group, and "math throughout the year" activities. These experiences emphasized the understanding and use of learning trajectories, assessment with the small-group record sheets, and questions from teachers.

Some sessions featured videos of best practice that helped make explicit how such practice exemplifies research-based principles and emphasized that teachers considered exemplary continue to struggle. In doing this it illustrates that high-quality teaching for understanding is both rewarding and challenging and everyone can continue to improve and contribute to the profession (Confrey, Bell, & Carrejo, 2002; Heck et al., 2002; Weiss, 2002). Such "virtual visits," actual visits, and videos were designed to communicate the vision of the curriculum in action and to make the ideas and processes accessible, memorable, engaging, and therefore usable. This type of intense, focused work with video examples, interactions with peers around the curriculum's activities, and support throughout the year was designed to positively change teachers' perceptions, because "most teachers report that they believe their children are capable of fine work, but what they think they know from daily experience often hedges that belief with limited expectations" (Ball & Cohen, 1999, p. 8).

Coaching, the second main component, was similarly conducted throughout the year. Teachers experienced in implementing the curriculum were recruited from participating programs and were trained in a 1½-day institute in California or New York. Coaches visited teachers on the average of once every month (more if a particular teachers were having any difficulties), focusing on (a) encouraging participation and communication; (b) one-to-one consultation, planning, reflection, and reinforcement (Bresler & Walker, 1990; Sarama et al., 1998); and (c) monitoring—reminding teachers that their attention is required for the program, that the project is a priority, that a commitment has been made to it, and that somebody cares about them (Hord, Rutherford, Huling-Austin, & Hall, 1987).

Curriculum implementation. Following pretesting, TRIAD teachers began implementing the mathematics curriculum, consisting of two components, one from the Building Blocks (BB; Clements & Sarama, 2003, 2007a) project and the *Pre-K Mathematics Curriculum* (PMC; Klein et al., 2002). BB is a mathematics curriculum designed to help children extend and mathematize their everyday activities, from blocks to art to songs to puzzles. Teachers in this study used mainly the software component and correlated off-computer activities, using on average 33 different computer activities (about 5-10 min each), repeating 10 activities at least once. Each software activities was demonstrated and discussed by the teacher, then used individually by children (with support from adults available). These activities were based on research that provided guidelines for creation of software that is accessible, motivating, and understandable to young children (e.g., each activity is introduced via talking characters that introduce how to perform tasks; Clements, 2002a). The software's management system leads each child through *learning trajectories* for particular topics (English or Spanish), giving more difficult tasks as each individual child learns. Teachers conducted 30 small-group mathematics sessions from the PMC twice per week for about 20 min per session per group of approximately four to six children and sent home 19 (of 21) home activities (English and Spanish). Each activity describes mathematical terms; requisite materials; scripts for teaching, including suggestions for scaffolding; and a sheet for teachers to record authentic assessments.

We integrated these components into a coherent program that follows research-based learning trajectories. As an illustration, the learning trajectory for counting specifies a developmental progression of levels of competence, including Pre-counter, Reciter (competent verbal counting, to 5 and later to 10), Corresponder (maintains a one-to-one correspondence between enunciation of counting words and objects), Counter—Small Numbers (counts sets of objects to 5, with cardinal understanding), Producer—Small Numbers (counts out a set of objects to 5), Counter and Producer—10+ (counts and counts out sets to

10 and beyond, keeps track of items counted even in unorganized arrays), and so forth. Activities from the two curricula were aligned with these levels, as well as similar developmental progressions for the other topics, and sequenced throughout the year. For example, verbal counting activities for the Reciter level included Count from BB, a software program in which children click the mouse and count along with the software as objects are produced. An activity to develop the Corresponder level was Count and Move from BB, in which children count as they clap, stamp, or jump. An activity for the Counter—Small Numbers was "Watch Me Count" from PMC, a small-group activity in which children count objects on counting strips. Producer—Small Numbers is similarly taught in PMC's "Watch Me Make a Set" and BB's software activities in which children put a specific number of toppings on a cookie. Counting with larger numbers are taught in similar activities, such as a PMC home activity, "Help Kitty Count" and BB's "Dinosaur Shop," a software activity in which children fill orders for toy dinosaurs.

Measures and Procedures

Three categories of activities were conducted beyond the procurement and random assignment of classrooms and implementation of the intervention (which were described previously): observations of teachers for fidelity and quality of the classroom environment and teaching, assessment of children, and collection of questionnaire data from teachers and parents.

Children's mathematical knowledge. Child outcomes were measured with the Research-based Early Mathematics Assessment (REMA), which uses an individual interview format, with explicit protocol, coding, and scoring procedures. It assesses children's thinking and learning along research-based developmental progressions within areas of mathematics considered significant for preschoolers, as determined by a consensus of participants in a national conference on early childhood mathematics (Clements & Conference Working Group, 2004), rather than mirroring objectives or activities from any curriculum or state. As listed in Table 1, topics in number include verbal counting, number recognition and subitizing, object counting and counting strategies, number comparison and sequencing, number composition and decomposition, and adding and subtracting; geometry topics include shape identification, shape composition and decomposition, congruence, construction of shapes, and turns; and finally there are items on measurement and patterning. Content validity was assessed via expert panel review; concurrent validity was established with a .86 correlation with another instrument (Klein, Starkey, & Wakeley, 2000). All sessions are videotaped, and the tapes and recording forms, including accuracy for each item and children's solution strategies and error types for relevant items, are coded by trained coders. Any discrepancies are resolved via consultation with

Table 1. Means and standard deviations for *Research-based Early Mathematics Assessment*

	Control		TRIAD	
	Pre	Post	Pre	Post
T-scores				
M total	44.13	51.61	42.70	55.30
SD	7.04	8.05	6.53	7.87
Raw subtest scores				
Number				
Verbal counting	1.00	1.24	0.89	1.36
Recognition of number, subitizing	2.85	3.74	2.34	3.61
Object counting, strategies	9.36	12.89	7.19	14.56
Comparing number and sequencing	9.31	12.85	6.48	13.62
Composition of number	0.37	0.49	.98	2.31
Arithmetic	0.50	1.39	0.45	2.15
Geometry, measurement, and patterning				
Shape				
Shape identification	77.21	78.74	76.62	85.00
Comparing shape	4.22	4.87	4.22	4.73
Representing shape	0.91	1.63	0.99	2.07
Composing shape	2.37	3.28	2.03	4.77
Transformations	0.22	0.24	0.24	0.29
Measurement of length	0.80	0.84	0.68	0.81
Patterning	2.11	2.98	2.30	3.65

the senior researchers. The assessment was refined in three pilot tests (Clements & Sarama, 2007b) and a Rasch model analysis computed, yielding a reliability of .94 (Clements, Sarama, & Liu, in press). All items are ordered by Rasch item difficulty; children stop the assessment after four consecutive errors. For our study, Rasch scores for the total instrument were computed on correctness scores and logits transformed to T-scores ($M = 50$, $SD = 10$) for ease of interpretation. These T-scores were used for all statistical analyses. In addition, the sum of raw scores ($1 =$ correct, $0 =$ incorrect) was computed for each mathematical topic for descriptive purposes.

Assessors for this study were doctoral students in educational or developmental psychology or other qualified staff hired for this purpose. During the previous spring and summer, they were trained in administration. They first read and discussed the instrument, then administered it to children (outside of the participants) until they achieved 100% accuracy on following the protocol. Assessors remained naïve to the treatment group of the study's participants throughout the project. The following fall, they pretested all selected children and administered the same instrument as a posttest at the end of the school year.

Child mathematics outcomes were coded and then scored by trained teams (including some assessors, but mostly separate individuals) naïve to the children's treatment group assignment.

Classroom teaching practices and environment. The neglect of valid measures of implementation fidelity is one of the most important deficits in the field of scaling up educational innovations (Borman, Hewes, Overman, & Brown, 2003). Two observational instruments, Fidelity of Implementation (Fidelity) and Classroom Observation of Early Mathematics—Environment and Teaching (COEMET), were created based on research on the characteristics and teaching strategies of effective teachers of early childhood mathematics (Clarke & Clarke, 2004; Clements & Conference Working Group, 2004; Fraivillig, Murphy, & Fuson, 1999; Galván Carlan, 2000; Galván Carlan & Copley, 2000; Horizon Research Inc., 2001; National Association for the Education of Young Children, 1991; Teaching Strategies Inc., 2001). In particular, this research corpus aided in identifying practices used by teachers with high-quality implementations of innovative curricula to meet research-based standards. Each item is connected to one or more of these studies. An example of a Likert item shared by both instruments in the section Mathematical Focus is "The teacher began by engaging and focusing children's *mathematical thinking* (i.e., directed children's attention to, or invited them to consider, a mathematical question, problem, or idea)." Also shared by both instruments in the section for an interactive mathematics activity titled "Organization, Teaching Approaches, Interactions" are items with the subheadings Expectations, Eliciting Children's Solution Methods, Supporting Children's Conceptual Understanding.

The Fidelity instrument evaluates the degree to which teachers are teaching the specific intervention curriculum. There were 61 items, all but 6 of which were generated by 4-point Likert scales from 1 (*strongly disagree*) to 4 (*strongly agre*). All instruments in this study were submitted to the Rasch model, with scores converted to T-scores. The Rasch T for the Fidelity instrument score includes all 55 Likert items and 6 additional variables (number of adults in the room; number of whole group activities; and duration of each of whole group, small group, computer, and center activities). An example of an item unique to the Fidelity measure in the Organization, Teaching Approaches, Interactions section is "The teacher conducted the activity *as written* in the curriculum, or made positive adaptations to it (*not* changes that violated the spirit of the core mathematical activity)." Further, as shown in Table 2, the Fidelity instrument includes specific sections for each component of the curriculum: whole group, small group, computer, center, "math throughout the year," and family activities. Project staff administered the Fidelity instrument in the fall, winter, and spring.

The Classroom Observation instrument measures the quality of the mathematics environment and activities with a full-day observation and is not connected to any curriculum. Thus, it also allows for intervention-control treatment contrasts, no matter what the source of the enacted curriculum. An example

Table 2. Means and standards deviations for the fidelity of implementation measure

	Observation*			Mean
	1	2	3	3 obs
MT-score	45.6	52.2	48.7	48.8
SD	12.6	9.7	12.0	8.9
General surriculum				
Schedule	3.4	3.5	2.8	3.2
Teacher within 2 weeks of schedule[a]	3.4	3.5	2.8	3.2
Family involvement	3.2	3.8	2.8	3.3
Activities were sent home[a]	3.2	3.8	2.8	3.3
Everyday activities	3.1	3.2	3.2	3.1
Materials were present	2.9	3.1	3.3	3.1
Teacher uses curriculum's every day	3.3	3.2	3.1	3.2
Extensions	2.7	2.7	2.6	2.7
Teacher extended activities	2.7	2.7	2.6	2.7
Whole group activity				
Mathematical focus	5.1	6.1	5.0	5.4
Teacher displayed understanding of concepts	5.1	6.1	5.0	5.4
Organization, teaching, approaches, interactions	4.8	6.0	4.4	5.1
Materials set up correctly	5.6	6.8	4.6	5.7
Teacher began by focusing thinking	4.0	6.0	4.5	4.8
Pace was appropriate	5.0	6.1	4.5	5.2
Teacher conducted activity as written	4.8	5.5	4.5	4.9
Management strategies enhanced quality	5.2	5.9	4.7	5.3
Discussion	4.1	5.7	3.8	4.5
Activity involved discussion	4.1	5.7	3.8	4.5
Small-group activity				
Mathematical focus	3.3	3.4	3.8	3.5
Teacher displayed understanding of concepts	3.3	3.4	3.8	3.5
Organization, teaching, approaches, interactions	3.1	2.9	3.3	3.1
Materials were set up correctly[a]	3.3	3.3	4.0	3.5
Teacher conducted activity as written	3.5	3.3	3.4	3.4
Pace was appropriate	3.6	3.4	3.7	3.6
Activity was completed with all children[a]	3.8	2.8	3.5	3.4
Management strategies high quality	3.6	3.0	3.2	3.3
Expectations	2.9	2.8	3.4	3.0
Teacher promoted effort, persistence	3.4	3.1	3.5	3.3
Teacher encouraged active reflection	2.3	2.4	3.2	2.6
Eliciting children's solution methods	2.5	2.5	2.7	2.5
Teacher asked children to share, justify	2.6	2.4	2.5	2.5
Teacher facilitated children's responding	2.6	3.0	3.0	2.9
Teacher encouraged children's listening	2.3	2.0	2.5	2.3
Supporting children's conceptual understanding	2.4	2.6	2.8	2.6
Supported describer's thinking	2.4	2.7	3.3	2.8
Supported listener's understanding	1.9	2.1	2.2	2.1
Gave just enough assistance	2.9	2.9	2.8	2.9
Extending children's mathematical thinking	2.1	2.3	2.9	2.4
Elaborated children's mathematical ideas	2.0	2.1	3.0	2.4
Went beyond initial solutions	1.6	1.9	2.2	1.9
Encouraged mathematical reflection	2.4	2.6	3.2	2.7
Cultivated love of challenge	2.6	2.6	3.2	2.8

Table 2. Means and standards deviations for the fidelity of implementation measure (Continued)

	Observation*			Mean
	1	2	3	3 obs
Assessment and instructional adjustment	3.2	3.2	3.2	3.2
Listened to children, taking notes	3.8	3.7	3.8	3.8
Adapted tasks to ability and development	2.7	3.0	3.0	2.9
Used scaffolding activities	3.5	3.3	3.6	3.5
Used upward and downward extensions[a]	2.8	2.8	2.6	2.7
Center activity				
Organization, teaching, approaches, interactions	3.5	2.9	3.6	3.3
Tasks engaged children	3.6	3.2	3.9	3.6
Task selected by child[a]	3.4	1.7	2.2	2.4
Materials set up correctly[a]	4.0	2.9	4.2	3.7
Teacher introduced activity as written	4.0	3.3	4.0	3.8
Teacher guided as needed	3.0	3.4	3.6	3.3
Management strategies enhanced quality	2.8	3.0	3.8	3.2
Computer activity				
Mathematical focus	2.4	4.0	4.0	3.5
Teaching strategies appropriate	2.7	4.0	4.0	3.6
Organization, teaching, approaches, interactions	2.9	2.7	2.6	2.7
Materials were set up correctly[a]	4.0	4.0	4.0	4.0
Child was “signed in” with correct name[a]	4.0	3.7	4.0	3.9
Teacher focused mathematical thinking	1.6	2.2	2.2	2.0
Teacher monitored, was available as needed	3.1	3.1	2.5	2.9
Management strategies enhanced quality	3.3	2.9	2.9	3.0
All children engaged in activity that week	3.1	3.0	2.8	3.0
Teacher was actively involved	2.3	1.3	1.5	1.7
% time teacher actively involved[b]	2.0	1.0	1.0	1.3
Expectations	2.5	4.0	4.0	3.5
High, realistic expectations	3.0	4.0	4.0	3.7
Teacher promoted effort	2.0	4.0	4.0	3.3
Supporting children’s conceptual understanding	2.0	4.0	4.0	3.3
Teacher gave just enough assistance	2.0	4.0	4.0	3.3
Assessment and Instructional Adjustment	1.9	3.5	3.5	3.0
Teacher monitored activity, taking notes	1.7	3.0	3.0	2.6
Teacher can access records[a]	2.0	4.0	4.0	3.3
Descriptive items				
Total adults in classroom	2.0	2.1	2.3	2.1
No. of whole group activities	1.1	1.8	1.2	1.4
Total whole group duration (up to 3 activities)[c,d]	2.5	4.2	2.6	3.1
Small-group duration[c]	22.8	22.7	17.2	20.9
No. of center activities	0.4	0.8	0.9	0.7
Total center duration (up to 3 activities)[c,e]	0.7	1.2	1.7	1.2
Computer activity duration[c]	22.6	26.3	46.7	31.9

Note. Scale for all items without footnotes a or b: 1 (*strongly disagree*), 2 (*disagree*), 3 (*agree*), 4 (*strongly agree*).

[a] 1 = No, 4 = Yes. [b] Scale for percentage items 1 = 0–24%, 2 = 25–49%, 3 = 50–74%, 4 = 75–100%. [c] In minutes. [d] 1–5 min = 1; 6–9 min = 2; 10+ min = 3. [e] 1–20 min = 1; 21–45 min = 2; 46+ min = 3.

*1 = Fall, 2 = Winter, 3 = Spring.

of one of the three items in a section unique to this measure, "Personal Attributes of the Teacher," is "The teacher appeared to be knowledgeable and confident about mathematics (i.e., demonstrated accurate knowledge of mathematical ideas and procedures demonstrated knowledge of connections between or sequences of mathematical ideas)." Table 3 lists the COEMET's sections.

Interrater reliability for the COEMET was 88%; 100% of the disagreements had the same polarity (i.e., if one was agree, the other was strongly agree). Coefficient alpha (interitem correlations) for the two instruments ranged from .95 to .97 in previous research. Both instruments were submitted to the Rasch model, yielding T-scores and reliability of .96 for the COEMET and .90 for Fidelity.

Retired teachers identified by administrators as expert in early childhood mathematics teaching (in New York) and doctoral students and staff (in California) were trained on the COEMET. They practiced administering the instrument until they reached a satisfactory reliability. They remained naïve to the experimental curriculum and to the classrooms' treatment group. They were to visit all classrooms three times—fall, winter, and spring. This schedule was followed in New York, but in California only winter and spring data were collected.

In summary, those working with the TRIAD teachers on achieving fidelity, and assessing fidelity, were familiar with the curriculum and with teachers' treatment group. Those assessing both TRIAD and control conditions with the COEMET were naïve to treatment group.

Teachers' knowledge, and beliefs. Teachers completed a questionnaire at the beginning and the end of the school year. The Teacher Questionnaire measures teacher's self-reported knowledge and beliefs pertaining to early childhood mathematics. It includes sections on demographics, education and experience, mathematics goals, children's learning, and teaching (interrater reliability of coders, 97%).

Parents' knowledge, beliefs, and practices. Parents were also asked to complete a parent questionnaire (PQ) at the beginning and the end of the school year. The PQ provides descriptive information about the children's family, including parents' beliefs, knowledge, and practices pertaining to early childhood mathematics and their child's school readiness. One aggregate score measures parents' opinions of their child's mathematical competencies, and another measures the child's mathematics learning environment in the home. Most of the items on the end-of-the-year questionnaires were the same as the fall administration. However, some items were omitted, such as background information that would not have changed, and some items were added, such as their reactions to the year's experience with the TRIAD project.

Table 3. Means and standards deviations for the classroom observations (COEMET)

	Control				TRIAD			
	Observation*			M	Observation*			M
Scored Items	1	2	3	3 obs	1	2	3	3 obs
T-score *M*	47.1	44.7	40.7	44.2	52.5	45.3	49.6	48.2
Classroom elements								
No. of computers running math activities	1.3	1.0	1.0	1.1	1.3	1.0	1.0	1.1
M no. of math activities	3.0	2.9	2.5	2.8	3.3	3.1	4.0	3.6
M duration of math activities[a]	20.7	9.1	8.3	12.2	24.7	13.1	15.3	15.4
% teachers stayed in classroom	62.5	100	75.0	79.2	100	94.1	76.5	86.3
Classroom culture	2.57	2.84	2.51	2.63	3.52	3.10	3.09	3.11
Environment and interaction	2.38	2.85	2.22	2.48	3.14	2.97	2.82	2.88
Interacted with children	3.50	3.63	3.38	3.50	4.00	3.82	3.82	3.82
Used teachable moments	2.75	2.88	2.13	2.58	3.57	2.41	2.82	2.66
% time children using computers	1.63	2.88	1.50	2.00	3.29	3.47	2.29	2.91
Math work displayed	1.63	2.00	1.87	1.83	1.71	2.18	2.35	2.12
Personal attributes of the teacher	2.75	2.83	2.79	2.79	3.90	3.24	3.35	3.34
Knowledgeable about math	2.88	2.87	2.88	2.88	4.00	3.24	3.35	3.36
Believed math learning enjoyable	2.75	2.87	2.75	2.79	4.00	3.29	3.47	3.44
Enthusiasm for math ideas	2.63	2.75	2.75	2.71	3.71	3.18	3.24	3.23
Specific math activities	3.35	2.96	3.18	3.15	3.73	3.04	3.19	3.12
Mathematical focus	3.68	3.13	3.38	3.36	3.96	3.37	3.75	3.63
Understanding of topic	3.68	3.17	3.38	3.38	3.93	3.58	3.72	3.71
Developmentally appropriate	3.68	3.08	3.38	3.34	3.98	3.16	3.78	3.54
Organization, teaching approaches, interactions	3.30	3.10	3.19	3.19	3.68	2.94	3.09	3.02
Engaged children's mathematical thinking	3.07	2.88	3.34	3.11	3.61	2.38	2.79	2.63
Pace was appropriate	3.48	3.10	3.48	3.33	3.86	3.37	3.60	3.53
Management strategies improved activity	3.26	3.17	3.23	3.21	3.69	3.47	3.62	3.59
Actively involved	3.79	3.79	3.16	3.59	3.82	3.28	3.40	3.35
% time involved	3.37	3.49	2.75	3.23	3.61	2.81	3.06	2.92
Appropriate strategies	3.46	3.17	3.24	3.28	3.90	3.31	3.41	3.49
Expectations	3.45	3.14	3.38	3.33	3.83	3.27	3.18	3.25
High, realistic expectations	3.42	3.03	3.38	3.32	3.90	2.92	3.03	3.00
Acknowledged effort	3.48	3.25	3.38	3.33	3.75	3.61	3.33	3.50
Eliciting children's solution methods	3.14	2.75	3.18	2.99	3.45	2.52	2.45	2.40
Asked children to share ideas	3.14	2.75	3.36	3.03	3.39	2.27	2.15	2.14
Facilitated children's responses	3.25	2.75	3.17	3.03	3.61	2.81	2.71	2.68
Encouraged evaluating others	3.02	2.74	3.00	2.90	3.36	2.48	2.49	2.37
Supporting children's conceptual understanding	3.24	2.87	3.07	3.06	3.50	2.78	3.05	2.78
Supported describer's thinking	3.32	2.92	3.10	3.11	3.39	2.74	3.46	2.88
Supported listener's understanding	3.05	2.69	2.97	2.90	3.36	2.53	2.96	2.63
Gave just enough assistance	3.35	3.00	3.14	3.16	3.75	3.07	2.72	2.84
Extending children's mathematical thinking	2.73	2.67	3.10	2.79	3.56	2.17	2.31	2.30

Table 3. Means and standards deviations for the classroom observations (COEMET) (Continued)

	Control				TRIAD			
	Observation*			M	Observation*			M
Scored Items	1	2	3	3 obs	1	2	3	3 obs
Elaborated children's ideas	2.64	2.69	3.10	2.75	3.64	2.19	2.27	2.32
Encouraged mathematical reflection	2.82	2.65	3.10	2.82	3.48	2.14	2.35	2.28
Assessment and Instructional Adjustment	3.14	2.90	2.98	3.04	3.65	3.08	3.07	3.07
Listened to children, taking notes	3.10	2.86	3.05	3.08	3.67	3.47	3.49	3.53
Adapted tasks to ability and development	3.17	2.93	2.90	3.00	3.62	2.68	2.65	2.60
Nonscored items								
No. children	15.1	15.5	15.5	15.4	15.1	15.9	15.4	15.7
No. volunteers	0.1	0.0	0.3	0.1	0.3	0.1	0.3	0.2
No. of adults	2.5	2.0	1.9	2.1	2.3	2.7	2.5	2.6

Note. Classroom culture and Specific math activities: Scale for percentage items 1 = 0–24%, 2 = 25–49%, 3 = 50–74%, 4 = 75–100%. Scale for all other items: 1 (*strongly disagree*), 2 (*disagree*), 3 (*agree)*, 4 (*strongly agree*). Control $n = 8$, TRIAD, $n = 17$ (7 for Observation 1). COEMET = Classroom Observation of Early Mathematics—Environment and Teaching.

[a] In minutes.

*1 = Fall, 2 = Winter, 3 = Spring.

Analyses

Because children were nested within classrooms, child outcome data were analyzed using a hierarchical linear model (HLM). This was a cluster randomized trial, with the classroom the unit of random assignment. HLM accounts for both child- and classroom-level sources of variability in outcomes by specifying a two-level hierarchical model (Raudenbush, 1997). Thus, two-level analyses on the Rasch scores were computed to assess the effectiveness of the curricula and to ascertain the effects of class-level (Level 2) and child-level (Level 1) predictors and interactions of those predictors with treatment group. The model is presented in the appendix. After running this full model, any effects that were not significant were omitted and a parsimonious, reduced model computed. An effect size for the TRIAD intervention was computed by dividing the regression coefficient by the pooled posttest standard deviation.

RESULTS

We expected our research-based implementation to result in substantive gains across multiple contexts, leading to our inclusion of the following contextual variables: location (New York and California), types of Pre-K programs (e.g.,

Table 4. HLM Model

Fixed effects	*Coeff.*	*SE*	*t*	*p*
Intercept	53.35	0.41	131.53	$.000^{+}$
TRIAD	4.94	0.82	6.00	$.000^{+}$
Pre	0.90	0.05	16.48	$.000^{+}$
Random effect	*SD*	*Var*	χ^2	*p*
Intercept	1.18	1.34	32.63	.09
Level 1	4.85	23.54		

Note. All $df = 23$ except for Pre, with $df = 183$. Var = variance component, the between-classroom variance. TRIAD = Technology-enhanced, Research-based, Instruction, Assessment, and professional Development.

Head Start and State Preschools), and child/family characteristics (e.g., child socioeconomic status [SES]). Implementation variables (Fullan, 1992) are features that we encourage (but can not control absolutely); the main one for this study is fidelity of implementation.

Table 1 presents means and standard deviations for the REMA child outcome measure. To evaluate the effectiveness of the curriculum, we computed the full HLM model. In that model, only the pretest score and TRIAD treatment were significant. There were no other main effects and no significant interactions involving type of program or state; therefore, these were omitted from the model. (For exploratory analyses, we also computed an HLM model that included main effects and interactions with the treatment between the Level 1 variables of child-level ethnicity and gender, that is, modeling the intervention variable on those Level 1 slopes; these similarly revealed no significant effects.) The results of the reduced model, shown in Table 4, reveal that the TRIAD group made significantly greater gains than the control group ($p = .000$). Indeed, this model fully explained the variance between classrooms (the intercept variance was no longer significant). The effect size for the TRIAD intervention was .62.

An examination of the means for the mathematics topics assessed on the REMA (see Table 1) indicate that the TRIAD condition was not noticeably more effective than the control condition in the areas of comparing shapes, transformations (only a single item on the REMA), and measurement. The TRIAD condition had somewhat higher means in recognition of number and subitizing, arithmetic, and patterning. It had higher means on object counting, counting strategies, comparing number, and sequencing. The TRIAD condition's highest means were on the topics of verbal counting, composition of number, shape recognition, representing shape, and composing shape.

Table 2 presents means and standard deviations for the measure of fidelity. All but 10 of the 61 Fidelity items were generated by 4-point Likert scales from 1 (*strongly disagree*) to 4 (*strongly agree*). The mean Likert score was

2.8 ($SD = .60$). A mixed factorial repeated measures analyses was computed on the total Fidelity score (this Rasch T-score also included the six additional variables, such as number and duration of activities) to ascertain whether fidelity increased or decreased in the TRIAD classrooms. However, the analysis of variance (ANOVA) revealed no significant effect for time. Finally, the mean of the Fidelity scores correlated positively, but without reaching significance, with average child gain scores ($r = .48$, $p = .11$, recall the low sample size).

Table 3 presents means and standard deviations for the COEMET. Factorial repeated measures analyses were conducted on the Rasch T-scores to test for differences in the quality and quantity of the mathematics environment and teaching in the treatment, compared to the control, classrooms. The repeated measures ANOVA computed on the COEMET's T-scores showed a significant treatment effect, $F(1, 23) = 7.14$, $p = .014$, $MSE = 78.23$, but no significant main effect for time and no significant Treatment × Time interaction (analyses were computed on only the two scores, winter and fall, for which both sites had data). The TRIAD group had the higher scores.

As Table 3 indicates, TRIAD means were higher than control means for number of mathematical activities (3.6 vs. 2.8) and duration (15.4 vs. 12.2). With few exceptions ("supported listeners' understanding" was one), TRIAD teachers outscored control teachers. This difference was largest on items assessing general classroom culture behaviors of using teachable moments, engaging children in computer software, and all three personal attributes (appearing knowledgeable, believing mathematics enjoyable, enthusiasm) and on items assessing specific mathematics activities, such as understanding and teaching developmentally appropriate mathematics and using effective management and instructional strategies.

In addition, the posttest REMA score was regressed on the COEMET after controlling for REMA pretest score to test whether the observations predicted children's learning. The COEMET correlated with average child gain scores, although with only marginal significance ($r = .35$, $p = .09$). In summary, the intervention increased the quantity and quality of the mathematics environment and teaching in preschool classrooms.

Two questionnaires were administered—one to teachers, another to parents. The Teacher Questionnaire contained hundreds of questions; here we highlight only relevant, salient results. TRIAD teachers reported teaching mathematics for an average of 257 min per week, compared to the control teachers' 151 min, consistent with the COEMET observations.

Consistent with previous research (Sarama, 2002), teachers rated certain topics as very important, especially counting (4.00 and 3.94 for TRIAD and control, respectively, on a Likert scale of 1 [*not*] to 4 [*very*]) and number concepts and relationships (4.00, 3.77). In contrast to previous research, teachers rated most other topics as important, including patterns and relationships (4.00, 3.67), and geometry and spatial sense (3.96, 3.61). The trend for TRIAD teachers to rate mathematics as more important than control teachers was especially clear on the topics of computation (3.87 vs. 3.11), place value (3.96 vs. 2.96), algebraic thinking (3.46 vs. 2.52), and probability (3.63 vs. 2.63). The trend was present, but to a lesser degree, on the other topics—estimation, measurement, data collection, and technology. TRIAD teachers were more likely (.92 vs. .33) to state that "there should be a standard list of math topics" for preschool.

The questionnaire also asked teachers to rate how important those topics were to others, how prepared teachers believed they were to teach the topic, and how capable their children were in learning and mastering the topic. Generally, the same ratings and trends emerged, with a few notable distinctions. Teachers consistently rated others' (colleagues, administration) opinions of importance as lower than their own. Similarly, they rated their preparedness to teach the topic slightly lower than their rating of its importance, more so with algebraic thinking, probability, and technology. TRIAD teachers were notably higher in rating their preparedness to teach these same three topics.

Asked about the TRIAD curriculum on a 5-point scale from 1 (*strongly disagree*) to 5 (*strongly agree*), teachers in that group averaged 4.83 for adequacy of skills practice, 3.92 for assessments, and 4.0 for focus on understanding; they averaged 4.17 for "teachers would enjoy the program," 4.42 for the preparation of their children to meet standards, and 4.58 for the program's clarity. They averaged neutral to small disagreement (2.92) with the statement that the program might be too difficult for most teachers. Consistent with previous research (Sarama, 2002), teachers did not view financial rewards as key motivators to engage with the new program but instead rated as strong motivators increased student learning (4.83) and performance (4.67), personal satisfaction (4.83), and acquiring new curriculum materials (4.58). Other motivations were rated lower, including influence from administrators (3.33) or colleagues (3.08), course credit (3.00), and the possibility of increased pay (2.67).

The PQ was used for contact and simple descriptive information and to assess TRIAD's impact on children's home learning environments. The majority of the forms were completed by children's mothers. Their low SES status, inferred from their children receiving free and reduced lunch, was supported by the result that only 17% of TRIAD and 24% of control parents having any college education.

Two Rasch T-scores were computed. The first score indicated the parents' opinion of their children's mathematical skills, and the second indicated their opinion of the experiences that they provided their child in mathematics. Significant correlations between the parents' skills rating score and the REMA scores (Fall PQ to REMA pre, $r = .35$, $p = .000$; Spring PQ to REMA post, $r = .45$, $p = .000$) provide validation for both. However, the correlations between the parent experiences scores and the REMA scores were small and all but one nonsignificant (Fall PQ to REMA pre $r = .07$, *ns;* post $r = .05$, *ns*; Spring PQ to REMA post,$r = .20$, $p = .03$). The correlation between parents' highest educational level and their children's REMA scores was significant but moderate (REMA pre, $r = .26$, $p < .01$; post, $r = .23$, $p < .01$; change $r = 0.0$. *ns*).

A repeated measures ANOVA computed on the skills rating scores showed a significant effect for time, $F(1, 107) = 85.10$, $p < .000$, $MSE = 73.71$, but no significant effects for treatment, $F(1, 107) = 1.71$, $p = .19$, $MSE = 194.87$, or for Treatment × Time interaction, $F(1, 107) = 0.053$, $p = .82$, $MSE = 73.71$. Parents' ratings increased from fall to spring, but there was no effect of the intervention Table 5 displays the means for individual items of interest. A few items showed higher gains for the TRIAD group consistent with the REMA results, such as Item 1, counting (see the percentage who could count higher than 50); Item 4, recognizing shapes; and Item 5, recognizing numerals (similar relative gains were not found for colors, Item 3, but were for recognizing letters, Item 2). Repeated measures ANOVA computed on the experience scores showed a significant effect for time, $F(1, 107) = 21.79$, $p < .000$, $MSE = 82.20$, but no significant effects for treatment, $F(1, 107) = .52$, $p = .47$, $MSE = 136.19$, or for Treatment × Time interaction, $F(1, 107) = 2.07$, $p = .15$, $MSE = 82.20$. Parents' ratings increased from fall to spring, but effects of the intervention missed significance. Again, some items showed slight but higher gains for the TRIAD group, including Item 8 on art activities involving patterns or symmetry (all these items asked about weekly frequency), Item 9 on origami, Item 12 on games, Item 15 on puzzles, and Item 21 on activity books on math. The control group gained more on Item 16, playing on computer, although the opposite held for Item 22 on electronic toys, so little should be concluded. Finally, more TRIAD parents believed that their children's teachers provided a great deal of opportunities for learning math (Item 25). However, they were more likely not to complete activities teachers sent home (Item 27) and were less satisfied with them (Item 28) than were the control parents.

DISCUSSION

This study evaluated a limited scale up of the research-based TRIAD intervention model. Children in all groups made significant gains in mathematics knowledge during the preschool year. However, children in the TRIAD group, compared to those in the control group, made gains that were statistically and practically greater. The effect sizes indicate that the TRIAD model has a substantial influence on children's knowledge of mathematics. The effect sizes could be considered moderate to large (J. Cohen, 1977, uses .8 as the benchmark for large effects; Rosenthal & Rosnow, 1984, uses .5). There was no evidence that contextual variables influenced these positive effects. That is, there was no evidence that the TRIAD intervention was more effective with either program type or was differentially effective in either state—New York or California. (Exploratory analyses revealed no evidence that effects differed by children's ethnicity or gender.)

Differences by topic reveal that the intervention was more effective on some mathematical topics than on others. These differences are probably due

Table 5. Percentages for responses on parent questionnaire

	Control		TRIAD	
Item/Response	Pre	Post	Pre	Post
1. About how high can your child count?				
a. Not at all	0.0	0.0	0.0	0.0
b. Up to 5	2.9	3.2	3.9	0.0
c. Up to 10	29.0	14.5	29.9	8.5
d. Up to 20	53.6	38.7	53.2	39.0
e. Up to 50	11.6	21.0	9.1	33.9
f. Up to 100	1.4	19.4	3.9	15.3
g. Over 100	1.4	3.2	0.0	3.4
2. Does your child recognize the letters of the alphabet when he/she sees them?				
a. No	5.8	4.8	6.5	0.0
b. Some of them	43.5	34.9	53.2	13.6
c. Most of them	17.4	20.6	19.5	28.8
d. All of them	33.3	39.7	20.8	57.6
3. Does your child recognize the colors red, yellow, blue, and green?				
a. No	0.0	1.6	0.0	0.0
b. Yes	94.2	92.1	84.4	100.0
c. Some of them	5.8	6.3	15.6	0.0
4. Does your child recognize the shapes circle, square, triangle, and rectangle?				
a. No	1.4	5.0	6.5	0.0
b. Yes	71.0	76.0	67.5	96.6
c. Some of them	27.5	19.0	26.0	3.4
5. Does your child recognize the numerals 1–10?				
a. No	4.3	3.3	11.7	1.7
b. Yes	52.2	70.5	45.5	86.4
c. Some of them	34.8	19.7	41.6	5.1
d. Most of them	8.7	6.6	1.3	6.8
6. Read stories?				
a. No	9.0	6.3	2.6	0.0
b. Yes, once or twice	41.8	49.2	48.1	54.2
c. Yes, three or more times	49.3	44.4	49.4	45.8
7. Worked on arts or crafts?				
a. No	21.7	28.6	24.7	30.5
b. Yes, once or twice	62.3	63.5	61.0	62.7
c. Yes, three or more times	15.9	7.9	14.3	6.8
8. Art activities that involve patterns or symmetry?				
a. No	48.5	29.5	40.3	19.0
b. Yes, occasionally	48.5	34.4	50.6	39.7
c. Yes, often	3.0	36.1	9.1	41.4
9. Does your child do origami (paper folding) or kirigami (paper cutting)?				
a. No	53.7	43.3	47.4	42.1
b. Yes, occasionally	43.3	21.7	43.4	17.5
c. Yes, often	3.0	35.0	9.2	40.4

Table 5. Percentages for responses on parent questionnaire (Continued)

	Control		TRIAD	
Item/Response	Pre	Post	Pre	Post
10. Practiced letters, words, or numbers?				
a. No	1.4	4.8	5.3	1.7
b. Yes, once or twice	44.9	44.4	46.7	37.3
c. Yes, three or more times	53.6	50.8	48.0	61.0
12. Played a board game or a card game?				
a. No	31.9	34.9	32.0	16.9
b. Yes, once or twice	46.4	46.0	52.0	67.8
c. Yes, three or more times	21.7	19.0	16.0	15.3
14. Played with blocks?				
a. No	33.3	24.2	32.5	23.7
b. Yes, once or twice	43.5	48.4	44.2	57.6
c. Yes, three or more times	23.2	27.4	23.4	18.6
15. Played with puzzles?				
a. No	30.9	22.2	27.3	15.3
b. Yes, once or twice	36.8	58.7	54.5	59.3
c. Yes, three or more times	32.4	19.0	18.2	25.4
16. Played on the computer?				
a. No	29.0	30.2	46.8	27.1
b. Yes, once or twice	44.9	39.7	35.1	32.2
c. Yes, three or more times	26.1	30.2	18.2	40.7
18. Counted different things?				
a. No	4.3	1.6	5.2	0.0
b. Yes, once or twice	46.4	36.5	55.8	27.6
c. Yes, three or more times	49.3	61.9	39.0	72.4
19. Played counting games or sang counting songs?				
a. No	11.6	6.3	11.7	3.4
b. Yes, once or twice	53.6	55.6	59.7	52.5
c. Yes, three or more times	34.8	38.1	28.6	44.1
20. Read books with counting or shapes?				
a. No	16.2	15.9	17.1	17.5
b. yes, Once or twice	57.4	63.5	55.3	52.6
c. Yes, three or more times	26.5	20.6	27.6	29.8
21. Activity books with math activities (dot-to-dot, matching shapes)?				
a. No	47.8	22.2	40.8	13.6
b. Yes, once or twice	37.7	58.7	44.7	62.7
c. Yes, three or more times	14.5	19.0	14.5	23.7
22. Played electronic toys (LeapFrog, VTech)?				
a. No	41.2	34.9	37.3	27.1
b. Yes, once or twice	36.8	46.0	44.0	42.4
c. Yes, three or more times	22.1	19.0	18.7	30.5

Table 5. Percentages for responses on parent questionnaire (Continued)

	Control		TRIAD	
Item/Response	Pre	Post	Pre	Post
25. How much do your child's teachers provide opportunities for learning math in their classroom?				
a. not at all		0.0		1.8
b. slightly		8.3		0.0
c. moderately		21.7		26.3
d. a great deal		41.7		59.6
e. don't know		28.3		12.3
26. How often do your child's teachers send home math activities for you to do with your child at home?				
a. at least once a week		41.4		54.5
b. several times a month		34.5		29.1
c. once a month		13.8		3.6
d. a few times this year		10.3		12.7
e. not yet		0.0		0.0
27. How often do you complete these activities with your child?				
a. every time		57.1		5.1
b. almost every time		0.0		3.4
c. about half the time		9.5		13.6
d. less than half the time		4.8		28.8
e. never		28.6		49.2
28. How satisfied are you with these activities in helping your child to learn about math?				
a. very satisfied		88.9		73.7
b. somewhat satisfied		11.1		26.3
c. somewhat dissatisfied		0.0		0.0
d. very dissatisfied		0.0		0.0

Note. TRIAD = Technology-enhanced, Research-based, Instruction, Assessment, and professional Development.

Questions 11, 13, 17, 23, and 24 were open-ended and did not yield percentages.

to relatively less attention in all aspect of the intervention (e.g., professional development, curricula) to certain topics compared to the traditional curricula in the control classrooms. Certainly, there was more emphasis in the intervention classrooms on topics such as counting strategies, comparing numbers, shape recognition, representing shape, and composing shape that were not emphasized in control classrooms. However, future research should examine the topics and components of the curricula in more detail, as well as alternative explanations, such as the possibility that some aspects of mathematics knowledge are more amenable to intervention at the preschool level than others.

The two observational measures provide additional evidence that greater scores in achievement by the TRIAD group resulted from the TRIAD intervention. Results on the Fidelity instrument indicate that the teachers implemented the curriculum with acceptable fidelity across the three measurement periods. There were no indications that fidelity changed across those periods, indicating that the professional development was adequate to support the teacher throughout the year. In addition, the higher the teacher's fidelity score, the higher the average child gain in achievement in her classroom, although the correlation did not reach statistical significance.

Further, the TRIAD intervention resulted in consistently greater scores on the quality and quantity of the mathematics environments and teaching in the TRIAD, compared to control, classes. There was a tendency for TRIAD to have higher average scores on general classroom behaviors such as teaching more mathematics; showing knowledge of, enjoyment in, and enthusiasm for mathematics; and using effective management and instructional strategies. The COEMET correlated with average child gain scores, although with only marginal significance ($r = .35$, $p = .09$). In summary, the intervention increased the quantity and quality of the mathematics environment and teaching in preschool classrooms.

There was no reason to reject a null hypotheses that (a) parents will not have a significant effect on their children's development of mathematics during the children's preschool year, at least as measured by the mathematics experience items on the parent questionnaire, and (b) parents in the two treatment groups will provide similar mathematical experiences during the preschool year. The lack of variance in income may have accounted for the first results, and the low return rate of the questionnaires may have mitigated all findings with this instrument. Therefore, conclusions are tentative and these issues should be addressed in future research.

CONCLUSIONS AND IMPLICATIONS

The goal of the TRIAD project was to increase knowledge of scaling up by conducting research that investigates the effectiveness of a research-based mathematics education intervention implemented in varied Pre-K settings with diverse student populations. The moderate effect size is notable considering that other comprehensive reform programs, including multitiered teacher support, sustained professional development, and in-class coaching, achieve effect sizes such as .24 only with great effort (Balfanz, Mac Iver, & Byrnes, 2006). Based on a synthesis of the literature, the TRIAD model includes collaboration among key groups, extensive, multifaceted professional development, and strategies to maintain the integrity of research-based Pre-K mathematics curricula. Perhaps most unique is TRIAD's consistent emphasis on teaching for understanding following developmental guidelines, or learning trajectories, as well as its use of technology at multiple levels.

The mathematics curriculum was implemented with good fidelity. In addition, because of the TRIAD implementation, the quality of the mathematics environment and teaching was significantly greater in the intervention than the control group. Finally, TRIAD had a strong positive effect on children's mathematics achievement. The contextual variables tested did not affect the outcomes of the study. There is no evidence the approach is differentially effective for participants in different states or types of programs, or for children of different SES, ethnic group, or gender.

According to the Glenn Commission (2000) report, "at the daybreak of this new century and millennium . . . the future well-being of our nation and people depends not just on how well we educate our children generally, but on how well we educate them in mathematics and science specifically" (p. 6). This education must begin in Pre-K, and it must be improved, especially for low-income children—research indicates, the sooner the better, and the more the better (Ramey & Ramey, 1998). Our empirically tested curriculum and research-based intervention for scaling up may provide a model for such improvement in Pre-K mathematics, and the study adds to the knowledge relevant to scaling up in radically diverse contexts. There are two final caveats. First, this was a "proof of concept" study of limited size and without longitudinal follow-up. Second, it was conducted, as most such studies have been, with volunteer teachers, limiting the generalizability of results. We are presently conducting longitudinal studies in greater numbers of classrooms, including all teachers (not just volunteers) to evaluate TRIAD's impact on deep and sustained reform (Campbell & Silver, 1999) as well as an expansion into more settings. In so doing, the project will integrate research and education by simultaneously creating knowledge, promoting teaching and learning, broadening the participation of underrepresented minorities in Pre-K and inservice settings, and building and strengthening school–university partnerships.

ACKNOWLEDGMENTS

This material is based in part on work supported by the National Science Foundation Research Grants REC-0228440, "Scaling Up the Implementation of a Pre-Kindergarten Mathematics Curricula: Teaching for Understanding with Trajectories and Technologies," and ESI-9730804, "Building Blocks—Foundations for Mathematical Thinking, Pre-Kindergarten to Grade 2: Research-based Materials Development." Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. It should be noted that, although the research is concerned with the scale-up model and not particular curricula, the curricula evaluated in this research have been published by the authors, who thus could have a vested interest in the results.

REFERENCES

Balfanz, R., Mac Iver, D. J., & Byrnes, V. (2006). The implementation and impact of evidence-based mathematics reforms in high-poverty middle schools: A multi-site, multi-year study. *Journal for Research in Mathematics Education, 37*, 33–64.

Ball, D. L., & Cohen, D. K. (1999). *Instruction, capacity, and improvement*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.

Baroody, A. J., with Coslick, R. T. (1998). *Fostering children's mathematical power: An investigative approach to K-8 mathematics instruction*. Mahwah, NJ: Lawrence Erlbaum Associates.

Berends, M., Kirby, S. N., Naftel, S., & McKelvey, C. (2001). *Implementation and performance in New American Schools: Three years into scale-up*. Santa Monica, CA: Rand Education.

Blevins-Knabe, B., & Musun-Miller, L. (1996). Number use at home by children and their parents and its relationship to early mathematical performance. *Early Development and Parenting, 5*, 35–45.

Bodilly, S. J. (1998). *Lessons from New American Schools' scale-up phase*. Santa Monica, CA: RAND Education.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievment: A meta-analysis. *Review of Educational Research, 73*, 125–230.

Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.

Bredekamp, S. (2004). Standards for preschool and kindergarten mathematics education. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 77–82). Mahwah, NJ: Lawrence Erlbaum Associates.

Bresler, L., & Walker, D. (1990). Implementation of computer-based innovation: A case study. *Journal of Computer-Based Instruction, 17*(2), 66–72.

Bryant, D. M., Burchinal, M. R., Lau, L. B., & Sparling, J. J. (1994). Family and classroom correlates of Head Start children's developmental outcomes. *Early Childhood Research Quarterly, 9*, 289–309.

Campbell, P. F., & Silver, E. A. (1999). *Teaching and learning mathematics in poor communities*. Reston, VA: National Council of Teachers of Mathematics.

Carpenter, T. P., Fennema, E. H., Peterson, P. L., Chiang, C.-P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*, 499–531.

Clarke, D. M., & Clarke, B. A. (2004). Mathematics teaching in K-2: Painting a picture of challenging, supportive and effective classrooms. In R. Rubenstein & G. Bright (Eds.), *Perspectives on teaching mathematics: 66th Yearbook* (pp. 67–81). Reston, VA: National Council of Teachers of Mathematics.

Clements, D. H. (1984). Training effects on the development and generalization of Piagetian logical operations and knowledge of number. *Journal of Educational Psychology, 76*, 766–776.

Clements, D. H. (2002a). Computers in early childhood mathematics. *Contemporary Issues in Early Childhood, 3*(2), 160–181.

Clements, D. H. (2002b). Linking research and curriculum development. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 599–636). Mahwah, NJ: Lawrence Erlbaum Associates.

Clements, D. H. (2004). Geometric and spatial thinking in early childhood education. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 267–297). Mahwah, NJ: Lawrence Erlbaum Associates.

Clements, D. H., & Conference Working Group. (2004). Part one: Major themes and recommendations. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 1–72). Mahwah, NJ: Lawrence Erlbaum Associates.

Clements, D. H., & Sarama, J. (2002). Design of microworlds in mathematics and science education. *Journal of Educational Computing Research, 27*(1&2), 1–6.

Clements, D. H., & Sarama, J. (2003). *DLM early childhood express math resource guide*. Columbus, OH: SRA/McGraw-Hill.

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*, 81–89.

Clements, D. H., & Sarama, J. (2007a). *Building Blocks—SRA real math teacher's edition, Grade PreK*. Columbus, OH: SRA/McGraw-Hill.

Clements, D. H., & Sarama, J. (2007b). Effects of a preschool mathematics curriculum: Summative research on the *Building Blocks* project. *Journal for Research in Mathematics Education, 38*, 136–163.

Clements, D. H., Sarama, J., & Liu, X. (in press). Development of a measure of early mathematics achievement using the Rasch model: The Research-based Early Math Assessment. *Educational Psychology*.

Cohen, D. K. (1996). Rewarding teachers for student performance. In S. H. Fuhrman & J. A. O'Day (Eds.), *Rewards and reforms: Creating educational incentives that work* (pp. 61–112). San Francisco, CA: Jossey Bass.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic.

Confrey, J., Bell, K., & Carrejo, D. (2002). *Systemic crossfire: What implementation research reveals about urban reform in mathematics*. Austin: Systemic Research Collaborative for Education in Mathematics, Science, and Technology (SYRCE), University of Texas at Austin.

Copley, J. V. (1999). *The early childhood mathematics collaborative project: Year one*. Unpublished manuscript, University of Houston, Houston Texas.

Copley, J. V. (2004). The early childhood collaborative: A professional development model to communicate and implement the standards. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 401–414). Mahwah, NJ: Lawrence Erlbaum Associates.

Darling-Hammond, L. (1996). Restructuring schools for high performance. In S. H. Fuhrman & J. A. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 144–192). San Francisco: Jossey-Bass.

Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. In. San Francisco: Jossey-Bass.

Denton, K., & West, J. (2002, March 20). *Children's reading and mathematics achievement in kindergarten and first grade* (NCES 2002125). Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002125

Elmore, R. F. (1996a). Getting to scale with good educational practices. *Harvard Educational Review, 66*, 1–25.

Elmore, R. F. (1996b). Getting to scale with good educational practices. In S. H. Fuhrman & J. A. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 294–329). San Francisco: Jossey-Bass.

Ely, D. P. (1990). Conditions that facilitate the implementation of educational technology innovations. *Journal of Research on Computing in Education, 23*(2), 298–305.

Farran, D. C., Silveri, B., & Culp, A. (1991). Public preschools and the disadvantaged. In L. Rescorla, M. C. Hyson, & K. Hirsh-Pase (Eds.), *Academic instruction in early childhood: Challenge or pressure? New directions for child development* (pp. 65–73). San. Francisco: Jossey-Bass.

Ferguson, R. F. (1991). Paying for publication education: New evidence on how and why money matters. *Harvard Journal on Legislation, 28*(2), 465-498.

Fraivillig, J. L., Murphy, L. A., & Fuson, K. C. (1999). Advancing children's mathematical thinking in *Everyday Mathematics* classrooms. *Journal for Research in Mathematics Education, 30*, 148–170.

Fullan, M. G. (1992). *Successful school improvement*. Philadelphia: Open University Press.

Fullan, M. G. (2000). The return of large-scale reform. *Journal of Educational Change, 1*, 5–28.

Fuson, K. C., Smith, S. T., & Lo Cicero, A. (1997). Supporting Latino first graders' ten-structured thinking in urban classrooms. *Journal for Research in Mathematics Education, 28*, 738–760.

Galván Carlan, V. (2000). *Development of an instrument to assess the use of developmentally appropriate practices in teaching mathematics in early childhood classrooms.* Unpublished doctoral dissertation, University of Houston.

Galván Carlan, V., & Copley, J. V. (2000). *Early childhood mathematics profile observational checklist*. Houston, TX: University of Houston.

Ginsburg, A., Cooke, G., Leinwand, S., Noell, J., & Pollock, E. (2005). *Reassessing U.S. international mathematics performance: New findings from the 2003 TIMSS and PISA* Washington, DC: American Institutes for Research.

Ginsburg, H. P., & Russell, R. L. (1981). Social class and racial influences on early mathematical thinking. *Monographs of the Society for Research in Child Development, 46*(6).

Glenn Commission. (2000). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century.* Washington, DC: U.S. Department of Education.

Griffin, S. (2004). Number Worlds: A research-based mathematics program for young children. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 325–342). Mahwah, NJ: Lawrence Erlbaum Associates.

Griffin, S., Case, R., & Capodilupo, A. (1995). Teaching for understanding: The importance of the Central Conceptual Structures in the elementary mathematics curriculum. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning* (pp. 121–151). Mahwah, NJ: Lawrence Erlbaum Associates.

Heck, D. J., Weiss, I. R., Boyd, S., & Howard, M. (2002). *Lessons learned about planning and implementing statewide systemic initiatives in mathematics and science education*. New Orleans, LA: American Educational Research Association.

Hiebert, J. C. (1999). Relationships between research and the NCTM Standards. *Journal for Research in Mathematics Education, 30*, 3–19.

Hinkle, D. (2000). *School involvement in early childhood.* Washington, DC: National Institute on Early Childhood Development and Education, U.S. Department of Education Office of Educational Research and Improvement.

Holloway, S. D., Rambaud, M. F., Fuller, B., & Eggers-Pierola, C. (1995). What is “appropriate practice” at home and in child care?: Low-income mothers’ views on preparing their children for school. *Early Childhood Research Quarterly, 10*, 451–473.

Hord, S., Rutherford, W., Huling-Austin, L., & Hall, G. (1987). *Taking charge of change*. Alexandria, VA: Association for Supervision and Curriculum Development.

Horizon Research Inc. (2001). *Classroom observation protocol*. Chapel Hill, NC: Author.

Huberman, M. (1992). Critical introduction. In M. G. Fullan (Ed.), *Successful school improvement* (pp. 1–20). Philadelphia: Open University Press.

Jacobson, S. L., & Battaglia, C. F. (2001). Authentic forms of teacher assessment and staff development in the US. In D. Middlewood & C. Cardno (Eds.), *Managing teacher appraisal and performance a comparative approach* (pp. 75–89). London: RoutledgeFalmer.

Jacobson, S. L., Emihovich, C., Petrie, H., Helfrich, J., & Stevenson, R. (1998). *Transforming schools and schools of education: A new vision for preparing educators*. Newbury Park, CA: Corwin.

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.

Kaser, J. S., Bourexis, P. S., Loucks-Horsley, S., & Raizen, S. A. (1999). *Enhancing program quality in science and mathematics*. Thousand Oaks, CA: Corwin.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Klein, A., & Starkey, P. (2004). Fostering preschool children’s mathematical development: Findings from the Berkeley Math Readiness Project. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 343–360). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, A., Starkey, P., & Ramirez, A. B. (2002). *Pre-K mathematics curriculum*. Glenview, IL: Scott Foresman.

Klein, A., Starkey, P., & Wakeley, A. (2000). *Child Math Assessment: Preschool Battery (CMA)*. Berkeley: University of California, Berkeley.

Mohrman, S. A., & Lawler, E. E., III. (1996). Motivation for school reform. In S. H. Fuhrman & J. A. O’Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 115–143). San Francisco: Jossey-Bass.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R. A., O’Connor, K. M., et al. (2000). *TIMSS 1999 international mathematics report*. Boston: The International Study Center, Boston College, Lynch School of Education.

National Association for the Education of Young Children. (1991). *Early childhood classroom observation—National Academy of Early Childhood Programs* (Rev. ed.). Washington, DC: Author.

Natriello, G., McDill, E. L., & Pallas, A. M. (1990). *Schooling disadvantaged children: Racing against catastrophe*. New York: Teachers College Press.

Ramey, C. T., & Ramey, S. L. (1998). Early intervention and early experience. *American Psychologist, 53*, 109–120.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2*, 173–185.

Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Sarama, J. (2002). Listening to teachers: Planning for professional development. *Teaching Children Mathematics, 9*, 36–39.

Sarama, J., Clements, D. H., & Henry, J. J. (1998). Network of influences in an implementation of a mathematics curriculum innovation. *International Journal of Computers for Mathematical Learning, 3*, 113–148.

Sarama, J., & DiBiase, A.-M. (2004). The professional development challenge in preschool mathematics. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 415–446). Mahwah, NJ: Lawrence Erlbaum Associates.

Saxe, G. B., Guberman, S. R., & Gearhart, M. (1987). Social processes in early number development. *Monographs of the Society for Research in Child Development, 52*(2).

Schoen, H. L., Cebulla, K. J., Finn, K. F., & Fi, C. (2003). Teacher variables that relate to student achievement when using a standards-based curriculum. *Journal for Research in Mathematics Education, 34*(3), 228–259.

Secada, W. G. (1992). Race, ethnicity, social class, language, and achievement in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 623–660). New York: Macmillan.

Starkey, P., & Klein, A. (1992). Economic and cultural influence on early mathematical development. In F. L. Parker, R. Robinson, S. Sombrano, C. Piotrowski, J. Hagen, S. Randoph, et al. (Eds.), *New directions in child and family research: Shaping Head Start in the 90s* (pp. 440). New York: National Council of Jewish Women.

Starkey, P., Klein, A., Chang, I., Qi, D., Lijuan, P., & Yang, Z. (1999). *Environmental supports for young children's mathematical development in China and the United States*. Albuquerque, NM: Society for Research in Child Development.

Teaching Strategies Inc. (2001). *A checklist for assessing your program's implementation of the Creative Curriculum for early childhood* (3rd ed.). Washington, DC: Author.

Tyack, D., & Tobin, W. (1992). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal, 31*, 453–479.

Weiss, I. R. (2002). *Systemic reform in mathematics education: What have we learned?* Research presession of the 80th annual meeting of the National Council of Teachers of Mathematics.

Zill, N., Resnick, G., Kwang, K., McKey, R. H., Clark, C., Pai-Samant, S., et al. (2001). *Head Start FACES: Longitudinal findings on program performance, Third progress report*. Washington, DC: U.S. Government Printing Office.

APPENDIX

HLM Model

The Level 1 model was

$$Y_{ij} = \beta_{0j} + \gamma_{01}(\Pr e)_j + r_{ij},$$

where

Y_{ij} is the posttest latent mathematical competence of child i in class j ($j = 1 \ldots 25$ classrooms);
β_{0j} is the mean outcome in class j;
Pre is the mean pretest score;
γ_{01} is the effect for the pretest; and
r_{ij} is the residual (level-1 random effect).

The level-2 model was

$$\begin{aligned}\beta_{0j} = {} & \gamma_{00j} + \gamma_{01}(PT)_j + \gamma_{02}(State)_j + \gamma_{03}(TRIAD)_j + \gamma_{04}(iTRPT)_j \\ & + \gamma_{05}(iTRSt)_j + u_{0j},\end{aligned}$$

where

γ_{00} is the mean achievement in the classrooms (intercept); and
PT is a dummy code for program type (Head Start or state funded);
γ_{01} is the main effect for PT;
State is a dummy code for the state (NY vs. CA);
γ_{02} is the main effect for State;
TRIAD is a treatment-indicator variable for the TRIAD intervention;
γ_{03} is the treatment effect for TRIAD;
iTRPT is the interaction of TRIAD and PT;
γ_{04} is that interaction effect;
iTRSt is the interaction of TRIAD and State;
γ_{05} is that interaction effect; and
u_{0j} is the residual (level-2 random effect).

All Level 2 predictors were centered around their grand means. All interactions were computed on mean-centered transformations of the variables involved.