Effects of a Preschool and Kindergarten Mathematics Curriculum: Big Math for Little Kids

Ashley Lewis Presser[1]
Education Development Center, Inc.

Margaret Clements[2]
Education Development Center, Inc.

Herbert Ginsburg[3]
Teachers College, Columbia University

Barbrina Ertle[4]
Adelphi University

This research was funded by Grant R305K040001 from the U.S. Department of Education, Institute for Education Sciences, to EDC. We thank the childcare centers, children, and teachers for their participation in this study. In addition, we would like to offer our sincere gratitude to Ellen Mandinach, Melissa Morgenlander, Leslie Manlapig, Maria Cordero, and Tracey Curran for their contributions to the research.

Correspondence concerning this article
should be addressed to Ashley Lewis Presser (alewis@edc.org)

1. Corresponding Author. 96 Morton Street, 7th Floor, New York, NY 10014. Fax 212-633-8804. Phone 212-807-4264. E-mail alewis@edc.org.

2. 96 Morton Street, 7th Floor, New York, NY 10014. Fax 212-633-8804. Phone 212-807-4268. E-mail pclemetns@edc.org.

3. Department of Human Development, 542 Grace Dodge Hall, Teachers College Columbia University, 525 W. 120th Street, New York, NY 10027. Fax 212-678-3837. Phone 212-678-3443. E-mail ginsburg@tc.edu.

4. Department of Human Development, 542 Grace Dodge Hall, Teachers College Columbia University, 525 W. 120th Street, New York, NY 10027. Present Address: Ruth S. Ammon School of Education, Harvey Hall, Room 3, Garden City, NY 11530-0701. Fax: 516-877-4097. Phone: 516-877-4064. E-mail: bertle@adelphi.edu.

Abstract

Research Findings: *Big Math for Little Kids* (*BMLK*) is a mathematics curriculum designed for 4- and 5-year-old children. In this study, the curriculum was evaluated for effectiveness over two years, using a cluster-randomized controlled study. Over 750 children participated in the study and experienced either the *BMLK* curriculum or business-as-usual instruction. Students' mathematics knowledge was assessed using the Early Childhood Longitudinal Study–Birth Cohort (ECLS-B) Direct Mathematics Assessment, an independent outcome measure not tied to the curriculum materials. The *BMLK* children significantly outperformed the business-as-usual control group, a difference that represents the equivalent of 1.6 months of additional instruction, with a medium effect size (*Cohen's d*=0.40). *BMLK* children also showed indications of improved mathematical language on piloted language tasks. *Policy or Practice*: These results suggest that the inclusion of thoughtful, developmentally appropriate mathematics curriculum can positively impact young students' achievement.

Keywords: mathematics, evaluation/outcome, inner-city, cognitive, center-based

Effects of a Preschool and Kindergarten Mathematics Curriculum:
Big Math for Little Kids

Early childhood mathematics education and the levels of young children's mathematics knowledge are receiving increased attention from educators, researchers, and policy makers. One concern motivating the development of early childhood education policy and funding is that American children's performance on mathematics assessments is weaker than that of children in a number of other developed countries (Gonzales et al., 2008; Miller & Parades, 1996) and below what experts deem proficient (Kilpatrick, Swafford, & Findell, 2001; Mullis et al., 1997; Mullis et al., 1998). For example, results from the National Assessment of Educational Progress (2011) indicate that 82% of fourth graders score at or above the "basic" category, 40% are "proficient," and only 7% are "advanced."

Children's mathematics achievement trajectories are established in the early primary grades, and children who begin elementary school behind their peers in mathematics tend to fall further behind over time (Duncan et. al., 2007; Entwistle & Alexander, 1989; Starkey & Klein, 2008). Recognizing the importance of mathematics education in early childhood, a number of national organizations have made clear recommendations about the need for better mathematics education for young children. In 2002, a joint statement from the National Association for the Education of Young Children (NAEYC) and the National Council of Teachers of Mathematics (NCTM) recommended that early childhood programs include a challenging, research-based, developmentally appropriate mathematics curriculum that will support effective mathematics learning (NAEYC & NCTM, 2002). A report from the National Research Council's Committee on Early Childhood Mathematics (NRC-CECM) has expanded upon these recommendations, focusing on learning, teaching, teacher education, and curriculum (Cross, Woods, &

Schweingruber, 2009). The report points to the need for high-quality early mathematics instruction and improved teacher preparation and training in order to provide all children with the mathematical foundation necessary for academic success in elementary school and beyond. In spite of this increased focus on the importance of early childhood mathematics education, teachers often provide little math teaching for all children at the preschool and kindergarten levels (Ginsburg, Lee, & Boyd, 2008) and the teaching that does take place is of poor quality, leading Pianta and LaParo to characterize early education environments as “socially positive but instructionally passive” (Pianta & La Paro, 2003, p. 28). Yet there is a renewed emphasis on mathematics in the early grades with the release of the Common Core mathematics standards, which start in the kindergarten year (National Governors Association Center for Best Practices, 2010).

This paper reports the findings from a two-year cluster-randomized controlled trial investigating the effectiveness of *Big Math for Little Kids (BMLK)*, a preschool mathematics curriculum intended to facilitate the mathematics learning of 4- and 5-year-old children. Children from mostly single-parent families living below the federal poverty threshold who attended publically-subsidized child care centers in New York City participated in the study with their teachers. To measure achievement, we employed a nationally-normed outcome measure from the ECLS, as well as a specially constructed language measure. We begin with some background on the need for the study and on the principles guiding construction of the mathematics curriculum.

SES Differences in Math Achievement

Math achievement among American children is below that of many other developed countries, perhaps as early as preschool (Miller & Parades, 1996) or kindergarten (Stevenson, Lee, & Stigler, 1986), but certainly by the 3rd or 4th grade, according to the Trends in

International Mathematics and Science Study (TIMSS) (Lemke & Gonzales, 2006). The picture is especially bleak for "disadvantaged children," meaning those from poor and poorly educated families, often living in unsafe neighborhoods, and disproportionally composed of African Americans and Latinos (Douglas-Hall, Chau, & Koball, 2006). Of particular concern is the fact that disadvantaged children receive a poorer education than do middle-socioeconomic-status (SES) children and are thus caught in a vicious cycle: "Poor educational attainment is a major cause of poverty, and poverty is a key influence on academic failure" (Arnold & Doctoroff, 2003, p. 518).

Over six million children—about 22% of the entire US child population under six years old—are from families with incomes below the federal poverty threshold (Chau, Thampi, & Wight, 2010). As a group, children from poor and low-income families begin kindergarten with a less–well-developed understanding of early mathematics concepts than do their higher-income peers (Rathbun, West, & Hausken, 2004; National Mathematics Advisory Panel, 2008), which does not bode well for their future. A recent study examining several large data sets found that early math test scores (at entry to school) predict later performance in both mathematics and reading with considerable accuracy (Duncan, et al., 2007). Similarly, "number sense" measures given during kindergarten and 1st grade predict later scores on a high-stakes mathematics achievement test at the 3rd grade (Jordan, Glutting, Ramineni, & Watkins, 2010). The results of both studies suggest that low-SES children who start behind continue to remain behind.

It is important to place low-SES students' performance in a wider perspective. Not only do they grow up in more challenging conditions than do their more affluent peers, but low-SES children also receive an inferior education (Arnold & Doctoroff, 2003; Lee & Burkham, 2002). For example, schools serving low-SES children receive lower public funding than schools serving higher-SES children (Arnold & Doctoroff, 2003), and provide inadequate educational opportunities (Lee & Burkham, 2002). "…[T]eachers have lower expectations and more negative perceptions of low-SES students than their higher-SES peers…" (Arnold & Doctoroff, 2003, p. 522). To determine if the inclusion of research-based early childhood mathematics curricula could ameliorate this risk for low-SES students, this study compared low-SES students in classrooms using a research-based early mathematics curriculum with students in classrooms teaching the usual mathematics activities.

Research on Mathematical Thinking

Fortunately, cognitive and motivational research findings point the way to the development of innovative programs of early childhood mathematics education (ECME) that may improve mathematics achievement in the early years and proactively prevent the achievement gap that exists between high-SES and low-SES children.

Everyday mathematics. For at least 25 years, researchers have accumulated a wealth of evidence showing that young children develop important forms of mathematical competence (Baroody, Lai, & Mix, 2006; Clements & Sarama, 2007; Ginsburg, 2006). Throughout the preschool years, children's everyday mathematics understanding develops in interesting ways, often without adult assistance. Young children develop an everyday mathematics entailing a variety of topics (e.g., space, shape, pattern, number and operations) and comprising several important features described below: interest, concrete and abstract thinking, and understanding and misconceptions (Ginsburg, Cannon, Eisenband, & Pappas, 2006). Furthermore, this emerging mathematics competence develops for both boys and girls, as at the preschool level several studies indicate no gender differences in mathematical competence (Dowker, 2005).

Below, we provide three examples of young children's early interest in and understanding about mathematics.

First, young children have a spontaneous and sometimes explicit interest in mathematical ideas. Naturalistic observation has shown that in the block area, for example, young children spend a good deal of time determining which tower is higher than another, creating and extending interesting patterns with blocks, exploring shapes, creating symmetries, and the like (Seo & Ginsburg, 2004). Everyday mathematics is not an imposition from adults; indeed, adults— including teachers—are often unaware of its existence.

Second, their thinking is both concrete and abstract. Young children can add 3 toy dogs to 4 toy dogs to get the sum, but they also have abstract ideas about counting objects, including the abstraction principle, such as the rule that different types of discrete objects can be counted, from stones to unicorns (Gelman & Gallistel, 1986).

Finally, children's mathematical knowledge demonstrates both understanding and misconceptions. Although they seem to understand basic ideas of addition and subtraction from an early age (Brush, 1978), they fail to realize that an odd-looking triangle (for example, an extremely elongated, non-right-angle, "skinny" triangle) is as legitimate a triangle as one with three sides the same length (Clements, 1999).

In conclusion, while young children's natural mathematical learning is impressive, it is also limited. To most effectively develop more comprehensive and abstract thinking about mathematics, children often need more than their natural, spontaneous learning. Instead, they need experiences that expose them to mathematical concepts in a progressive and developmental fashion.

Factors associated with young children's mathematics learning. There are a number of factors associated with children's mathematics learning, including language, SES, and motivation.

Recent research stresses the importance of language in mathematical thinking (Rudd, Lambert, Satterwhite, & Zaier, 2008). Young children have not only to acquire the standard mathematical vocabulary—for example, words for quantity ("bigger," "less")—but, importantly, they must also learn how to employ this language to express and justify their mathematical thinking. With development, children become increasingly aware of their own thinking and begin to express it in words (Kuhn, 2000). These kinds of metacognitive skills are as necessary for mathematics as for other topics, and begin to develop in children as young as four or five years of age (Pappas, Ginsburg, & Jiang, 2003). In addition, growth of preschoolers' conventional mathematical knowledge over the school year is significantly impacted by the amount of math-related talk their teachers use (Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006). Clearly, language is deeply imbedded in mathematics learning and teaching.

During piloting of the *BMLK* curriculum, teachers reported being impressed by the vocabulary and language development of their students (Ertle et. al., 2008). Subsequent classroom observations supported teachers' anecdotal reports, observing that children's use of mathematical language increased in frequency and their mathematical explanations became more robust (Greenes, Ginsburg, & Balfanz, 2004). Based on these observations, the authors hypothesized that "the learning of mathematics … is inevitably tied to the development of language and communication skills" (Greenes, Ginsburg, & Balfanz, 2004, pg. 165).

Beyond mathematical language, children's socio-economic background is also associated with their mathematics learning. A close examination of the cognitive literature reveals that the pattern of how children from different SES backgrounds vary is complex and interesting. First, although lower-SES children's performance on informal addition and subtraction problems often lags behind that of middle-SES children, the two groups often employ similar strategies to solve problems (Ginsburg & Pappas, 2004). Both groups use methods such as counting on from the larger number or "derived facts." An example of this type of thinking would be "4 and 5 equals 9 because I know that 4 and 4 is 8 and 5 is just 1more, so the answer has to be 9."

Second, although lower-SES children exhibit difficulty with *verbal* addition and subtraction problems, they perform as well as middle-SES children on *non-verbal* forms of these tasks (Jordan, Huttenlocher, & Levine, 1994). They do not lack the basic skills or concepts of addition and subtraction, but are not able to verbally articulate these understandings. Third, lower- and middle-SES children exhibit few, if any, differences in the everyday mathematics they spontaneously employ in free play (Ginsburg, Pappas, & Seo, 2001). In brief, although their performance on early childhood mathematics assessments may be weaker, low-SES children exhibit a good deal of competence upon which early childhood mathematics education (ECME) can build.

Motivation is another factor in learning mathematics. Young children generally begin schooling with high academic interest in learning. Arnold and Doctoroff (2003) describe this as "the cluster of variables that includes interest, motivation, engagement, goals, values, and self-efficacy" (p. 520). At the start of school, young children are self-confident and have positive attitudes towards school. But their motivation decreases within the first few years of school, most likely because of educational factors such as boring and developmentally inappropriate teaching (Arnold & Doctoroff, 2003).

Summary

In the ordinary environment, young children develop an everyday mathematics that encompasses more than numeracy; it entails a variety of topics, including space, shape, and pattern, as well as number and operations. It is both concrete and abstract, involves both skills and concepts, and may be learned spontaneously as well as with adult assistance. The question of whether young children are "ready" to learn mathematics is beside the point. Learning mathematics occurs naturally and is a developmentally appropriate activity for young children. Without much direct adult assistance, they are motivated to learn and are already aquiring some real mathematical skills and ideas. It is also true that children from lower-SES family backgrounds demonstrate less-proficient mathematical performance than do their middle-SES peers, particularly when metacognition is required, but that they enter early childhood education environments with the basic skills and concepts they need to become mathematically proficient.

What Can be Done?

Numerous studies show that high-quality early education experiences provide children with the foundation for later academic success, especially in the short term and arguably in the years thereafter (Bowman, Donovan, & Burns, 2001; Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Gormley, 2007; Ludwig & Phillips, 2007; Reynolds & Ou, 2003). The basic rationale is that early education may provide a kind of "cognitive multiplier" (Siegler, 2010): The more a child knows at the outset, the more he or she can form associations and construct meaning, especially since mathematics is often structured in a hierarchical manner.

Early education may even be seen as a good financial investment, resulting in economic benefits over the long term (Heckman, 2000). "…[T]he value of benefits is very large relative to costs, even for very costly programs" (Pianta, Barnett, Burchinal, & Thornburg, 2009, p.50). Fortunately, cognitive and motivational research findings point the way to the development of innovative programs of ECME that not only prepare for the future but engage children in active, meaningful, challenging, and often enjoyable learning.

The value of ECME programs. Several mathematics curricula drawing on contemporary research have been shown to produce learning gains in young children. These include *Building Blocks* (Clements & Sarama, 2007); the *Measurement-Based* approach (Sophian, 2004); the *Number Worlds* curriculum (Griffin, 2007); the *Pre-K Mathematics Curriculum* (Klein, Starkey, & Ramirez, 2002); and *Storytelling Sagas* (Casey, Kersh, & Young, 2004). Other authors (e.g. Young-Loveridge, 2004) also have found that classroom use of games and storybooks can improve children's numerical skills and enhance their motivation for mathematics learning. Thus, thoughtfully designed and executed activities can help to further develop the mathematical understanding children already have and to challenge their misconceptions.

In this study, *BMLK* was compared to a business-as-usual control group, which included use of the *Creative Curriculum* (*CC*), *High Scope*, and other homegrown curricula. The majority of these control classrooms used *CC* (Dodge, Colker, & Heroman, 2002), which is extremely popular in early childhood programs. It is a play-based curriculum that seeks to aid children's learning in four developmental domains, including social-emotional, cognitive, language, and physical. Children spend 60 minutes a day or more in self-selected interest areas (e.g., blocks, dramatic play, table toys, art, sand/water, library, science, music/movement, cooking, computers, and outdoors). The *CC* outlines learning objectives and related developmental milestones that children should reach in the four developmental domains (see Table A1, Appendix). The teacher's primary role is that of an observer. After making written observations, teachers plan whole-group, small-group, and individual instruction to help each child reach the next stage of development. The methods, scope, sequence, and dosage of instruction are left to the teacher's discretion.

The *Big Math for Little Kids* Program. This paper presents an evaluation of *Big Math for Little Kids* (*BMLK*), a pre-kindergarten and kindergarten curriculum developed with funding from the National Science Foundation (Ginsburg, Greenes, & Balfanz, 2003). The program was constructed on theory- and research-based principles as discussed above, including the following.

- Young children are already engaged in learning (informal) mathematics. They do not need to be made ready to learn.
- Young children already possess many basic informal mathematical ideas upon which instruction can be built.
- Play is not enough for optimal mathematics learning.
- Sensitive adult guidance can help children engage in complex forms of mathematics learning and to realize their learning potential.
- The mathematics curriculum should stress not only basic ideas and procedures, but also the verbal expression of mathematical thinking.
- Low-income children in particular need help in describing their mathematical thinking and making explicit their mathematical competence.

BMLK offers a separate curriculum for prekindergarten (approximately age 4) and kindergarten (age 5) children. At each age level, the curriculum presents a systematic approach to teaching mathematics. *BMLK* offers a structured sequence of activities designed to promote challenging mathematical learning and related verbal expression. *BMLK* is designed to be used at least 20 to 30 minutes each day, for a total of approximately 32 week—the length of the typical academic year at these age levels. A teacher guide describes the *BMLK* lessons in detail to support teachers as they implement with large groups and small groups, and with individual children. The lessons take the form of games, activities with manipulatives, stories, and various other activities. Some of the activities include a very small amount of work related to reading and writing about mathematics.

The teachers' guide contains background information on the program, explicit learning goals, a planning chart, take-home activities for parents (in English and Spanish), and suggestions on how to assess children's mathematical learning and thinking in the context of instruction. In an effort to inform and involve families in the learning process, activities and letters are sent home. Each unit also addresses mathematical learning through literacy with a list of mathematical terms that teachers should use, introduce to children, and encourage children to use and a storybook that engages the children in the mathematical ideas and uses the relevant vocabulary. A black-and-white version of the storybook is given to each child to take home and share with his or her family.

The curriculum covers six units: number, shape, patterns and logic, measurement, number operations, and spatial relations. Each of these math concepts is first introduced in the pre-K curriculum and then further developed in the kindergarten curriculum. In the *number* unit, children learn to say the counting sequence, to use a number to tell how many (cardinality), and to use ordinal numbers to identify positions in a line (ordinality). In the *shape* unit, children learn the names and important attributes of two- and three-dimensional shapes, as well as the concept of symmetry. The *patterns and logic* unit gives children experience with patterns involving sound, color, shape, letters, and numbers. Children also learn to reason logically through the use of clues. In the *measurement* unit, children develop basic measurement principles as they investigate length, weight, capacity, temperature, time, and money. The *number operations* unit extends children's understanding of number by introducing addition, subtraction, and introductory multiplication and division concepts. In the *spatial relations* unit, children learn to identify positions in space, navigate through space, and represent space using maps.

Research Questions

Our primary research question for the study was whether *BMLK* produces gains in achievement above and beyond the mathematics instruction that typically occurs in preschools for low-SES children ("business as usual"). The program was developed with the assumption that low-SES four- and five-year-olds can learn math when it is taught in an organized way that builds on what is known about young children's mathematical competence.

Given the developing evidence regarding the potentially important role of language in children's mathematics learning, the study also investigated a secondary research question: Do children exposed to *BMLK* develop a greater ability to use and understand mathematical language compared to children who participate in their preschool's typical mathematics instruction?

Methods

This study was a two-year longitudinal cluster-randomized control trial. The study team collaborated with the New York City Administration for Children's Services (ACS), which supported the study and helped recruit childcare centers from among those it administered. We randomly assigned participating childcare centers to either the intervention group, which used *BMLK*, or the business-as-usual (BAU) comparison group, which used their existing mathematics curriculum, usually *Creative Curriculum* (Dodge, Colker, & Heroman, 2002).

The study followed children from the beginning of their prekindergarten year through the end of their kindergarten year, collecting data on children's early mathematics knowledge in the fall and spring of each year (four waves of data).

Recruitment and Random Assignment

Knowing that that it was common for children to attend an ACS childcare center for their prekindergarten year and transfer to a public elementary school for their kindergarten year, the focus in recruitment was on centers that served the same group of children across both years. To ensure a sufficient number of children eligible to participate in both years of the study, our two recruitment criteria were (1) that a center contain at least two prekindergarten classes and one kindergarten class, and (2) that it enroll at least 20 prekindergarten students total. Of the 115 ACS centers operating at the time of recruitment, 31 centers both met the study's selection criteria and expressed a willingness to participate.

From these centers, the study team randomly selected 16 childcare centers to participate in the study and then randomly assigned them to either the *BMLK* intervention group or the BAU comparison group. We notified all remaining interested and eligible centers that we would place them on a waiting list. After random assignment, four centers withdrew from the study. To replace those centers in the sample, the study team randomly selected centers from the wait-list group and these centers were invited to participate in the study.

The *BMLK* teachers received the curriculum and targeted professional development on use of *BMLK* during their year of participation in the randomized trial. Meanwhile, the BAU teachers received the standard professional development offered by ACS during the trial. After the conclusion of the research, all the BAU teachers received the *BMLK* curriculum and were invited to receive the professional development. Finally, childcare centers that were not selected to participate in the study—the wait-list group—were invited to send their center teachers to the *BMLK* teacher-training sessions. Thus, all teachers who were recruited were eventually provided with both the curriculum itself and the opportunity for the professional development to accompany it, although not all teachers in the control group and wait-list group decided to take advantage of that opportunity.

Child Sample

Children attending prekindergarten (during Year 1 of the study) or kindergarten (during Year 2) in a participating childcare center composed the sample. Over the two years of the study, a total of 762 children participated in data collection (N = 646 in prekindergarten and N = 385 in kindergarten), with 268 children participating for both years of the study (prekindergarten and kindergarten), 378 children participating in prekindergarten only, and 116 participating in kindergarten only.

The study team collected data on children's race/ethnic background, age, gender, language spoken, and monthly family income from centralized records that ACS maintains on children attending its childcare centers. In cases where these data were not available from ACS, study team members attempted to collect it directly from the childcare centers. Table A2 (Appendix) presents descriptive statistics for child demographic characteristics for the sample as a whole, as well as for the *BMLK* and BAU groups separately. Most children in the study came from small, single-parent, poor families. Specifically, the vast majority of children (93%) lived with a single parent in households that, on average, included 3.0 family members and reported a monthly family income of $1,356, which was essentially the federal poverty threshold ($1,354) for a single-parent family with two children in 2006.

There was baseline equivalence for the *BMLK* and BAU groups on these demographic variables. To create the child race/ethnic group background variable, we combined the data from two variables—one indicating whether the child's ethnic background was Hispanic and the second recording the child's race (white, black, American Indian or Alaskan Native, Asian, or native Hawaiian/other Pacific islander)—to indicate whether a participating child was black/not Hispanic (53.4%) or Hispanic/any race (46.6%). Close to one-quarter of the child sample spoke Spanish at home. Children's average age in the fall of the first (prekindergarten) year of the study was 4.3 years old[1] and 50.7% of the children were female. There was baseline equivalence between the *BMLK* and BAU groups for age and language spoken at home. However, the two groups differed in terms of race/ethnic group and gender, with the *BMLK* group including more girls and Hispanic children, and the BAU group including more boys and black/non-Hispanic children.

Teacher Sample

Participating prekindergarten and kindergarten teachers filled out a questionnaire that included questions about their education, teaching experiences, and demographic background. Table A3 (Appendix) presents descriptive statistics for these variables for the sample of teachers as a whole, as well as for the *BMLK* and BAU teachers separately. Of the 48 participating teachers (32 prekindergarten teachers and 16 kindergarten teachers), 47 completed the questionnaire.

The majority of the participating teachers were female (89.4%). Their race/ethnic backgrounds were Black/non-Hispanic (53.2%), Hispanic (29.8%), or another race/ethnic group (17.0%). For the majority of participating teachers, the language they spoke at home was English

[1] We calculated children's ages by subtracting their birth dates from October 15, 2005.

(71.7%), with the remaining teachers reporting that they spoke Spanish (10.9%), English and Spanish (6.5%), or some other language (10.8%). On average, teachers had taught for 11.0 years in the ACS system (SD = 9.0) and taught for 14.9 years (SD = 9.1) overall. Participating teachers had a range of education backgrounds representing a variety of academic degrees, with 47.9% reporting an earned master's degree, 37.5% reporting an earned bachelor's degree, and the remaining teachers reporting an associate's degree or an associate's degree in progress (14.6%). In addition to their academic degree, 25% had their teaching certification. There were not statistically significant differences between the treatment and the BAU teachers on any of the demographic or background variables.

Professional Development Workshops for *BMLK*

One of the *BMLK* authors and his colleagues developed a series of workshops for teachers implementing *BMLK* (Ginsburg et. al., 2006). This series was a revision of a program developed and piloted several years earlier with early childhood educators in low-income communities in New York and New Jersey.

For each year of this study, there were nine workshops in total. The first was a full-day session in the summer preceding the school year. Beginning in October of the school year, eight *BMLK* workshops were offered, at roughly monthly intervals, on Friday mornings for approximately 2½ hours each. During the prekindergarten year, six teachers missed one session per year, whereas in the kindergarten year one teacher missed four sessions, one teacher missed two sessions, and four teachers missed one session.

Each workshop was devoted to one of the units (i.e., number, number operations, shape, space, pattern, measurement). Two debriefing workshops also were held, one mid-year and the other at the end of the year. Each workshop after the first began with a group discussion about the previous month's workshop. These discussions were designed to serve as debriefing sessions in which teachers had an opportunity to discuss problems, issues, and successes they had encountered in the preceding month of teaching the *BMLK* curriculum. After this discussion, the workshop focused on the following four topics for each mathematics unit being covered.

Topics 1 and 2: What is the math? What do children know about the math? Helping teachers understand the mathematics of the target unit was often discussed simultaneously with a review of research on the development of children's understanding of the mathematics. The goal in covering these two topics was to help teachers understand the depth of these "simple" mathematical ideas, understand the children's capabilities, and understand children's common misconceptions and difficulties.

Topic 3: What are the goals of the *BMLK* curriculum for this unit? This part of the workshop included a discussion of both children's informal knowledge of the mathematics and the mathematics of the unit itself. The mathematical goals of the curricular unit were explicitly stated, then broken down into more specific goals in order to relate them to both the mathematics (topic 1) and children's mathematical thinking (topic 2).

Topic 4: What are the major activities of the *BMLK* curriculum for this unit? In the final part of the workshop, key curricular activities in the unit were discussed. This topic was explored by having teachers engage in role-playing, demonstrations, and discussions, and by examining video-clip examples of other teachers conducting the activities with children. The discussion continued to emphasize the mathematics, issues of pedagogy, methods of assessment, grouping children for instructional purposes, and materials needed for the activities. One purpose of these discussions was to help teachers understand the activities within the context of the earlier discussions of the mathematics itself and the research and theory relating to children's mathematical thinking and learning. A second purpose was to help teachers deal with practical aspects of preparing for and managing the activities in their own classrooms.

Measures

ECLS-B mathematics knowledge assessment. The primary outcome measure for the study is the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Direct Mathematics Assessment (Nejarian, Snow, Lennon, & Kinsey,2010). The U.S. Department of Education, National Center for Education Statistics (NCES) conducted the study and developed the child cognitive assessments. The ECLS-B assessments were designed to broadly survey children's mathematical knowledge and skills. Test specifications cover five content strands: (1) number sense, properties, and operations; (2) measurement; (3) geometry and spatial sense; (4) data analysis, statistics, and probabilities; and (5) patterns, algebra, and functions (Nejarian et.al., 2010). The mathematical domains covered on the ECLS-B are comprehensive, and map closely onto the mathematical content strands central to the *BMLK* curriculum. Nejarian and colleagues report that the reliabilities for the ECLS-B mathematics assessment are .89 and .92 for the prekindergarten and kindergarten years, respectively. In addition, they conducted gain score analysis that demonstrated the test is sensitive to growth in mathematics achievement, and a differential item functioning analysis that indicated there were no substantial differences in performance on the assessment for the major groups on which the test was normed and pilot-tested. Specifically, the differential item functioning analysis compared a group of white children with a number of other racial/ethic groups (black, Hispanic, and Asian) in an effort to determine if the assessment was biased (Nejarian et.al., 2010).

The ECLS-B mathematics assessment has a two-stage design in which a first-stage routing section is followed by one of several alternative second-stage forms that vary in difficulty. Children's performance on the routing section determined whether they received the low-, middle-, or high-difficulty second-stage option. The purpose of using this adaptive assessment design is to minimize administration time while maximizing the accuracy of the score produced by the assessment. ECLS-B scores are based on all of the assessment items administered (the first and second stages) and are calculated using Item Response Theory (IRT) procedures. IRT uses the pattern of correct and incorrect answers, along with the difficulty of each item, to place children on a continuous ability scale. As a result of this method, children do not have to answer the full set of test items, which would result in excessively long testing sessions involving many problems that are either too easy or too hard. In addition to reducing the testing burden on children, IRT scores can be compared across waves to determine a child's gain in achievement over time, even though the tests are not identical at each time point. Table A4 (Appendix) presents the means and standard deviations for each wave of assessment data for the sample as a whole and for the *BMLK* and BAU groups separately.

To administer the assessment for the study, the study team hired and trained assessors to administer the ECLS-B mathematics assessment. Each assessor was engaged in or had completed graduate study in fields related to early childhood (for example, early childhood education or developmental psychology). Assessors also attended a one-day training session that focused on administering the mathematics assessment to individual students. Researchers coordinated with center teachers and administrators to schedule testing days in order to minimize the extent to which students would miss activities (e.g., circle time, snacks, lunch, naptime, and outdoor activities). If a student was absent on the scheduled testing date, the study team returned as often as needed to administer the assessment. We excluded children from data collection if they had serious disabilities, did not speak English, or would not cooperate with the assessor.

However, these exclusions were rare and almost all students were eligible and able to complete the assessment.

The ECLS-B is administered using a computer-assisted interview protocol that provides directions and records the results. The testers administered the assessment items one-on-one to children in either a quiet area of the classroom or a nearby space. Testers presented questions to children in one of two formats, either by referring to a set of pictures on an 8-inch by 8.5-inch easel placed in front of the child or by providing manipulatives for the child to use to answer the question. An example of a question using pictures presented on the easel is a page divided into four sections, each showing pictures of different numbers of soccer balls. The student is asked to select and point to the picture showing a given number of soccer balls. An example of a question using manipulatives is a small number of counting chips with which students could show the number *five*. The easel presented pictured answer choices, such as different objects or different numbers of objects, and children were asked to point to the answer choice that corresponded to the number they created with manipulatives. Each assessment session was untimed, but typically lasted approximately 15 minutes. At the conclusion of the assessment, children were allowed to choose a sticker as a reward.

Mathematics language. We developed a short measure to conduct a preliminary test of the hypothesis that children exposed to the *BMLK* curriculum would show evidence of better mathematical language skills than children not exposed to the *BMLK* curriculum during the final testing administration in the spring of the kindergarten year. The measure consisted of two items. (A third item was administered, but was eliminated because the ambiguity of the question resulted in a wide variety of responses that could not be usefully coded). For each item, a visual was prepared that portrayed a mathematical idea or phenomenon. Prompts were developed to have children identify the mathematical idea/phenomenon that was portrayed, and then provide justification for their response.

The first item showed a scalene right triangle with a diagonal orientation. The prompt asked the child to identify the shape. The second item showed a repeating color pattern of alternating red and yellow circles—three repetitions for a total of six circles. The prompt required the tester to name the colors in order while pointing to the circles, and then have the child identify what this sequence was called.

The language measure was administered immediately upon completion of the ECLS-B, in the same testing session. This measure was administered only during the last round of testing, at the end of the kindergarten year. There were a total of 342 students tested on this measure, including 157 control students and 185 treatment students. An additional 7 students took the ECLS-B assessment, but did not progress on to take the language assessment due to fatigue.

Two independent raters then coded all responses, with inter-coder reliability of 97%. Items were coded for accuracy of the mathematical label (in response to the initial prompt), and then accuracy of the justification. Justifications were coded as being accurate if the child explained the main ideas. Justifications were coded as incomplete or irrelevant if they were faulty or did not apply to the question. Answers could also be coded as having no mathematical justification at all.

Fidelity of implementation observation measure. A fidelity of implementation observation measure was created based on a content analysis of the *BMLK* curriculum. The measure included ratings for critical areas of implementation, specifically (1) content coverage, (2) quality of teacher directions, (3) student engagement, (4) materials, (5) vocabulary, and (6) a global rating of adherence to the intent of the activity. Trained observers visiting the treatment classrooms, who visited twice in the prekindergarten year and twice in the kindergarten year, completed this measure. Observers were trained on the measure content and completed reliability training at the beginning of the study. The results were used as a rough indicator of the teachers' ability to implement the curriculum as intended, as we did not have enough cases or an adequate distribution of scores to examine the relation between fidelity and student achievement. We were primarily interested in getting preliminary information concerning the difficulty of program implementation.

Child-level covariates. The data analysis for this study included a number of child-level demographic variables as covariates. As described above, child demographic data were collected from records maintained by ACS. The covariates included in the analyses were a child's age at the beginning of the prekindergarten year, gender, family monthly income, and a dichotomously coded variable indicating whether the child was black/not Hispanic or Hispanic/any race.

Results

We used latent growth modeling (LGM) to analyze the impact of the *BMLK* curriculum on children's mathematics learning over the course of the two-year study. Latent growth modeling estimates a growth trajectory for each child, producing an intercept and a slope, or rate of growth. For this analysis, we estimated the intercept to represent each child's initial ECLS-B scale score (prekindergarten fall) and the slope to represent the total growth in scores over the course of the study. Because children were clustered within childcare centers, it was necessary for us to conduct analyses that would take into account the non-independence of observations. One approach would have been to conduct a three-level analysis that would account for the fact that observations are nested both within children (the four time points) as well as within childcare centers. However, because our sample size at the center level was 16, we did not have enough degrees of freedom to conduct a multi-level analysis. Instead, we conducted a TYPE=COMPLEX analysis using Mplus, which computes the standard errors taking into account the non-independence of observations, but did not allow us to include any center-level variables in the analysis.

Our longitudinal study involves missing scores, a situation that can create problems for statistical analysis and inference, including introducing bias, influencing the estimation of model parameters, and effecting statistical analysis and inference (Enders, 2010; Schafer & Graham, 2002). We utilized full information maximum likelihood (FIML) using Mplus (Mplus Software, Version 5, 2007), which yields efficient and consistent estimates in the presence of data that are missing at random (Enders, 2010; Schafer, 1999).

To determine post hoc the statistical power of our analysis, we calculated the minimum detectable effect (MDE), which is the smallest true effect that has an 80% chance of being found to be statistically significant at the $p < 0.05$ level for a two-tailed test. When a minimum effect size is expressed in standard deviation units, it is referred to as a minimum detectable effect size (MDES). An MDES of 0.25 is considered necessary for the effect of an intervention to have an "educational significance" (Bloom, Hill, Black, & Lipsey, 2008).

We used *Optimal Design for Multi-level and Longitudinal Research* power analysis software (Spybrook, Bloom, Congdon, Martinez, & Raudenbush, 2011) to conduct the power analysis. In addition to the assumptions described above (i.e., 80% power and a statistical significance level of $p < .05$ for a two-tailed test), we conducted the power analysis to account for the following: The study randomly assigned 50% of the 16 recruited childcare centers to the *BMLK* group; an observed intra-class (ICC) correlation of .04; and an average of 30 students in each childcare center. This post hoc power calculation indicates that an analysis of the data collected in the prekindergarten year is powered to detect an MDES of 0.40.

Latent Growth Models

We began by conducting two preliminary analyses. The goal of the first analysis was to identify which child-level covariates should be included in the model, as well as whether they should be included as predictors for the intercept, the slope, or both. It was necessary for us to determine which of the theoretically important child-level covariates to include in the analysis, rather than including all of them, in light of the multi-level structure of the data. In multi-level analysis, the degrees of freedom available for the analysis is based on the number of clusters—in this case, the number of childcare centers (N = 16), not the number of children in the sample. As a result, in order to estimate the model, we conducted a set of preliminary analysis to determine which of the available covariates were significant predictors of the intercept or slope estimated in the LGM. To do this, we estimated a latent growth model that included all of the available, theoretically important child-level covariates, and determined that the intercept was significantly predicted by age (at the beginning of the prekindergarten year), gender, whether the child was Hispanic, and family income. Not significant were whether the child spoke English, household size, and the child’s attendance record. We also found that none of the child-level characteristics predicted the slope. We conducted a second preliminary analysis to determine whether treatment predicted the intercept, which would have indicated that the *BMLK* and the BAU groups had significantly different initial ECLS-B mathematics scores. The treatment indicator variable did not significantly predict the intercept, indicating that there was baseline equivalence for the *BMLK* and the BAU groups, and therefore we did not include the treatment indicator variable as a predictor of the intercept.

Impact of the *BMLK* intervention

The LGM used to estimate the impact of the *BMLK* intervention on children's mathematics scores regressed the intercept on four child-level covariates as predictors, as described above, and regressed the slope on the treatment indicator. The model indicated that the slopes for the ECLS-B mathematics scores of *BMLK* children were significantly larger than the slopes for BAU children, such that the average increase in mathematics scores for children in *BMLK* classrooms was 1.8 points higher than BAU classrooms ($Slope_{BMLK} = 18.5$, $Slope_{BAU} = 16.7$; $B_{Treatment} = 1.8$, $p < .05$). This difference represents a medium effect size, *Cohen's d* = 0.40, and can be interpreted as the equivalent of an additional 1.6 months of instruction.

We conducted a regression analysis to determine if the difference between the *BMLK* and the BAU groups had emerged by the end of the prekindergarten year (using TYPE=COMPLEX to account for the fact that children were nested in childcare centers). The regression analysis included the same child-level covariates as the LGM, prekindergarten fall ECLS-B mathematics scores, and the treatment indicator, and found that treatment was not a significant predictor of children's ECLS-B mathematics scores at the end of prekindergarten, $B = 0.8$, $\hat{a} = .05$, p = n.s. In other words, the significant difference between the *BMLK* and BAU groups on the ECLS-B mathematics scores did not emerge until the kindergarten year.

Comparison of Sample Scores to National Sample

Table A4 shows the standardized t-scores for each group across time. The t-scores place the scores on a normal curve, so that 50 is the mean for the national sample at each time point. Both the treatment ($M_{BMLK} = 47.8$, $SD_{BMLK} = 7.8$) and control group ($M_{BAU} = 47.6$, $SD_{BAU} = 7.8$) scored below the national mean at the beginning of the prekindergarten year and above the national mean by the end of prekindergarten ($M_{BMLK} = 56.3$, $SD_{BMLK} = 7.6$ and $M_{BAU} = 55.2$,

SD_{BAU} = 7.5). A similar pattern was observed in the subsequent year, with the treatment (M_{BMLK} = 47.2, SD_{BMLK} = 6.1) and control groups (M_{BAU} = 46.7, SD_{BAU} = 6.7) scoring below the national mean in the fall of the kindergarten year and above it in the spring (M_{BMLK} = 54.5, SD_{BMLK} = 6.3 and M_{BAU} = 52.3, SD_{BAU} = 6.8).

Mathematics Language Measure

Our mathematics language measure consisted of two items, one involving triangles and the other patterns. In both cases, children's responses were categorized in terms of verbal *identification* and *justification*. Table A5 (Appendix) shows that a higher percentage of *BMLK* children (85%) than BAU children (72%) were able to identify triangles ($\chi^2 = 8.490$, $p < .004$), and that a significantly higher percent of *BMLK* children (55%) than BAU children (28%) were able to identify patterns ($\chi^2 = 24.548$, $p < .001$). In addition, a significantly higher percentage of *BMLK* children (34%) than BAU children (14%) were able to justify judgments concerning triangles ($\chi^2 = 19.507$, $p < .001$) by describing relevant properties. Both groups had approximately the same percentages of incomplete justifications (19% for BAU and 20% for *BMLK*).

In the case of pattern, a low percentage of children in each group (9% for BAU and 8% for *BMLK*) were able to offer accurate justifications, and the percentages of incomplete justifications were similar (64% for BAU and 72%—slightly, but not significantly, more—for *BMLK*).

Fidelity of Implementation Observation Measure

Observations showed that, in general, the *BMLK* teachers held quite closely to the curriculum. Table A6 (Appendix) is based on the composite scores for each observation on the fidelity measure, which included 6 items on a scale of 0–2, for a total possible composite score of 12. In Year 1, 69% of treatment teachers had high fidelity scores and 31% of the rest obtained medium fidelity. In Year 2, 85% of treatment teachers had high fidelity scores, 5% had medium fidelity scores, and 10% had low fidelity scores. However, although the curriculum does provide detailed descriptions of the activities, and often includes suggested questions or wordings, the activities are far from scripted. This leaves plenty of room for individual interpretation and variation, even when the teachers follow the general activity guidelines. Further, these differences in implementation could affect the quality of instruction.

In some cases, for example, observations showed teachers asking questions or using specific materials that were well aligned with the mathematical ideas being addressed. In other cases, however, the specific choices of questions or materials did not seem in line with the intent of the curriculum. Yet both groups were considered in compliance with the overall activity as described in the manual.

An example is a counting activity in which children were given zipper-locking bags, each labeled with a numeral between 0 and 5, and were instructed to count the appropriate number of counting objects into the bags. The curriculum guide lists some possible counting objects that teachers may use, such as counters, tiles, buttons, and plastic people, animals, and cars. Some teachers offered children a variety of plain objects to choose from, such as counting bears, beads, tiles, and so on. But other teachers offered items that were mathematically problematic. One teacher offered children a variety of colored buttons—red, blue, yellow, green, etc. The buttons were in different shapes, such as circles, squares, triangles, and butterflies, and had differing numbers of holes—two, three, or four. Some children were confused as to whether they were to count the buttons themselves or the holes in the buttons. Another teacher offered the children connecting cubes, some of which were already connected. One child was seen to count two connected cubes as one object, but the teacher counted it as two. In both of these cases, the objects selected by the teachers—though technically items that could be counted—presented features that were problematic for children as they attempted to count. Given these problematic features, and the resulting confusion, the quality of instruction was impaired, despite the literal fidelity to the curriculum.

The final rating on the fidelity measure—the global rating of adherence to the intent of the activity—went beyond the strict adherence of activity guidelines to evaluate the quality of these decisions as they related to the curricular intent. The global ratings then provided an indication not only of the strict fidelity of implementation, but also of the teacher's abilities to implement the curriculum effectively.

Table A7 (Appendix) shows the global ratings, which is the last of the six items on the fidelity of implementation measure with points ranging from 0–2. Of the 52 observations over the two years, all but four showed that teachers adhered to the lesson guidelines with medium-to-high fidelity on the global ratings question. In Year 1, 53% of treatment teachers had high fidelity scores on the global rating, 44% had medium fidelity scores, and 3% had low fidelity scores. In Year 2, 70% of treatment teachers had high fidelity scores on the global rating, 15% had medium fidelity scores, and 15% had low fidelity scores.

Discussion

Overview

The purpose of this study was to test the efficacy of the *Big Math for Little Kids* (*BMLK*) curriculum compared to a business-as-usual (BAU) comparison group in a randomized controlled trial. The study followed children from the beginning of their prekindergarten year through the end of kindergarten. Treatment teachers attended a professional development series designed specifically to help them implement *BMLK*, and control teachers were offered the same opportunity at the completion of the research study. Mathematics achievement was measured using a distal outcome measure, the ECLS-B, which is a nationally normed standardized measure. This measure was developed independently of the intervention and is therefore not directly or intentionally aligned with the mathematics presented in the curriculum. As a result, children in the *BMLK* group held no inherent advantage with regard to the outcome measure.

Conclusions

The primary research question was whether *BMLK* produces gains in achievement above and beyond business as usual in preschools for low-SES children. The program was designed to build upon the mathematical knowledge prekindergarten and kindergarten children bring to the classroom. The impact analysis shows that *BMLK* children learned more mathematics—the equivalent of approximately 1.6 months of instruction—and learned it at a faster pace than did children in the BAU control condition. In comparison to national norms, all the students in the sample began the academic year below the national mean and ended it above the national mean, showing that instruction in both the BAU and the *BMLK* groups is effective at improving students' mathematical achievement. All children did learn mathematics over the course of the year, but the children in the BMLK group learned significantly more mathematics, which demonstrates the overall efficacy of *BMLK*.

Implications

It is important to note that our use of the ECLS-B as the outcome measure is unique among other efficacy studies of preschool mathematics curricula in that it was not developed by the curriculum developers or in direct relation to the curriculum (Ginsburg, Lewis, & Clements, 2008). Our decision to use the ECLS-B as the outcome measure was based on the fact that it is a nationally norm-referenced assessment with strong psychometric properties. It is our opinion that using the ECLS-B as the outcome measure makes this study a very rigorous evaluation of the curriculum's effectiveness. In light of the fact that we used an outcome measure that was not developed in conjunction with the mathematics curriculum, we believe that the treatment's effect size (*Cohen's d* = 0.40) clearly demonstrates the effectiveness of *BMLK* in increasing young children's mathematics knowledge.

While the curriculum's success in improving ECLS-B scores at a faster rate than the control group attests to the efficacy of *BMLK*, the study does not indicate how the curriculum impacts student understanding of key mathematical concepts. Although a distal indicator of mathematics achievement increases the rigor of the randomized control trial and provides a global measure of mathematics achievement, the ECLS-B may also miss the more nuanced improvements in mathematical knowledge and ability that are unique to the curriculum. A logical next step is to conduct further research to determine specifically how the curriculum and professional development support learning of specific mathematical skills. To get some insight into this issue, we attempted to capture one aspect of curriculum that we think may be important: mathematical language.

In order to determine whether the *BMLK* curriculum helped children improve their ability to use mathematical language, we administered two tasks in the spring of the kindergarten year as a pilot to determine whether there was any evidence to support the idea that the curriculum supports mathematical language development. A higher percentage of children in the *BMLK* group than in the BAU group were able to use the correct language for shapes and, most importantly, to justify their responses. These results suggest that *BMLK* may indeed promote effective use of descriptive mathematical language, as well as analytic language, as shown by the data on shape justification. (The negative result in the case of pattern justification may reflect the fact that this idea is notoriously difficult to define.) In brief, this study offers modest support for the proposition that *BMLK* children are more adept than BAU children in employing both mathematical words and justifications.

Finally, we investigated the extent to which treatment teachers implement the curriculum with fidelity after attending the professional development series designed specifically for this curriculum. Fidelity of implementation cannot be merely a mindless adherence to a set of guidelines. It should involve understanding of the content, how to teach it, and thoughtful application that follows the activity guidelines. Roughly half of the teachers demonstrated a level of fidelity that seemed to employ such thoughtful application. The other half typically held closely to the curriculum, but did not necessarily implement it in a way that was true to the intent of the curriculum. In brief, although observed fidelity was high, quality was variable. This fact helps to put the efficacy of *BMLK* in perspective. *BMLK* is a comprehensive curriculum that requires teacher training. *BMLK* worked reasonably well, even though about half of the teachers who had received professional development did not implement precisely as our model suggested.

We draw two conclusions from these results. First, although our evaluation of *BMLK* was not conducted under conditions of ideal implementation, the program performed reasonably well. Second, our findings suggest that *BMLK* professional development has to be improved, particularly when the teachers have had little experience or motivation in teaching mathematics and when they operate under the very difficult conditions of urban education. Providing extensive and effective professional development is one of the most important challenges of American education.

Overall, the *BMLK* curriculum made a difference in the mathematics competence of low-SES, minority, preschool children. The program helped them learn in an enjoyable, developmentally appropriate format. It is important to note that all students in the study improved over time compared to national norms, starting out below average and ending above average in both years. This lends support for the value of early math education overall, and *BMLK* specifically. Yet, children who received the BMLK curriculum learned significantly more.

There is mounting evidence that mathematics education for preschool students effectively supports their acquisition of key foundational content knowledge in a way that need not displace the active play and reading activities that also are critical in the prekindergarten and kindergarten years. Mathematics instruction can succeed in teaching mathematics in a developmentally appropriate manner, and can be effectively integrated into current educational efforts by introducing well-designed activities and teacher professional development. It is clear that young children can, and do, learn mathematics, and that teaching it is developmentally appropriate.

Further research is needed to better understand how specific mathematical skills are supported by the *BMLK* curriculum and how they relate to understanding of more advanced mathematical concepts later. Second, future research should seek to determine how to improve professional development in order to increase implementation fidelity and, ultimately, improve the program's efficacy. Third, research can be designed to determine whether the intervention has a differential effect for subgroups of children (i.e., whether the intervention works better for children at different levels of SES or of different genders). Finally, future research should focus on the long-term impacts of early mathematics learning experiences on subsequent mathematics achievement in elementary school and beyond. Since achievement differences tend to increase over time, quality early intervention is key to addressing problems of mathematics achievement in our schools.

References

Arnold, D. H., & Doctoroff, G. L. (2003). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology, 54*, 517–545.

Baroody, A. J., Lai, M.-L., & Mix, K. S. (2006). The development of young children's early number and operation sense and its implications for early childhood education. In B. Spodek & O. Saracho (Eds.), *Handbook of research on the education of young children* (pp. 187–221). Mahwah, NJ: Lawrence Erlbaum Associates.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). *Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions* (Working Paper). New York, NY: MDRC.

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

Brush, L. R. (1978). Preschool children's knowledge of addition and subtraction. *Journal for Research in Mathematics Education, 9*, 44–54.

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology, 37*(2), 231–242.

Casey, B., Kersh, J. E., & Young, J. M. (2004). Storytelling sagas: an effective medium for teaching early childhood mathematics. *Early Childhood Research Quarterly, 19,* 167–172.

Chau, M., Thampi, K., & Wight, V. R. (2010). *Basic facts about low-income children, 2009: Children under age 6*. New York: National Center for Children in Poverty.

Clements, D. H. (1999). Geometric and spatial thinking in young children. In J. V. Copley (Ed.), *Mathematics in the early years* (pp. 66–79). Reston, VA: National Council of Teachers of Mathematics.

Clements, D. H., & Sarama, J. (2007). Early childhood mathematics learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 461–555). Charlotte, NC: Information Age Publishing.

Cross, C. T., Woods, T. A., & Schweingruber, H. (Eds.). (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academy Press.

Dodge, D. T., Colker, L. J., & Heroman, C. (2002). *The creative curriculum for preschool* (4th ed.). Washington, DC: Teaching Strategies, Inc.

Douglas-Hall, A.,Chau, M. & Koball, H. (2006, September). *Basic facts about low-income children: Birth to age 18*. New York, NY: Columbia University, National Center for Children in Poverty (NCCP).

Dowker, A. (2005). *Individual differences in arithmetic: Implications for psychology, neuroscience and education*. Hove, UK: Psychology Press.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446.

Enders, C. K. (2010) *Applied missing data analysis*. New York: The Gilford Press.

Entwisle, D. R., & Alexander, K. L. (1989). Early schooling as a “critical period“ phenomenon. In K. Namboodiri and R. G. Corwin (Eds.), *Sociology of Education and Socialization, 8*, 27–55.

Ertle, B. B., Ginsburg, H. P., Cordero, M. I., Curran, T. M., Manlapig, L., & Morgenlander, M. (2008). The essence of early childhood mathematics education and the professional development needed to support it. In A. Dowker (Ed.), *Mathematical difficulties: Psychology, neuroscience and Interventions* (pp. 59–83). Amsterdam, Netherlands: Elsevier.

Gelman, R., & Gallistel, C. R. (1986). *The child’s understanding of number*. Cambridge, MA: Harvard University Press.

Ginsburg, H. P. (2006). Mathematical play and playful mathematics: A guide for early education. In D. Singer, R. M. Golinkoff, & K. Hirsh-Pasek (Eds.), *Play = Learning: How play motivates and enhances children’s cognitive and social-emotional growth* (pp. 145–165). New York, NY: Oxford University Press.

Ginsburg, H. P., Cannon, J., Eisenband, J. G., & Pappas, S. (2006). Mathematical thinking and learning. In K. McCartney & D. Phillips (Eds.), *Handbook of early child development* (pp. 208–229). Oxford, UK: Blackwell.

Ginsburg, H. P., Greenes, C., & Balfanz, R. (2003). *Big math for little kids.* Parsippany, NJ: Dale Seymour Publications, Pearson Learning Group.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report—Giving Child and Youth Development Knowledge Away, 22*(1), 1–24.

Ginsburg, H. P., Lewis, A. E., & Clements, M. (2008, October). *School readiness and early childhood education: What can we learn from Federal investments in research on mathematics programs?* Paper presented at Working Meeting on Recent School Readiness Research: Guiding the Synthesis of Early Childhood Research, Washington, DC.

Ginsburg, H., & Pappas, S. (2004). SES, ethnic, and gender differences in young children's informal addition and subtraction: A clinical interview investigation. *Journal of Applied Developmental Psychology, 25*, 171–192.

Ginsburg, H. P., Pappas, S., & Seo, K.-H. (2001). Everyday mathematical knowledge: Asking young children what is developmentally appropriate. In S. L. Golbeck (Ed.), *Psychological perspectives on early childhood education: Reframing dilemmas in research and practice* (pp. 181–219). Mahwah, NJ: Lawrence Erlbaum Associates.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., and Brenwald, S. (2008). *Highlights From TIMSS 2007:Mathematics and Science Achievement of U.S. Fourth- and Eighth-Grade Students in an International Context* (NCES 2009–001 Revised). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Gormley, W. T. (2007). Early childhood care and education: Lessons and puzzles. *Journal of Policy Analysis and Management, 26*(3), 633–671.

Greenes, C., Ginsburg, H. P., & Balfanz, R. (2004). Big math for little kids. *Early Childhood Research Quarterly, 19*(1), 159–166.

Griffin, S. (2007). *Number Worlds: A mathematics intervention program for grades preK–6*. Columbus, OH: SRA/McGraw-Hill.

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics, 54*(1), 3–56.

Jordan, N. C., Glutting, J., Ramineni, C., & Watkins, M. W. (2010). Validating a number sense screening tool for use in kindergarten and first grade: Prediction of mathematics proficiency in third grade. *School Psychology Review*, *39*, 181–195.

Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1994). Assessing early arithmetic abilities: Effects of verbal and nonverbal response types on the calculation performance of middle- and low-income children. *Learning and Individual Differences, 6*(4), 413–432.

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

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

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology, 42*(1), 59–69.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*(5), 178–181.

Lee, V. E., & Burkham, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

Lemke, M., & Gonzales, P. (2006). *U.S. student and adult performance on international assessments of educational achievement: Findings from the condition of education 2006.* Washington, DC: National Center for Educational Statistics.

Ludwig, J. & Phillips, D. A. (March, 2007). *The benefits and costs of Head Start* (NBER Working Paper 12973). . Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w12973.pdf

Miller, K. F., & Parades, D. R. (1996). On the shoulders of giants: Cultural tools and mathematical development. In R. J. Sternberg & T. Ben-Zeev (Eds.), *The nature of mathematical thinking* (pp. 83–117). Mahwah, NJ: Lawrence Erlbaum Associates.

Mplus (Version 5) [Computer software]. Los Angeles, CA: Muthén & Muthén. Retrieved from http://www.statmodel.com/

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*. Chestnut Hill, MA: Boston College.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1998). *Mathematics and science achievement in the final years of secondary school.* Chestnut Hill, MA: Boston College.

Najarian, M., Snow, K., Lennon, J., and Kinsey, S. (2010). *Early childhood longitudinal study, birth cohort (ECLS-B), preschool–kindergarten 2007 psychometric report* (NCES 2010-009). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Assessment of Educational Progress (2011). *The Nation's Report Card: Mathematics 2011* (NCES 2012–458). Washington, DC: Institute of Education Sciences, U.S. Department of Education,

National Association for the Education of Young Children (NAEYC) and National Council of Teachers of Mathematics (NCTM). (2002, April; updated 2010). Early childhood mathematics: Promoting good beginnings. (Position statement). Retrieved from NAEYC website: www.naeyc.org/files/naeyc/file/positions/psmath.pdf

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards: Mathematics standards*. Retrieved from Common Core State Standards Initiative website: http://www.corestandards.org/the-standards

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

Pappas, S., Ginsburg, H. P., & Jiang, M. (2003). SES differences in young children's metacognition in the context of mathematical problem solving. *Cognitive Development, 18*(3), 431–450.

Pianta, R. C., & La Paro, K. (2003). Improving early school success. *Educational Leadership, 60*(7), 24–29.

Pianta, R. C., Barnett, W. S., Burchinal, M., & Thornburg, K. R. (2009). The effects of preschool education: What we know, how public policy is or is not aligned with the evidence base, and what we need to know. *Psychological Science in the Public Interest, 10*(2), 49–88.

Rathbun, A., West, J., & Hausken, E. G. (2004). *From kindergarten through third grade: Children's beginning school experiences.* (NCES 2004-007). Washington, DC: U.S. Department of Education, NCES.

Reynolds, A. J. & Ou, S. (2003). Promoting resilience through early childhood intervention. In S. S. Luthar (Ed.), *Resilience and vulnerability: Adaptation in the context of childhood adversities.* New York: Cambridge University Press.

Rudd, L. C., Lambert, M. C., Satterwhite, M., & Zaier, A. (2008). Mathematical language in early childhood settings: What really counts? *Early Childhood Education Journal, 36*(1), 75–80.

Seo, K.-H., & Ginsburg, H. P. (2004). What is developmentally appropriate in early childhood mathematics education? Lessons from new research. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 91–104). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schafer, J. L. (1999) Multiple imputation: A primer. *Statistical methods in medical research*, *8,* 3–15.

Schafer, J. & Graham, J. W. (2002). Missing data: Our view of state of the art. *Psychological Methods, 7*(2)*,* 147–177.

Siegler, R. L. (2010, March). *Teaching the meaning of numbers*. Paper presented at the Teacher's College, Columbia University, Annual Tisch Lecture, New York, NY.

Sophian, C. (2004). Mathematics for the future: Developing a Head Start curriculum to support mathematics learning. *Early Childhood Research Quarterly, 19*, 59–81.

Spybrook, J. Bloom, H., Congdon, R., Hill, C., Martinez, A., & Raudenbush, S. (2011). *Optimal design for longitudinal and multilevel research: Documentation for the "Optimal Design" software (Version 3.0)*. Retrieved from http://sitemaker.umich.edu/group-based/optimal_design_software.

Starkey, P. & Klein, A. (2008). Sociocultural influences on young children's mathematical knowledge. In *Contemporary perspectives on mathematics in early childhood education* (pp. 253–276). Charlotte, NC: Information Age Publishing.

Stevenson, H., Lee, S. S., & Stigler, J. (1986). The mathematics achievement of Chinese, Japanese, and American children. *Science, 56*, 693–699.

Young-Loveridge, J. M., (2004). Effects on early numeracy of a program using number books and games. *Early Childhood Research Quarterly, 19*, 82–98.

Appendix

Table A1. *Comparison of the Big Math for Little Kids curriculum and the Creative Curriculum*

	Big Math (Greenes, Ginsburg, & Balfanz, 2004)	Creative Curriculum (Dodge, Colker, & Heroman, 2002)
Research Foundations	Piaget—children construct their own understanding of the world. Ginsburg & Baron—during preschool years children develop mathematical concepts, strategies, and skills that they are not yet able to put into words.	Piaget—children learn through hands-on exploration of concrete objects. Maslow & Erikson—"Basic Needs" Theories: Children learn when teachers create environments where children feel safe and emotionally secure with a sense of belonging. Children also learn when teachers provide activities that help them to feel competent, make decisions, and direct their own learning. Vygotsky—Children learn in community. Each member of the classroom is considered a learner and a teacher.
Role of Play and Environment	Children spontaneously engage in mathematics during ordinary play—pattern, shape, symmetry, magnitude comparison, and use of number. Children can learn much about math when given the opportunity to explore mathematical ideas in rich educational environments with adult guidance.	Children learn through 4 types of play—functional, constructive, dramatic, and rule-governed. Teachers provide new materials, interact with and extend children's ideas during play, as well as facilitate games that involve physical activity. Children learn when given opportunities to pursue talents and demonstrate strengths. Interest areas—blocks, dramatic play, table toys, art,

		sand/water, library, science, music/movement, cooking, computers, and outdoors—are provided to offer each child an opportunity to use his/her unique intelligences.
Guiding Principles	Play is not enough—children can go only so far on their own. This does not mean that we should deny children the opportunity to engage in unguided free play. Instead, it means we should teach them using a challenging and comprehensive math curriculum.	Children's physical, social/emotional, cognitive, and language development typically unfolds in a sequence of developmental milestones. Knowing where each child is developmentally helps teachers plan and individualize supports. The Creative Curriculum Developmental Continuum for Ages 3–5 lays out the progression of development in each developmental area in four stages, and provides goals and objectives.
Structure	Group instruction, small group teaching, and individual exploration.	Child-initiated learning, teacher scaffolding, and direct teaching.

Table A2. *Descriptive statistics for the sample overall and by treatment group*

	All		BMLK (treatment)		BAU (control)		Participated in pre-K and K		Participated pre-K only		Participated K only	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
BMLK group							148	55.2	184	48.7	54	46.6
BAU group							120	44.8	194	51.3	62	53.4
Gender (% female)	386	50.7	210	54.4	176	46.8	123	45.9	190	50.3	63	54.3
Race/ethnic group												
Black/not	318	51.7	145	47.2	173	56.2	130	56.3	143	53.2	45	47.4
Hispanic												
Hispanic/all races	277	45.0	149	48.5	128	41.6	101	43.7	126	46.8	50	52.6
Neither black nor Hispanic	20	3.3	13	4.2	7	2.3	3	1.3	16	5.6	1	1.0
Family structure												
Two-parent	46	6.4	25	6.8	21	6.0	18	6.9	22	6.3	6	6.4
Single parent	671	93.5	341	93.2	330	94.0	243	93.1	328	93.7	100	94.3
Language spoken at home												
Speaks English	443	76.5	217	75.1	226	77.9	170	78.0	216	78.5	57	66.3

Speaks Spanish	136	23.5	72	24.9	64	22.1	48	22.0	59	21.5	29	33.7
	N	*Mean (SD)*	*N*	*Mean (SD)*	*N*	*Mean (SD)*	*N*	*Mean (SD)*	*N*	*Mean (SD)*	*N*	*Mean (SD)*
Age (mos.) [a]	720	51.8 (4.1)	366	51.7 (3.9)	354	51.9 (4.2)	261	52.3 (3.7)	352	51.6 (4.4)	107	51.0 (3.7)
Family income/month (monthly)	719	1355.5 (949.3)	367	1334.2 (956.0)	352	1377.9 (943.0)	261	1415.9 (940.0)	351	1300.9 (936.8)	107	(1387.6) (1009.3)
Household size	719	3.0 (1.0)	367	3.0 (1.0)	352	3.1 (1.0)	261	3.0 (1.0)	351	3.0 (1.0)	107	3.1 (0.9)

Notes.

[a] Mean age for children at the beginning of their prekindergarten year (whether or not they participated in the prekindergarten year of the study).

Table A3. *Descriptive statistics for teachers overall and by treatment group*

	All		BMLK (treatment)		BAU (control)	
	N	%	*N*	%	*N*	%
Female	42	89.4	21	91.3	21	87.5
Black/not Hispanic	25	53.2	11	47.8	14	58.3
Hispanic/all races	14	28.8	9	39.1	5	20.8
Other race/ethnic group	8	17.0	3	13.0	5	20.8
Primary language: English	33	71.7	16	69.6	17	73.9
Primary language: Spanish or other	13	28.4	7	30.4	6	26.1
Age: <= 40	14	29.8	9	39.1	5	20.8
Age: 41–50	15	31.9	7	30.4	8	33.3
Age: >= 51	18	38.3	7	30.4	11	45.8
Highest degree: AA	4	8.9	1	4.3	3	13.6
Highest degree: BA	18	40.0	10	43.5	8	36.4
Highest degree: MA	23	51.1	12	52.2	11	50.0
Teaching certificate	12	25.0	4	17.4	8	32.0

	N	*Mean (SD)*	*N*	*Mean (SD)*	*N*	*Mean (SD)*
Years experience (total)	44	14.9 (9.1)	21	12.2 (8.2)	23	17.4 (9.4)
Years experience (in an ACS childcare center)	47	11.0 (9.0)	23	8.0[a] (7.4)	24	13.8[a] (9.6)

Note

[a] $p < .05$

Table A4. *ECLS-B scores for sample and by treatment group: Scale scores and standardized scores*

	All	BMLK	BAU
IRT Scale Scores			
Pre-K fall	27.0	27.1	26.9
	(7.4)	(7.5)	(7.3)
Pre-K spring	35.3	35.8	34.7
	(7.9)	(8.0)	(7.7)
K fall	37.1	37.3	36.8
	(6.7)	(6.4)	(7.0)
K spring	44.0	45.2	42.8
	(7.2)	(6.9)	(7.3)
Standardized T-scores			
Pre-K fall	47.7	47.8	47.6
	(7.6)	(7.8)	(7.8)
Pre-K spring	55.8	56.3	55.2
	(7.5)	(7.6)	(7.5)
K fall	47.0	47.2	46.7
	(6.4)	(6.1)	(6.7)
K spring	53.4	54.5	52.3
	(6.6)	(6.3)	(6.8)

Notes.

IRT = Item response theory.

Table A5. *Mathematics language measure response accuracy*

	BAU (n = 157)	BMLK (n = 185)
Accurate Triangle Identification	72% (n=113)	85% (n=157)
Accurate Pattern Identification	28% (n=44)	55% (n=101)
Accurate Triangle Justification	14% (n=22)	34% (n=62)
Incomplete Triangle Justification	19% (n=30)	20% (n=37)
Accurate Pattern Justification	09% (n=14)	08% (n=14)
Incomplete Pattern Justification	64% (n=101)	72% (n=133)

Table A6. *Fidelity of implementation composite score by fidelity category*

	High Fidelity (9-12 points)	Medium Fidelity (5-8 points)	Low Fidelity (0-4 points)
Year 1	69% (n=22)	31% (n=10)	0% (n=0)
Year 2	85% (n=17)	5% (n=1)	10% (n=2)

Table A7. *Fidelity of implementation global rating scores*

	High Fidelity (2)	Medium Fidelity (1)	Low Fidelity (0)
Year 1	53% (n=17)	44% (n=14)	3% (n=1)
Year 2	70% (n=14)	15% (n=3)	15% (n=3)