STRENGTHENING CHILDREN'S MATH SKILLS WITH ENHANCED INSTRUCTION

The Impacts of Making Pre-K Count and High 5s on Kindergarten Outcomes

mdrc
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

Shira K. Mattera
Robin Jacob
Pamela A. Morris

March 2018

Strengthening Children's Math Skills with Enhanced Instruction

The Impacts of Making Pre-K Count and High 5s on Kindergarten Outcomes

Shira K. Mattera
Robin Jacob
Pamela A. Morris

March 2018

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY

This report is made possible through funding from the Robin Hood Foundation, the Heising-Simons Foundation, the Overdeck Family Foundation, and the Richard W. Goldman Family Foundation.

Dissemination of MDRC publications is supported by the following funders that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Annie E. Casey Foundation, Charles and Lynn Schusterman Family Foundation, The Edna McConnell Clark Foundation, Ford Foundation, The George Gund Foundation, Daniel and Corinne Goldman, The Harry and Jeanette Weinberg Foundation, Inc., The JPB Foundation, The Joyce Foundation, The Kresge Foundation, Laura and John Arnold Foundation, Sandler Foundation, and The Starr Foundation.

In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

For information about MDRC and copies of our publications, see our website: www.mdrc.org.

Overview

Early math skills are a strong predictor of later achievement for young children, not only in math, but in other domains as well. Exhibiting strong math skills in elementary school is predictive of later high school completion and college attendance. To that end, the Making Pre-K Count and High 5s studies set out to rigorously assess whether providing high-quality math instruction, aligned across prekindergarten (pre-K) and kindergarten, could lead to long-term gains across a variety of domains for students growing up in low-income communities in New York City.

In Making Pre-K Count, pre-K programs were randomly assigned to receive an evidence-based early math curriculum (Building Blocks) and associated professional development or to a pre-K-as-usual control condition. Pre-K in New York City changed rapidly during the study, with teachers overall conducting substantially more math than had previously been documented — a factor that may have played a role in the lack of impacts from Making Pre-K Count on children's math learning at the end of the pre-K year. In the High 5s study, students who had been in Making Pre-K Count program classrooms in pre-K were individually randomly assigned within schools in the kindergarten year to supplemental small-group math clubs, which took place outside of regular instructional time, or to a business-as-usual kindergarten experience. A companion report describes the High 5s program in more detail. This report focuses on the effects in kindergarten of the two math programs.

Key Findings

- **Making Pre-K Count:** At the end of kindergarten, there was a small, positive, but not consistently statistically significant effect for the Making Pre-K Count program on one of two measures of math skills, a measure that is more sensitive to children's skill levels than the more global test used in pre-K and kindergarten. Making Pre-K Count led to positive impacts on children's attitudes toward math at the end of kindergarten and to about two months' greater growth in kindergartners' working memory skills.

- **Making Pre-K Count plus High 5s kindergarten supplement:** Two years of aligned, enhanced math experiences led to positive impacts on the more sensitive measure of children's math skills, both above and beyond Making Pre-K Count alone (equivalent to 2.5 months' growth) and compared with no math enrichment in pre-K and kindergarten (equivalent to 4.2 months' growth); effects were positive but not statistically significant on the more global measure. The effect of two years of enhanced math translates into closing more than a quarter of the achievement gap between low-income children and their higher-income peers at the end of kindergarten. Children who were offered two years of math enrichment also had more positive attitudes toward math than children with no enrichment.

These findings suggest that early enriched math instruction, particularly when aligned across years, can have a positive effect on children's math skills, math attitudes, and working memory. The amount of math already in place was associated with the magnitude of the estimated effects of these programs. In addition, the sensitivity of the math measures used in the study may have played a role in how well each assessed math skills. The studies will continue to follow children into third grade to better understand the long-term effects of these early math programs.

Contents

List of Exhibits

Table

Table

Figure

Box

Preface

Making Pre-K Count and High 5s were developed in a partnership between the Robin Hood Foundation, the Overdeck Family Foundation, the Heising-Simons Foundation, and other funders and MDRC to tackle the question of how to improve the quality of early instruction in a meaningful and long-lasting way, and on a large scale. The approach consists of an evidence-based preschool math curriculum, with training and coaching for teachers, followed in kindergarten by small math clubs designed to align pre-K and kindergarten in content and pedagogy. The study asked important questions about scale, instructional alignment from year to year, and the power of math to change children's outcomes across multiple domains.

The findings show that the preschool and kindergarten math programs tested in this study can be widely implemented and that they can have a positive impact on children's outcomes compared with business as usual — even at a time when New York City was investing heavily in pre-K, and business as usual involved an emphasis on math. The combined effect of Making Pre-K Count and High 5s is meaningful and potentially trajectory-changing, comparable to closing more than a quarter of the achievement gap between low-income children and their higher-income peers in kindergarten. And both programs individually make a difference for children's outcomes, although the effects of the preschool math curriculum may have been tamped down by a combination of how much math was already being conducted in preschool and the measures available for that age group.

These results come at a time when investments in preschool are being hotly debated within the field. Some large, highly publicized studies have found few impacts of preschool when provided across children at all income levels, while other studies have calculated positive effects and large returns on the investment. As the debate rages, localities continue to move forward with funding and expanding preschool for their constituents. This study speaks to both the debate and those funding and implementing preschool by asking not whether preschool works but how the quality of preschool can be strengthened to ensure positive, meaningful experiences for children.

The Making Pre-K Count and High 5s findings suggest that aligned, developmentally informed math instruction bolsters children's early learning experiences. The study will continue to follow children into elementary school to explore whether these effects are maintained as children move into new instructional environments.

Gordon L. Berlin
President, MDRC

Acknowledgments

The Making Pre-K Count and High 5s studies were a joint effort with many collaborators. We thank the teachers, facilitators, and site administrators who helped make this study possible. Our thanks also to the children, families, and teachers who gave so generously of their time and cooperation, without whom the study would not have been possible.

Both programs could not have been run without the dedicated support and hard work of our partners at Bank Street College, who provided coaching, facilitation, and supervision across Making Pre-K Count and High 5s. We thank Katherine Baldwin and Josh Thomases for their guidance, thoughtful partnership, and review of the report. We also thank the current and former leadership at the New York City Department of Education's Division of Early Childhood Education and the Division of Child Care and Head Start at the Administration for Children's Services, who worked with us as this study unfolded. Additionally, this work could not have been possible without the deep dedication of the developers of Building Blocks, Drs. Doug Clements and Julie Sarama, who helped us support Building Blocks implementation and conceptualize the High 5s program, and who provided ongoing consultation. We also thank the teams of Building Blocks and New York City-based trainers, who trained teachers, and Anna Johnson and Kristi Hanby at the University of Michigan, who helped develop the High 5s intervention and training.

The execution of this study was made possible by our research partners at RTI International, including Jean Lennon, Jennifer Keeney, Joe Simpson, and the many dedicated data collectors. We also extend our appreciation to the steering committees for Making Pre-K Count and High 5s and our academic partners, Peg Burchinal of the University of North Carolina; Greg Duncan of the University of California, Irvine; Dale Farran of Vanderbilt University; and Christina Weiland of the University of Michigan.

The project benefited tremendously from the ongoing support and commitment of the Robin Hood Foundation, the Heising-Simons Foundation, the Overdeck Family Foundation, and the Richard W. Goldman Family Foundation.

At MDRC, we thank Gordon Berlin, JoAnn Hsueh, Sharon Huang, John Hutchins, Rob Ivry, Ginger Knox, Michelle Maier, and Marie-Andrée Somers for their guidance and review of the report. We thank the operations and technical assistance team for their tireless work providing site support and technical assistance. We thank the dedicated data team that made the many analyses possible, particularly Desiree Alderson and Lyndsay McDonough for their indispensable data management and Alvin Christian, Ezra Cohn, Adam Greeney, Anne Kou, Aimee Mun, Mirjana Pralica, and Max Snyder for their work on programming and analyses. Thanks to Mervett Hefyan for providing research assistance and coordinating the production of the report; Jennie Kaufman, who ably edited the report; and Carolyn Thomas, who prepared it for publication.

The Authors

Executive Summary

Recent research suggests that early math proficiency is one of the strongest predictors of later achievement for young children, not only in math, but in other domains as well.[1] Enhancing early math instruction, therefore, is a promising area for supporting low-income children's long-term outcomes. Moreover, aligning instruction from year to year may help sustain or even enhance the impacts from early math enrichment.[2] To that end, the Making Pre-K Count and High 5s studies set out to rigorously assess whether providing aligned, high-quality math instruction during prekindergarten (pre-K) and kindergarten could provide a critical boost that would lead to long-term achievement gains across a variety of domains.

Specifically, the studies assessed the impact of two programs on short- and longer-term child outcomes, including math, language, and executive function, for students growing up in low-income communities in New York City.[3] In the Making Pre-K Count study, pre-K programs in both public schools and community-based organizations were randomly assigned either to receive an evidence-based early math curriculum (Building Blocks) and associated professional development or to a pre-K-as-usual control condition. In the High 5s study, students who had been in Making Pre-K Count program classrooms in pre-K (in the public school sites) were individually randomly assigned within schools to small-group supplemental math clubs, which met outside of regular instructional time, or to a business-as-usual kindergarten experience. High 5s was designed to sustain the gains from the pre-K program and build on the same developmental trajectories and approach to learning that formed the basis of the Building Blocks program.

The studies were developed as part of the Robin Hood Early Childhood Research Initiative, which was established to identify and rigorously test promising early childhood interventions. That initiative is a partnership between Robin Hood, one of New York City's leading antipoverty organizations, and MDRC, a nonprofit, nonpartisan education and social policy

---

[1]Greg J. Duncan, Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, and Jeanne Brooks-Gunn, "School Readiness and Later Achievement," *Developmental Psychology* 43, no. 6 (2007): 1428-1446; Greg J. Duncan and Katherine J. Magnuson, "The Nature and Impact of Early Skills, Attention, and Behavior" (paper presentation, Russell Sage Foundation Conference on Social Inequality and Educational Outcomes, New York City, 2009).

[2]Valerie E. Lee and Susanna Loeb, "Where Do Head Start Attendees End Up? One Reason Why Preschool Effects Fade Out," *Educational Evaluation and Policy Analysis* 17, no. 1 (1995): 62-82.

[3]Executive function consists of the set of skills underlying children's ability to regulate themselves, their behavior, and their emotions. In early childhood, executive function skills include the ability to stop a primary response in favor of a more appropriate response (inhibitory control), the ability to shift attention and thinking from one rule or topic to another (cognitive flexibility), and the ability to manipulate small amounts of information (working memory).

research organization. Its flagship project, Making Pre-K Count and High 5s, was conducted in collaboration with Bank Street College of Education and RTI International and supported with lead funding from the Heising-Simons Foundation, the Overdeck Family Foundation, and the Richard W. Goldman Family Foundation. This is one of four reports based on this set of studies.

Findings from the end of the pre-K period show that Making Pre-K Count was implemented with fidelity to the program model.[4] The program had modest, positive impacts on the amount and quality of teacher math instruction at the end of the pre-K year. Program teachers conducted an additional 12 minutes of math per morning, although, notably, there was already a large amount of math — 35 minutes per morning — taking place in New York City pre-K classrooms. Despite the impacts on teachers' math practices, and perhaps because of the relatively high degree of math in pre-K-as-usual classrooms, there were no consistent effects on child outcomes at the end of pre-K. High 5s was also implemented with strong fidelity to the program model (as discussed in a companion report being published concurrently).[5]

This report is designed to explore the impact of Making Pre-K Count and High 5s on children's math skills, language skills, executive function, and attitudes toward math at the end of kindergarten. The main findings for outcomes at the end of kindergarten are as follows:

- **Making Pre-K Count:** At the end of kindergarten, small, positive, but not consistently statistically significant impacts were found on one of two measures of math skills for the Making Pre-K Count program. Making Pre-K Count led to positive impacts on children's attitudes toward math and to about two months' greater growth in kindergartners' working memory skills compared with the skills of kindergartners who did not receive enriched math in pre-K. Making Pre-K Count did not have statistically significant impacts on children's language or inhibitory control skills.

- **Making Pre-K Count plus High 5s kindergarten supplement:** At the end of kindergarten, two years of aligned, enhanced math led to positive and statistically significant impacts on one of two measures of math skills for students, both above and beyond Making Pre-K Count alone (equivalent to 2.5 months of growth) and compared with no math enrichment in either pre-K or kindergarten (equivalent to 4.2 months of growth). The effect of two years of enhanced math experiences translates into closing more than a quarter of the

[4]Pamela A. Morris, Shira K. Mattera, and Michelle F. Maier, *Making Pre-K Count: Improving Math Instruction in New York City* (New York: MDRC, 2016).

[5]Robin Jacob, Anna Erickson, and Shira K. Mattera, *Launching Kindergarten Math Clubs: The Implementation of High 5s in New York City* (New York: MDRC, 2018).

achievement gap between low-income children and their higher-income peers at the end of kindergarten.[6] Children who were offered two years of enhanced math experiences had more positive attitudes toward math than children with no enrichment. However, the programs did not have statistically significant impacts on children's language or executive function skills, or on the global measure of math skills that was likely influenced by children's language skills.

The size of the estimated effects from the pre-K program may have been due in part to the large amount of math already taking place in New York City and the sensitivity of the measures used to assess children's math skills. One of the math measures used in kindergarten was more sensitive to children's skill levels than the more global test used in pre-K and kindergarten. The sensitivity of that measure was due to a number of factors, including its depth and breadth of math skill assessment; it was not due to overalignment with the math programs.

The Math Programs

The Making Pre-K Count program provided pre-K teachers in New York City with a high-quality math curriculum (Building Blocks) and ongoing teacher training and coaching over two years. Building Blocks is a 30-week, evidence-based curriculum designed to take into account children's natural developmental progression in math skills.[7] Teachers received 11 days of training over the two years and met with coaches two to four times a month. In previous studies, Building Blocks has had positive impacts on teachers' math practices and children's preschool math outcomes in locales with little math instruction in place.[8]

---

[6]The effect sizes in this study are standardized measures of the difference in outcomes at the end of the kindergarten year for the control and program groups. The effect sizes were compared with standardized measures of the difference in outcomes at the end of kindergarten for children in the 90th income percentile and children in the 10th income percentile in the Early Childhood Longitudinal Study (ECLS) conducted in the 2010-2011 year (as described in Sean F. Reardon and Ximena A. Portilla, "Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry," *AERA Open* 2, no. 3 [2016]). In spring 2011, the income achievement gap was equivalent to 1.046 standardized units. The effect size of two years of enhanced math experiences in this study (0.30) is equivalent to 29 percent of that gap. The income achievement gap can also be calculated by comparing low-income children with their middle-income peers (a gap equivalent to 0.556 at kindergarten entry in 2010). The effect found here is equivalent to closing over half of the low-income–middle-income achievement gap.

[7]Douglas H. Clements and Julie Sarama, *Building Blocks: Teacher's Edition* (Columbus, OH: McGraw-Hill, 2013).

[8]Douglas H. Clements and Julie Sarama, "Effects of a Preschool Mathematics Curriculum: Summative Research on the Building Blocks Project," *Journal for Research in Mathematics Education* 38, no. 2 (2007): 136-163; Douglas H. Clements and Julie Sarama, "Experimental Evaluation of the Effects of a Research-Based Preschool Mathematics Curriculum," *American Educational Research Journal* 45, no. 2 (2008): 443-493; Kerry G. Hofer, Mark W. Lipsey, Nianbo Dong, and Dale C. Farran, "Results of the Early Math Project: Scale-Up Cross-Site Results," working paper (Nashville: Peabody Research Institute, Vanderbilt University, 2013).

The High 5s kindergarten supplement was developed to offer math enrichment in kindergarten to children who had received Making Pre-K Count in pre-K. A number of studies have shown that early impacts of preschool interventions fade out over time as children move into elementary school.[9] One hypothesis for this phenomenon is that the instruction that children receive in kindergarten and beyond is not well aligned, in instructional content or pedagogical approach, with the instruction they received in preschool.[10] Therefore, the High 5s program was designed to align with both the content and format of the activities students had been exposed to in pre-K. The program was developed by Robin Jacob and staff members at the University of Michigan with support from the developers of Building Blocks. Bank Street College of Education hired, trained, and supervised the facilitators, and MDRC provided ongoing operational support to both the participating schools and to Bank Street. The High 5s supplement paired three to four children with one facilitator for math clubs that met three times a week for 30 minutes each. The clubs offered math enrichment in a setting outside of regular classroom instruction, using engaging, developmentally appropriate activities. A companion report attests that the clubs were implemented well, with fidelity to the curricular materials.

Impacts of Early Math Programs on Kindergarten Outcomes

These two studies set out to explore whether providing young children with enhanced math instruction in early childhood could lead to learning gains that would translate into a sustained achievement boost across a variety of domains. As shown in Figure ES.1, analyses examine impacts at the end of kindergarten for three different comparisons:

- Comparison 1: Making Pre-K Count in pre-K versus no math enrichment
- Comparison 2: Making Pre-K Count plus the High 5s kindergarten supplement versus Making Pre-K Count only
- Comparison 3: Making Pre-K Count plus the High 5s kindergarten supplement versus no math enrichment in either pre-K or kindergarten

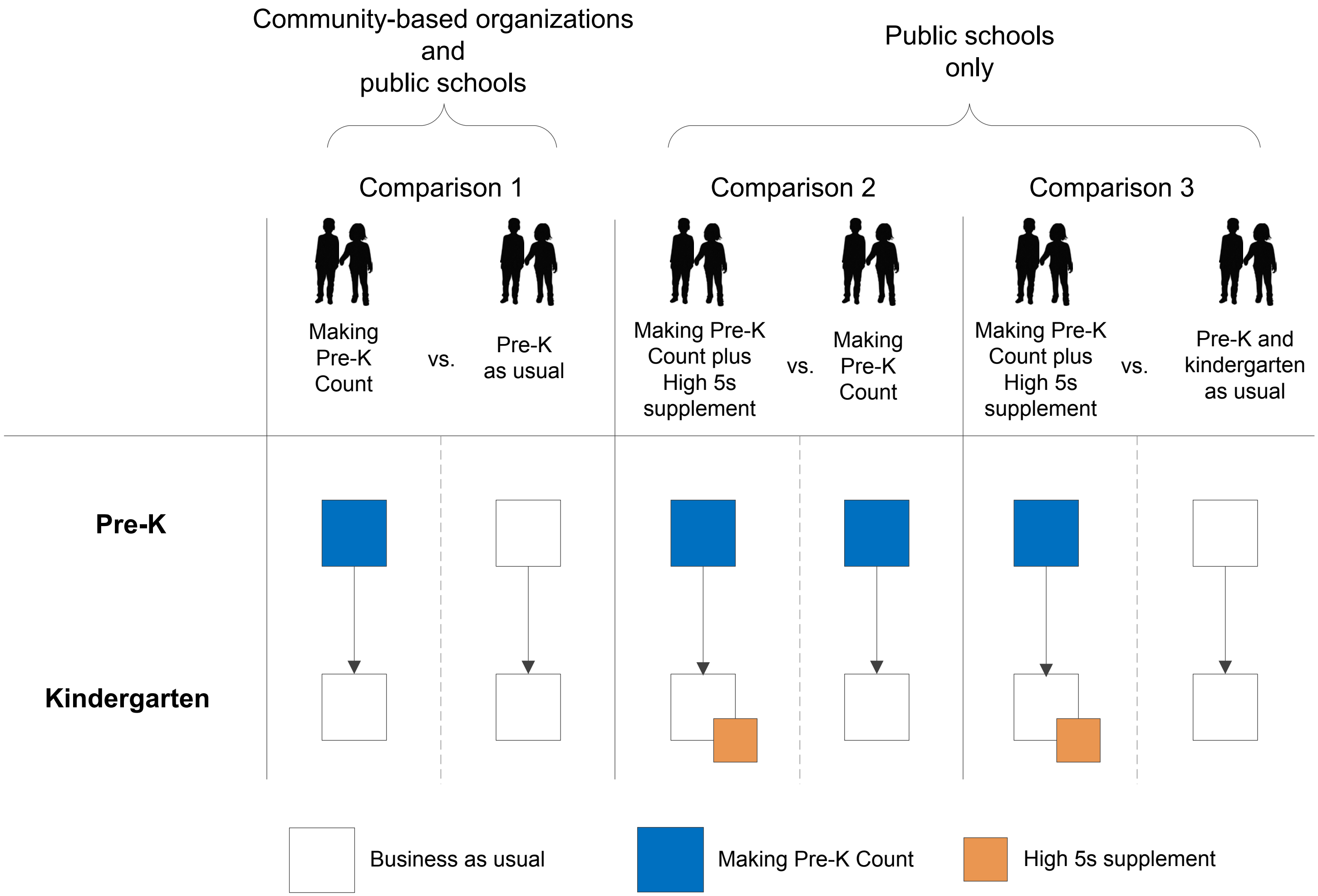
Math was assessed directly through two measures (the Research-Based Early Math Assessment–Kindergarten, or REMA-K, and the Woodcock-Johnson Applied Problems), which are described in greater detail in Box ES.1. Math attitudes were measured with an MDRC-developed question that drew on prior work: a show card displaying a range of five sad (1) to smiling (5) faces to describe how happy or unhappy math made children feel. Children's

[9]Mark W. Lipsey, Kerry G. Hofer, Nianbo Dong, Dale C. Farran, and Carol Bilbrey, *Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and First Grade Follow-Up Results from the Randomized Control Design* (Nashville: Peabody Research Institute, Vanderbilt University, 2013).

[10]Lee and Loeb, "Where Do Head Start Attendees End Up?"

Figure ES.1

What Are the Comparisons?
Impacts Measured at the End of Kindergarten



Box ES.1

Measuring Children's Kindergarten Math Skills in Making Pre-K Count and High 5s

Two measures were used to assess children's math skills in these studies:

- The Research-Based Early Math Assessment–Kindergarten (**REMA-K**) is a direct assessment used in kindergarten that measures thinking and learning related to a child's developmental progression along research-based mathematical learning trajectories. It was developed by the authors of the Building Blocks curriculum. These studies used a modified version of the assessment with a total of 48 items from the bank of over 100 items that constitute the full REMA.*
- The **Woodcock-Johnson Applied Problems** subscale, used in kindergarten and pre-K, is a standardized assessment of mathematical thinking for ages 2 through 90 that has been used widely in other studies of math interventions.†

What is the difference between these two measures? The REMA-K may be a more sensitive measure than the Woodcock-Johnson for a number of reasons. First, the REMA-K is a lengthier test that includes more items with the intention of assessing granular information about children's abilities in each mathematical skill, such as numbers or geometry, at each level of that skill. For example, the REMA-K includes nine questions about geometry or geometric measurement, while the Woodcock-Johnson has only three geometry items. Second, most items on the REMA-K assess only one specific skill at a time. Many of the Woodcock-Johnson items assess a mix of math content areas within one item. For example, one item on the Woodcock-Johnson asks children both to add and to understand coin value. Finally, the Woodcock-Johnson draws upon children's language skills in addition to math, as exemplified by items that ask children to count a certain animal in an array of animals. A child with strong language and math skills will do better on this test than one with commensurate math skills but for whom language and vocabulary may be lagging slightly.

Although the REMA-K was developed by the creators of Building Blocks, Douglas H. Clements and Julie Sarama, analyses suggest that the measure is not closely aligned with either of the math programs tested in this study. Ten percent of the items on the REMA-K (5 out of 48) look substantially similar in questions and materials to the types of activities to which children were exposed in Building Blocks or High 5s. (The High 5s program was not built off of Building Blocks activities and did not substantially align with Building Blocks.) Sensitivity analyses were conducted estimating program impacts on the REMA-K without the five more-aligned items; the magnitude and statistical significance of the effects stayed substantively the same without those items.

---

*Douglas H. Clements, Julie Sarama, and Xiufeng H. Liu, "Development of a Measure of Early Mathematics Achievement Using the Rasch Model: The Research-Based Early Maths Assessment," Educational Psychology 28, no. 4 (2008): 457-482.

†Richard W. Woodcock, Kevin S. McGrew, and Nancy Mather, Woodcock-Johnson III Tests of Achievement (Itasca, IL: Riverside, 2001).

language ability was measured using the Receptive One-Word Picture Vocabulary Test (ROWPVT-4), a standardized assessment of children's receptive vocabulary, or their ability to understand spoken language.[11] Executive function was measured two ways: with Hearts and Flowers, a computerized task that measures inhibitory control (the ability to stop an automatic response), and the Corsi Blocks task (backward span), which assesses short-term memory.[12]

Figure ES.2 presents the effects from each of these comparisons on children's math skills, attitudes toward math, language ability, and executive function in kindergarten.

- **One year of enhanced math in pre-K (Making Pre-K Count in pre-K) led to a small, positive, but not consistently statistically significant effect on one out of two measures of children's kindergarten math skills.**

The magnitude of the effect does not appear to be sensitive to how the measure or sample was specified, with impact estimates ranging in size from 0.08 to 0.12 standard deviations across different specifications. The estimates are statistically significant in three out of five specifications. These estimated effects are smaller than what has been found on a version of the REMA previously; in an earlier study of Building Blocks that took place in a similar context, researchers found an effect size of 0.19 standard deviations on the REMA at the end of kindergarten.[13] There was no effect on the Woodcock-Johnson Applied Problems, which is thought to be a less sensitive measure of children's math skills.

Pre-K enhanced math instruction led to more positive attitudes toward math and stronger working memory skills at the end of kindergarten than were found among children who had received pre-K as usual, equivalent to about two additional months of growth in working memory skills. The program had no impact on children's receptive language at the end of kindergarten.

High 5s math clubs took place in public school sites only. It was therefore important to examine the effect of Making Pre-K Count separately for those sites. In public school sites (Appendix Table G.1 in the full report), Making Pre-K Count had a positive effect on the more sensitive measure of children's kindergarten math skills, equal to nearly two months of

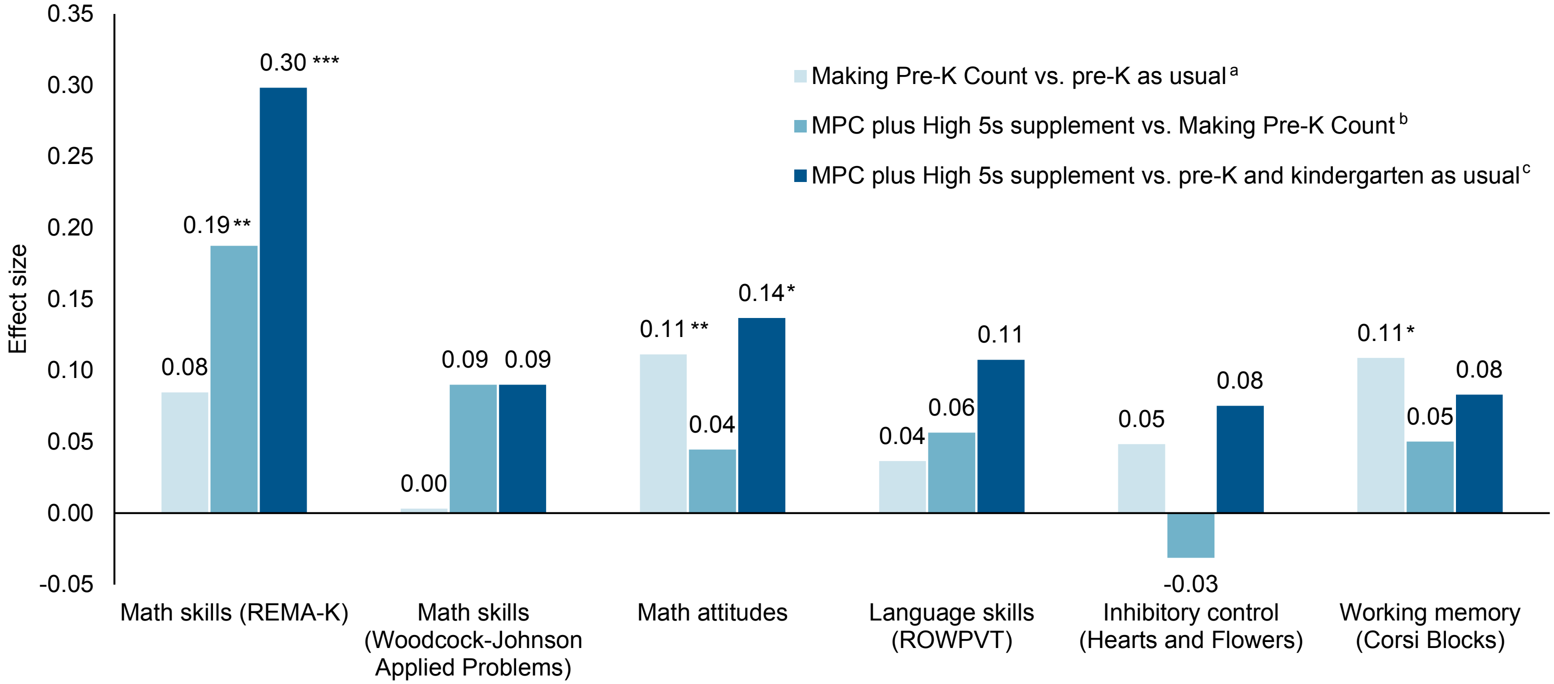
[11]Nancy A. Martin and Rick Brownell, *Receptive One-Word Picture Vocabulary Test*, 4th ed. (Novato, CA: Academic Therapy Publications, 2011).

[12]Andy Wright and Adele Diamond, "An Effect of Inhibitory Load in Children While Keeping Working Memory Load Constant," *Frontiers in Psychology* 5 (2014); Philip Michael Corsi, "Human Memory and the Medial Temporal Region of the Brain," PhD diss. (Montreal: McGill University, 1972); Muriel Deutsch Lezak, *Neuropsychological Assessment* (New York: Oxford University Press, 1983).

[13]Douglas H. Clements, Julie Sarama, Carolyn Layzer, Fatih Unlu, Carrie Germeroth, and Lily Fesler, "Effects on Mathematics and Executive Function Learning of an Early Mathematics Curriculum Synthesized with Scaffolded Play Designed to Promote Self-Regulation Versus the Mathematics Curriculum Alone" (unpublished paper, 2016), PDF.

Figure ES.2

Impacts in the Spring of the Kindergarten Year



SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[a]The Making Pre-K Count (MPC) program group received Making Pre-K Count in pre-K. The pre-K-as-usual control group did not receive math enrichment.

[b]The MPC plus High 5s supplement group received Making Pre-K Count in pre-K and High 5s in kindergarten. The Making Pre-K Count group received only Making Pre-K Count in pre-K. Both groups consist of public school children only.

[c]The MPC plus High 5s supplement group received Making Pre-K Count in pre-K and High 5s in kindergarten. The pre-K-and-kindergarten-as-usual control group did not receive math enrichment. Both groups consist of public school children only.

additional learning (effect size = 0.13 standard deviations), as measured at the end of kindergarten. The program also had a positive impact on children's math attitudes in kindergarten at public school sites. Therefore, High 5s math clubs took place on top of a documented impact of the Making Pre-K Count program on both math skills and math attitudes.

- **Compared with enhanced pre-K math only (Making Pre-K Count alone), the addition of the High 5s kindergarten supplement led to positive, statistically significant impacts on one of two measures of kindergarten math skills.**

These impacts for the High 5s kindergarten supplement translate into two and a half additional months of math learning (effect size = 0.19 standard deviations) on the REMA-K. There was no effect of the program on the Woodcock-Johnson Applied Problems, a global measure of math. There were no effects of the kindergarten math clubs on children's language, math attitudes, or executive function above and beyond the effects of Making Pre-K Count.

- **Two years of enhanced math instruction (Making Pre-K Count plus High 5s) led to positive impacts on one of two measures of children's math skills at the end of kindergarten, compared with no enhanced math instruction in pre-K or kindergarten.**

Two years of enhanced math experiences led to 4.2 additional months of math learning compared with business-as-usual math instruction in pre-K and kindergarten (effect size = 0.30 standard deviations) on the REMA-K. There were positive but not statistically significant effects on the Woodcock-Johnson assessment. The two years of math programming also led to more positive attitudes toward math compared with business-as-usual pre-K and kindergarten.

Contributing Factors

Some of the findings from this newest wave of analyses are consistent with expectations about the benefits of two years of math enrichment on children's math learning and attitudes. Findings from the enhanced pre-K math instruction, however, are inconsistent across pre-K and kindergarten, and the pattern of findings does not match prior expectations or most of the previous studies of the Building Blocks program. Prior Building Blocks studies have shown positive effects on pre-K teachers' math instruction, which have led to positive impacts on children's math skills at the end of pre-K (in the range of 0.47 to 1.47 standard deviations), with effects persisting — albeit with some fade-out in magnitude — into kindergarten.[14] In this study, the

[14]Clements and Sarama, "Effects of a Preschool Mathematics Curriculum"; Clements and Sarama, "Experimental Evaluation of the Effects of a Research-Based Preschool Mathematics Curriculum"; Hofer, Lipsey, Dong, and Farran, "Results of the Early Math Project."

positive effects on teachers' math practices in Making Pre-K Count did not lead to impacts on children's skills at the end of pre-K (with the exception of a small, positive impact on children's inhibitory control skills). Yet impacts of the program on both math attitudes and working memory skills emerged at the end of kindergarten. A number of differences between this study and prior studies suggest factors that may have contributed to this lack of alignment in findings:

- **Amount of math in the pre-K environment.** The small impacts that resulted from the pre-K program appear to have been due in part to the large amount of math instruction already occurring in New York City's pre-K programs. In earlier studies of Building Blocks, less math instruction took place in control group preschools — as low as 12 to 16 minutes per morning. Therefore, an additional 5 to 10 minutes of math may have added substantially more and different math content for children. In Making Pre-K Count, teachers in pre-K-as-usual classrooms were engaging in 35 minutes of math per morning on average. In this kind of environment, with teachers engaging in so much math content, even 12 additional minutes of math may not have substantially changed children's math experience. Exploratory analyses provide some support for this hypothesis; in Making Pre-K Count, impacts on children's pre-K math skills were larger in places where less math was taking place in pre-K-as-usual classrooms.

- **Measurement.** Math skills in kindergarten (and in prior Building Blocks studies) were measured using the REMA.[15] As discussed in Box ES.1, analyses demonstrate that the REMA-K is not highly aligned with the Building Blocks curriculum. Rather, the REMA-K proved to be a more sensitive measure of children's math skills, capturing more detailed information about each skill level, than either the Woodcock-Johnson assessment or the other measure of children's math skills used in pre-K. The more limited sensitivity in the pre-K measures may have contributed to the lack of measurable effects and therefore the lack of consistency with prior studies or with the kindergarten impacts.

- **Sample.** The sample in Making Pre-K Count also differed from study samples in prior Building Blocks studies. The Making Pre-K Count sample was more heavily Hispanic and entered pre-K with higher cognitive scores on

[15]Douglas H. Clements, Julie Sarama, Mary Elaine Spitler, Alissa A. Lange, and Christopher B. Wolfe, "Mathematics Learned by Young Children in an Intervention Based on Learning Trajectories: A Large-Scale Cluster Randomized Trial," *Journal for Research in Mathematics Education* 42, no. 2 (2011): 127-166; Douglas H. Clements, Julie Sarama, Carolyn Layzer, Fatih Unlu, Christopher B. Wolfe, Mary Elaine Spitler, and Daniel Weiss, "Effects of TRIAD on Mathematics Achievement: Long-Term Impacts" (paper presentation, Society for Research on Educational Effectiveness Spring Conference, Washington, DC, March 2-5, 2016).

average. While there was little evidence that either of these sample characteristics directly led to differentially small impacts, it is still possible that some combination of sample characteristics play a role in how the findings from this study align with prior work.

Implications

To date, the findings from these studies suggest a number of takeaways:

- **Early math enrichment appears to support children's math skills at the end of kindergarten.**

Two years of math enrichment via both Making Pre-K Count and High 5s had an impact equivalent to over four months of additional growth compared with pre-K and kindergarten as usual. This is equivalent to closing more than a quarter of the achievement gap between low-income children and their higher-income peers at the end of kindergarten. There was also some (although more limited) evidence that Making Pre-K Count alone had a small, positive impact on children's math skills, especially in the public school programs. The math enrichment children were offered in kindergarten via High 5s led to effects on math skills equivalent to about two and a half months over and above the impacts of Making Pre-K Count alone.

- **Sustained, aligned instruction in mathematics over multiple years is potentially important for children's math development.**

The alignment between pre-K and kindergarten math instruction may have been one factor that contributed to the impacts of the High 5s program. Small-group enrichment programs, like High 5s, may be one short-term approach to improving alignment in both content and pedagogy between early elementary school and what many children experience in pre-K.

- **Early math enrichment may have positive implications in domains beyond math.**

At the outset of this study, short-term improvements in math were hypothesized to spill over into other domains of children's learning, like executive function. Findings from this study demonstrate that a math intervention may have effects on other domains, such as working memory. These effects hold the potential for longer-term impacts, because executive function is hypothesized to be important for supporting longer-term effects and reducing fade-out of early childhood programs.[16]

---

[16]Susan E. Gathercole, Susan J. Pickering, Camilla Knight, and Zoe Stegmann, "Working Memory Skills and Educational Attainment: Evidence from National Curriculum Assessments at 7 and 14 Years of Age,"
(continued)

- **Places where a limited amount of math is being taught may benefit more from early math programs.**

Despite taking place in a shifting pre-K environment where the amount of math was growing from year to year, Making Pre-K Count led to an additional 12 minutes of math instruction per morning. Unlike prior studies, however, these 12 minutes of math did not lead to better outcomes for children at the end of pre-K. One of the reasons for the limited impacts observed at the end of pre-K may have been the relatively high amount of math taking place in pre-K-as-usual classrooms. This suggests that more high-quality math may matter most in an environment with very little math to begin with.

What's Next

Many questions remain about which impacts will be sustained into later elementary school for the children who participated in these two programs. The Making Pre-K Count and High 5s studies will uniquely be able to track the trajectory of effects across time and across domains. The studies will continue to follow children into third grade to better understand the long-term effects of these early enhanced math programs. These data will also help inform the field about the longitudinal relationships between early math measures and later outcomes for children.

These studies also raise questions about the timing of enhanced math experiences and the relative benefits of pre-K and kindergarten enrichment. It is unclear what the effect of kindergarten enrichment alone (without enhanced pre-K instruction) would have been or to what extent the enhanced math instruction in pre-K contributed to effects observed at the end of kindergarten.

At the same time, High 5s was a new program, and its theory of change is not yet well understood. Specifically, it is unclear whether the added math time, the wider math content, the instructional climate, or the individualized instruction — or some combination of the four — contributed to the impacts on children's math skills. Given the potential strength of High 5s, it is worth considering how it could serve as a model for integrating more small group work and hands-on learning opportunities into kindergarten math instruction in the classroom. Replication of the program with various adaptations (for example, in contexts where children do not enter kindergarten with a strong math background, or as a "push-in" or "pull-out" program slotted into the school day and run by paraprofessionals) could help identify the most important components of the program and identify contexts in which the program is most likely to be effective.

Applied Cognitive Psychology 18 (2004): 1-16; Megan M. McClelland, Alan C. Acock, Andrea Piccinin, Sally Ann Rhea, and Michael C. Stallings, "Relations Between Preschool Attention Span-Persistence and Age 25 Educational Outcomes," *Early Child Research Quarterly* 28, no. 2 (2013): 314-324; Jens Ludwig, "On What Should We Focus Our Early Interventions to Maximize Benefit-Cost Ratios?" (paper presentation, Robin Hood Early Childhood Institute planning meeting, New York City, January 2011).

Chapter 1

Introduction

Early math skills have been identified as potentially important predictors of children's later outcomes across a broad set of domains, including reading ability, high school completion, and college attendance.[1] Evidence from a set of rigorous studies that were led by program developers demonstrates that young children's math competencies can be improved by training prekindergarten (pre-K) teachers to deliver fun, interactive math curricula.[2] Short-term improvements in math could potentially affect other domains of children's learning, such as language and the skills underlying self-regulation, or what is known as executive function.[3]

However, few studies of pre-K math curricula follow children over the long term. Those that do often show a fade-out of math impacts as children move into elementary school.[4] In a prior seminal study of an early childhood intervention (Perry Preschool), initial impacts on children's cognitive outcomes faded by the time the children were in elementary school.[5] Long-term follow-up of the children showed a reemergence of impacts on other important outcomes in adulthood, such as high school graduation and employment.[6]

Various theories exist as to how or why effects from early childhood interventions are sustained or fade over time. Some have theorized that, along with the cognitive effects that are often measured in studies, early childhood programs may lead to impacts on "noncognitive" abilities — like executive function — that support learning across domains.[7] Noncognitive skills are associated with later outcomes for children and could contribute to a reemergence of program impacts, even after the original cognitive effects have faded out. Improvements in children's early math skills have been hypothesized to spill over into such important noncognitive domains as executive function, making math a potentially foundational outcome for

---

[1]Duncan et al. (2007); Duncan and Magnuson (2009).

[2]Clements and Sarama (2007, 2008); Clements et al. (2011); Hofer, Lipsey, Dong, and Farran (2013); Klein et al. (2008).

[3]Blair and Razza (2007); Ginsburg, Lee, and Boyd (2008). Executive function consists of the set of skills underlying children's ability to regulate themselves, their behavior, and their emotions. In early childhood, executive function skills include the ability to stop a primary response in favor of a more appropriate response (inhibitory control), the ability to shift attention and thinking from one rule or topic to another (cognitive flexibility), and the ability to manipulate small amounts of information (working memory).

[4]Clements et al., "Effects of TRIAD" (2016); Hofer, Lipsey, Dong, and Farran (2013).

[5]Weikart, Deloria, Lawser, and Wiegerink (1970).

[6]Schweinhart et al. (2005).

[7]Ludwig (2011).

sustaining effects into the longer term.[8] The shifting educational contexts children experience as they move from preschool to elementary school and into middle school and high school may also contribute to the stability or fade-out of effects from early programming.[9] Aligning children's instructional experiences as they move across school years may help maintain gains from pre-K. Building on these hypotheses, the Making Pre-K Count and High 5s studies set out to explore whether increasing students' exposure to high-quality math instruction that is aligned across pre-K and kindergarten could provide a critical boost that would lead to sustained long-term achievement gains across a variety of domains.

The Making Pre-K Count program began in fall 2013 and provided pre-K teachers in New York City with a high-quality math curriculum (Building Blocks) and ongoing teacher training and coaching. Building Blocks is a 30-week, evidence-based math curriculum designed to take into account children's natural developmental progression in math skills.[10] Previous studies of Building Blocks have shown positive impacts of the curriculum on teachers' math practices and children's preschool math outcomes in locales with little math instruction in place.[11] At the conclusion of this study in fall 2015, New York City began to roll out the Building Blocks program at pre-K sites throughout the city.

The High 5s program was developed to offer supplemental math enrichment outside of regular instructional time to kindergarten children who had received Making Pre-K Count in pre-K. High 5s was designed in an effort to sustain the gains from the pre-K program and built on the same development trajectories that formed the basis of the Building Blocks program. Both the content and format of the activities were aligned with what students had been exposed to in pre-K, upwardly extending children's numeracy, geometry, patterning, and measurement skills. The program was developed by Robin Jacob and staff members at the University of Michigan with input from Doug Clements and Julie Sarama, the developers of Building Blocks, and was administered by staff members from Bank Street College of Education. High 5s paired three to four children with one facilitator for math clubs that met three times a week for 30 minutes each. The clubs offered math enrichment in school but in a setting outside of regular classroom instruction — usually after or before school or during lunch.

The Making Pre-K Count and High 5s studies were designed to rigorously assess the impact of the two programs on short- and longer-term child outcomes, including math, language, and executive function skills, for students growing up in low-income communities in New York City. In the Making Pre-K Count study, pre-K programs in both public schools and

---

[8]Blair and Razza (2007); Ginsburg, Lee, and Boyd (2008).

[9]Lee and Loeb (1995).

[10]Clements and Sarama (2013).

[11]Clements and Sarama (2007, 2008); Hofer, Lipsey, Dong, and Farran (2013).

community-based organizations were randomly assigned either to receive the Building Blocks program and associated professional development or to continue with pre-K as usual. In the High 5s study, students who had been in Making Pre-K Count public school program classrooms in pre-K were individually randomly assigned within schools to participation in the High 5s program or to a business-as-usual kindergarten experience. The design of these two studies helps ensure that any impacts that are observed can be causally attributed to the programs themselves.

Making Pre-K Count and High 5s were developed as part of the Robin Hood Early Childhood Research Initiative, which was established to identify and rigorously test promising early childhood programs. The initiative is a partnership between Robin Hood, one of New York City's leading antipoverty organizations, and MDRC, a nonprofit, nonpartisan education and social policy research organization. Making Pre-K Count and High 5s, conducted in collaboration with Bank Street College of Education and RTI International, were also supported with lead funding from the Heising-Simons Foundation, the Overdeck Family Foundation, and the Richard W. Goldman Family Foundation.

Earlier findings (along with additional analyses presented in this report) demonstrated that Making Pre-K Count was generally implemented with fidelity to the program model. The program had positive impacts on the amount and quality of teacher math practices at the end of the pre-K year.[12] But at the time of the study, as described in more detail later, math instruction in New York City pre-K programs was changing substantially, and by the end of the last year of the study, teachers in pre-K classrooms were delivering a large amount of math. Therefore, the amount of math exposure students in pre-K-as-usual schools received was much larger than what was documented in prior studies of the Building Blocks program.[13] Despite the impacts on teachers' math practices, and perhaps because of the relatively high degree of math exposure in pre-K-as-usual classrooms, there were no consistent effects on child outcomes at the end of pre-K.

High 5s was also implemented with strong fidelity to the program model (as discussed in a companion report being published concurrently).[14] But that report finds that the typical kindergarten classroom math environment that children experienced was markedly different from the pre-K math environment. Observations of kindergarten classrooms showed that math instruction took place largely in whole-group settings or with children seated at desks and was relatively didactic in nature. This meant that High 5s, which involved small-

[12]Morris, Mattera, and Maier (2016).

[13]Clements and Sarama (2008); Clements et al. (2011).

[14]Jacob, Erickson, and Mattera (2018).

group instruction with hands-on activities outside of regular class time, provided a unique math experience for children.

The current report is designed to explore the impact of these two programs (Making Pre-K Count and High 5s) on children's math skills, language skills, executive function, and attitudes toward math at the end of kindergarten (see Chapter 3 for more information on the child outcome measures).[15] Although at the end of pre-K, the study found no effects of the program on children's math skills, strong positive impacts of the program were found at the end of kindergarten for students who were offered two years of enhanced math experiences via both Making Pre-K Count and High 5s. This report describes these end-of-kindergarten findings and then explores the potential reasons for the positive findings, specifically the amount of math already being conducted and the sensitivity of the measures used for assessing math. As shown in Figure 1.1, analyses examine the effects at the end of kindergarten of three different comparisons:

- Making Pre-K Count in pre-K compared with no math enrichment (that is, the effect of one year of math enrichment in pre-K)

- Making Pre-K Count plus the High 5s kindergarten supplement compared with Making Pre-K Count only (that is, the effect of adding an additional year of supplemental math enrichment in kindergarten)

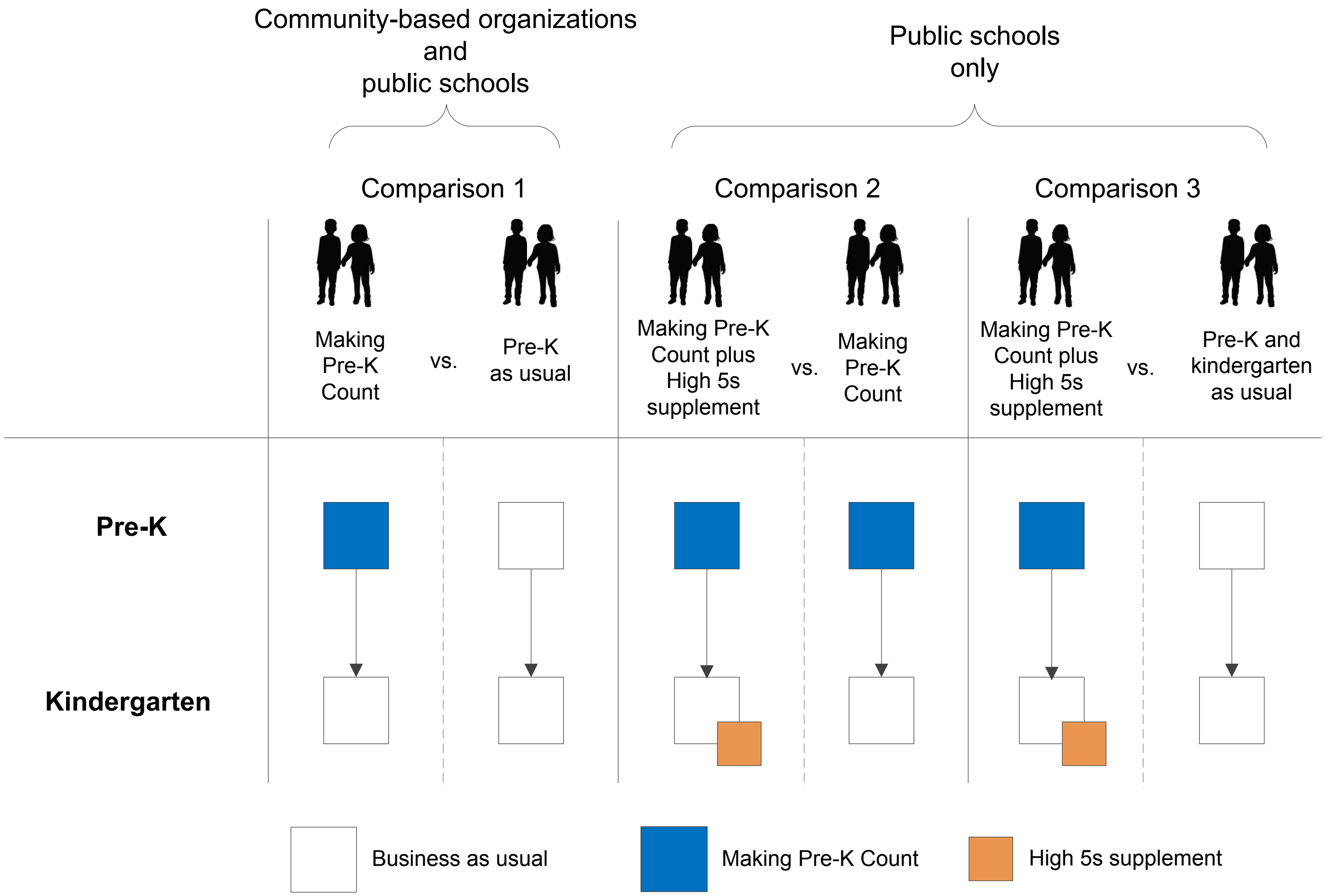
- Making Pre-K Count plus the High 5s kindergarten supplement compared with no math enrichment (that is, the effect of two years of math enrichment in pre-K and kindergarten)

Findings show that the Making Pre-K Count program alone (compared with pre-K as usual) led to small, positive, but not consistently statistically significant effects on one out of two measures of children's math skills and led to positive and statistically significant impacts on children's attitudes toward math and working memory. Two years of aligned, enhanced math experiences (Making Pre-K Count plus the High 5s kindergarten supplement) led to positive impacts on one of two measures of math achievement for students, above and beyond Making Pre-K Count alone or no math enrichment in pre-K and kindergarten. Neither program had statistically significant impacts on children's language or inhibitory control skills as measured in this study or on a global measure of math skills that simultaneously measured language skills.

[15]A preliminary report released in 2017 indicated positive effects of Making Pre-K Count and High 5s on children's math skills, attitudes toward math, and executive function skills at the end of kindergarten (Mattera and Morris, 2017). The analyses in that report included some children who had received only Making Pre-K Count in pre-K and some who had received both Making Pre-K Count in pre-K and High 5s in kindergarten, and compared them with children who had received no math enrichment in either pre-K or kindergarten. The report did not parse the separate contributions of Making Pre-K Count and High 5s to those impacts.

Figure 1.1

What Are the Comparisons?
Impacts Measured at the End of Kindergarten



The different math instruction that children experienced in pre-K and in kindergarten may have helped drive the pattern of effects across the programs. Although Making Pre-K Count had an effect on the amount of math teachers engaged in, the amount of math children were already experiencing each day (35 minutes per morning, on average) may have meant that the program's additional 12 minutes of math per morning was less of a contrast than in previous studies where teachers did not teach as much math.[16] In comparison, the High 5s clubs were substantially different from the kindergarten-as-usual experience, with children working in small groups with a facilitator and receiving focused attention and individualized instruction. In addition, measurement may have played a role, with the two assessments (the Woodcock-Johnson Applied Problems and the Research-Based Math Assessment–Kindergarten, or REMA-K) capturing different sets of skills. The REMA-K was not highly aligned with the math programs but was more sensitive to children's skill levels than the Woodcock-Johnson assessment.

It remains to be seen whether the effects from the pre-K and kindergarten programs will be sustained into later elementary school. Children participating in these studies will be followed into third grade to determine whether the impacts of these early math initiatives persist in the longer term. The findings for the new High 5s program are promising and suggest the need for future research about how and under what conditions these positive effects could be replicated. If adapted into a "push-in" or "pull-out" service during the school day, High 5s could serve as a model for ways of reconceptualizing kindergarten math instruction. More information is needed to determine whether similar effects from High 5s can be generated in other contexts — for example, in places such as Tulsa or Boston, where Building Blocks is already taught in pre-K, or in situations in which children did not receive a strong foundation of skills in pre-K via Making Pre-K Count.

This report lays out the above findings in greater detail. Chapter 2 describes the math programs and the context in which they were implemented, as well as the research design. Chapter 3 presents the end-of-kindergarten impacts of enhanced math experiences in pre-K and in kindergarten. Chapter 4 concludes with factors that may have contributed to the findings to date and a discussion of the findings' implications.

---

[16]Clements and Sarama (2008); Clements et al. (2011).

Chapter 2

Programs, Context, and Study Design

This chapter describes the Making Pre-K Count and High 5s programs that were tested and the conditions in which the programs were implemented. Making Pre-K Count comprised a prekindergarten math curriculum and supporting professional development, taking place at a time when math instruction in New York City pre-K programs was changing substantially. Pre-K teachers overall were delivering a large amount of math by the last year of the study. High 5s, for kindergartners, was a small-group math enrichment program that was provided in addition to the existing classroom instruction. Unlike the pre-K environment, the regular kindergarten math environment was more didactic and less likely to involve the small-group instructional format used in High 5s.

The two enhanced math programs were tested using a rigorous two-stage random assignment design. In the Making Pre-K Count study, pre-K programs in both public schools and community-based organizations were randomly assigned either to receive Making Pre-K Count or to continue with pre-K as usual. Comparing outcomes for children in the two kinds of programs constitutes a test of the effects of one year of math enrichment. In the High 5s study, students who had been in Making Pre-K Count program classrooms in pre-K in public school sites only were individually randomly assigned to participation in the High 5s program or to a business-as-usual kindergarten experience. As described in more detail below, this two-stage design allowed outcomes for children in the High 5s program to be compared with outcomes for (a) children who received Making Pre-K Count only and (b) children who received no math enrichment in pre-K or kindergarten.

The Making Pre-K Count Program

Making Pre-K Count provided pre-K classrooms with an evidence-based 30-week math curriculum (Building Blocks) developed by Douglas Clements and Julie Sarama, along with ongoing professional development. The Building Blocks curriculum includes a manualized, sequenced set of learning activities focused on numeracy and geometry. Curricular activities are organized based on the natural progressions by which children learn and develop math competencies over time, referred to as learning trajectories.[1] The program has an implicit focus on language skills and metacognition (encouraging children to reflect on and articulate their mathematical thinking) by asking questions like "How do you know?" It is hypothesized that

[1]Clements and Sarama (2004).

enhanced math instruction may also support working memory and inhibitory control, both components of children's executive function, by having children engage with complex math concepts such as problem solving.[2]

Building Blocks includes four main components: daily whole-group math lessons; hands-on math materials for children to play with in classroom play centers; small-group math activities for teachers to conduct with three to four children; and computer games that adapt to each child's skill level. Implementation of the math curriculum was supported with ongoing teacher training and coaching relating to the implementation of the program for lead and assistant teachers two to four times a month. To familiarize themselves with the program, pre-K teachers implemented Building Blocks for two years. Coaching and training were provided in both years of implementation. Program impacts were assessed in the second year of implementation, after teachers were comfortable with the curriculum.

Prior findings demonstrate that the Making Pre-K Count program was generally implemented well, with strong training, coaching, and use of the curriculum. As described in the pre-K report, teachers implemented three of the four Building Blocks components (whole-group activities, hands-on centers, and small-group activities) at levels prespecified by the research team; implementation of the computer math software lagged behind the other components but improved over the course of the academic year.[3] The levels for computer usage were prespecified by the research team for optimal implementation, which might not have been achievable in a real-world test of the program without extensive developer support. Further analysis of data collected by the online software shows levels of computer use in Making Pre-K Count comparable to that of another implementation of the Building Blocks program conducted under the direction of the developer of the program.[4]

Teachers in Making Pre-K Count classrooms conducted about 12 more minutes of math during a three-hour observation than teachers in pre-K-as-usual classrooms, despite the substantial amount — nearly 35 minutes per morning — already provided by pre-K-as-usual classrooms. This translates to an additional hour of math per week, on top of a base of about three hours of math per week. The additional math instruction focused on a range of math content areas: numbers, operations, and geometry. Making Pre-K Count teachers also demonstrated slightly higher-quality math instruction, although math instructional quality was generally low and there was no impact on global instructional quality when it was measured across all instructional contexts (not just math). Further analyses of the classroom quality data show that teachers in Making Pre-K Count classrooms were rated slightly higher on their use

[2]Blair and Razza (2007); Bull, Espy, and Wiebe (2008).
[3]Morris, Mattera, and Maier (2016).
[4]Anthony, Farran, and Hofer (2013).

of open-ended questions during math instruction, a particular focus of the Building Blocks curriculum as noted above.[5]

There were early, positive impacts of Making Pre-K Count on children's math and executive function scores in the fall of the pre-K year, but these impacts generally faded by the spring of that same year. Math in pre-K was measured using the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) assessment, a nationally normed test of math skills that was available in both Spanish and English, and the Woodcock-Johnson Applied Problems, a standardized measure of math used across the whole lifespan. At the end of pre-K, outcomes for children in Making Pre-K Count sites were generally similar to outcomes for children in pre-K-as-usual sites.[6] This may have been due in part to the large amount of math taking place in New York City pre-K programs during the course of the study. The ECLS-B and Woodcock-Johnson also may not have been sensitive enough to capture the impact of the program.

The High 5s Program

High 5s, a supplemental program for kindergarten children who received the Making Pre-K Count program in pre-K, used small-group math clubs held outside of regular instructional time to complement children's regular classroom instruction. Each club was typically led by one facilitator and included four children. Clubs were designed to meet three times a week for approximately 28 weeks starting in the fall of 2015. Facilitators were paid a paraprofessional salary and were hired and supervised by Bank Street College of Education. The High 5s program was developed by Robin Jacob and staff members at the University of Michigan with input from Doug Clements and Julie Sarama, the developers of Building Blocks.

The High 5s program was designed to pick up where Building Blocks left off. Many of the activities were adapted from Clements and Sarama's *Learning and Teaching Early Math*.[7] Activities were designed to move students along four mathematical learning trajectories: counting, composition of numbers, early addition and subtraction, and geometry. High 5s also included some measurement and patterning activities. Each activity contained recommendations about how to tailor or differentiate the activity for students at different points on the learning trajectory. The pacing of the program was designed to roughly align with the Go Math! curriculum used in most New York City kindergarten classrooms; so, for example, High 5s focused on

[5]Teachers in the program group were more likely to conduct a math activity at all, and therefore to be scored on this item. The difference in scores on this item includes only teachers who conducted a math activity and cannot be interpreted as a true impact.

[6]Making Pre-K Count had a small, positive, and statistically significant impact on children's inhibitory control — a facet of executive function — at the end of pre-K.

[7]Clements and Sarama (2014).

counting at roughly the same time children would have covered counting in their classrooms, and it focused on addition and subtraction at roughly the same time those topics were covered in their classrooms. However, the material in High 5s was designed to be somewhat more advanced than what students would have received in their kindergarten classrooms. As described in the companion report, High 5s clubs covered a wider range of content and included more hands-on learning than kindergarten classrooms.[8]

Clubs were scheduled either before school, during lunch, or after school in an effort not to disrupt classroom instructional time, and each club lasted 30 to 45 minutes and involved two start-up activities and one main activity. Activities were delivered in a game-like format and were intended to be fun, engaging, and developmentally appropriate. Every fourth club was a Game Day, which allowed children to choose among several different activities. Information was sent home to parents each Game Day, with an overview of what children were learning in clubs and suggestions for things to do at home. High 5s involved a substantial amount of training and supervision, led by Bank Street College and MDRC, designed to support facilitators over the course of the year, including 16 days of training before the start of the year, 8 additional days of training through the year, weekly group supervision, individual meetings with supervisors as needed (but intended to occur about once every two weeks), coaching in the field, and monitoring of data to help guide training and supervision.

High 5s was implemented with fidelity to the program model.[9] Attendance, student engagement, and adherence to the High 5s model were all high. Average student attendance was 87 percent for the year, and 92 percent of scheduled sessions were completed. This is noteworthy given that High 5s was a voluntary program that in some schools required children to arrive before school or stay after school to attend. Club observations indicated that instructional materials were set up correctly and facilitators were familiar with the activities they were implementing and understood the mathematical concepts they were teaching. Facilitators and observers rated student engagement as high.

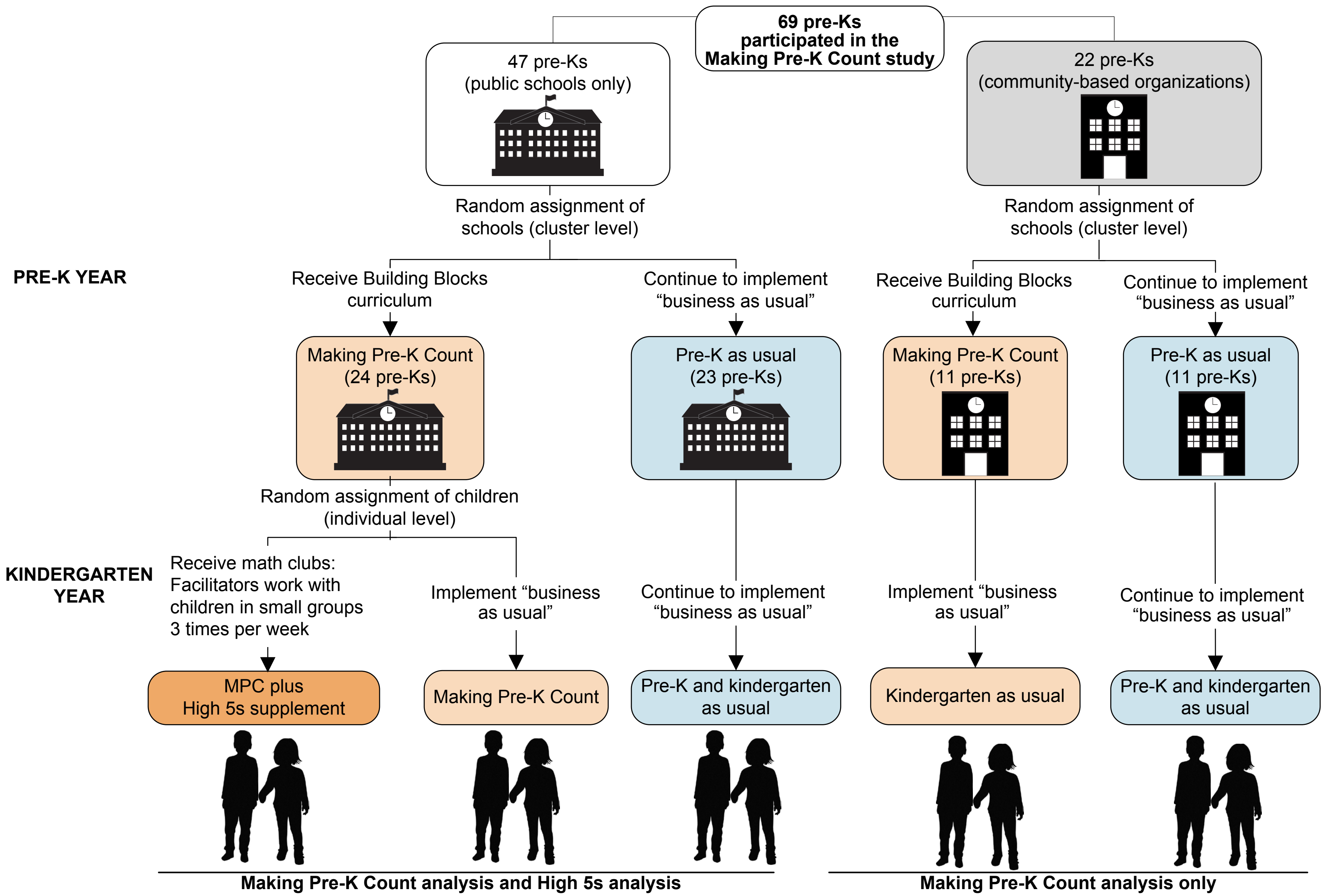
Study Design

The two enhanced math programs (Making Pre-K Count and High 5s) were tested using a rigorous random assignment design. The design, shown in Figure 2.1, allows for three key questions and comparisons that are addressed in this report.

[8]See the companion report for more details (Jacob, Erickson, and Mattera, 2018).
[9]See the companion report for more details (Jacob, Erickson, and Mattera, 2018).

Figure 2.1

Making Pre-K Count (MPC) and High 5s Design



First, in the pre-K year, entire pre-K program sites at public schools and community-based organizations were randomly assigned to receive the Making Pre-K Count program (MPC) or to continue business as usual. Comparing outcomes for children in pre-K sites that implemented Making Pre-K Count with outcomes for children in pre-K sites that continued business as usual allows for an investigation of the effects of *one year of math enrichment* (Making Pre-K Count) compared with *no math enrichment* (the pre-K-as-usual control condition).

Second, in the year after pre-K, when children went to kindergarten, children within a subset of sites — the public school sites only — participated in a rigorous test of the High 5s program in kindergarten. Children in community-based preschool settings were not included in the High 5s design for logistical reasons: They were less likely to move together into kindergarten, making it harder to group them together into a sufficient number of clubs for the research design. As shown in Figure 2.1, the High 5s random assignment sample was drawn from the public schools that were part of the original Making Pre-K Count program group. In these sites, children were randomly assigned to either the High 5s program group in kindergarten (MPC plus High 5s supplement group) or a kindergarten-as-usual group (Making Pre-K Count only group). Children in the remaining public school sites, which did not implement any math enrichment, constituted the pre-K-and-kindergarten-as-usual control group. This two-stage sequential random assignment design thus created three experimental groups that can be used to investigate two additional comparisons: (a) the effects of *two years of math enrichment* (MPC plus High 5s Supplement) compared with *one year of math enrichment* (Making Pre-K Count) and (b) the effects of *two years of math enrichment* (MPC plus High 5s supplement) compared with *no math enrichment* (pre-K-and-kindergarten-as-usual control condition).

Sample

Making Pre-K Count took place in 69 pre-K sites, including 173 classrooms, across New York City. Thirty-five sites were randomly assigned to receive the Building Blocks curriculum, coaching, and training. Thirty-four sites continued with their usual pre-K practices. Groups of four to five sites were randomly assigned within random assignment blocks. Sites were blocked by borough, venue (public school or community-based organization), and racial/ethnic composition (serving over 60 percent Hispanic children or serving children from other racial/ethnic backgrounds). This study follows the children who participated in Making Pre-K Count in pre-K as they make the transition into kindergarten the following year. Over 95 percent of eligible children who were in Making Pre-K Count sites in pre-K were tracked successfully into kindergarten, with similar child attrition across groups.

To examine the effect of *one year of math enrichment*, kindergarten outcomes for children who were in pre-K sites that implemented Building Blocks (Making Pre-K Count, n = 641) were compared with outcomes for children who were in pre-K sites that continued with pre-K as usual (pre-K-as-usual control group, n = 684).[10] The Making Pre-K Count program group in this analysis excludes the subset of children in public schools who received High 5s (MPC plus High 5s supplement group in Figure 2.1), so that the effect of one year of pre-K math enrichment *only* can be estimated.[11] Children in the program group and the control group in the resulting sample were similar on measured baseline demographic characteristics (see Appendix Table B.2 in Appendix B for baseline equivalence). A majority of the sample was Hispanic (57 percent) or non-Hispanic black (35 percent).

As described above, the High 5s study was embedded in the larger Making Pre-K Count design — taking place only within *public school sites*. Within the 24 Making Pre-K Count program public schools, children who had received Making Pre-K Count and remained in the same public school between pre-K and kindergarten were randomly assigned early in the fall of the kindergarten year to either the High 5s program or a control condition in which they would receive business-as-usual kindergarten instruction. As a result, the High 5s analytic sample includes public school children who stayed in the same school, some of whom were randomly assigned to two years of math enrichment (MPC plus High 5s supplement, n = 303) and some of whom received only one year of math enrichment (Making Pre-K Count, n = 310).[12] This analysis estimates the effect of math enrichment in kindergarten. Children in the MPC plus High 5s supplement group and the Making Pre-K Count group were similar on measured baseline demographic characteristics (see Appendix Table B.3).

Because the High 5s study involved two stages of randomization (random assignment of sites in pre-K and random assignment of individuals in kindergarten), the MPC plus High 5s supplement group can also be compared with a third group of students: those in public schools who received no math enhancements in either pre-K or kindergarten (the pre-K-and-kindergarten-as-usual control group, n = 345). The comparison builds on the cluster-

---

[10]A random subset of children in each center was selected to participate in data collection.

[11]Of the children selected in sites that implemented Building Blocks, approximately one-quarter had also been randomly assigned to participate in High 5s. Those children were removed from the Making Pre-K Count only analysis sample and replaced with other children from the same schools who did not receive High 5s in kindergarten. See Appendix B for further details on the sample and Appendix C for more information about the analysis.

[12]Parental consent was another condition for participation; 97 percent of parents gave positive consent in the pre-K year for their children to participate in High 5s. Children who stayed in the same school but did not consent to participate in the High 5s program (n = 18) were still part of the Making Pre-K Count study and were selected for kindergarten data collection. For analysis purposes, these children were randomly assigned either to the MPC plus High 5s supplement group or to the Making Pre-K Count only group to maintain the internal validity of the random assignment design when making comparisons with the no-math-enrichment control group. See Appendix B for further details about the sample and Appendix C for more information about the analysis.

randomized design and compares outcomes for children who stayed in the 24 Making Pre-K Count *program group* public schools (and were randomly assigned to the High 5s program) with outcomes for children who stayed in the 23 Making Pre-K Count *control group* public schools.[13] This analysis estimates the effect of two years of math enrichment in pre-K and kindergarten. The MPC plus High 5s supplement group and the pre-K-and-kindergarten-as-usual control group were similar at baseline on measured demographic characteristics (see Appendix Table B.4 for baseline equivalence).

As with the analytic sample for the unique effect of Making Pre-K Count, children in these analyses were predominantly Hispanic (56 percent) or non-Hispanic Black (36 percent).

The Pre-K and Kindergarten Environments

To understand the impact of these programs it is important to understand the counterfactual condition, or what children would have experienced in the absence of the program. Children in Making Pre-K Count experienced different math instructional contexts as they moved from pre-K to kindergarten. As described in a previous report, the New York City pre-K landscape was undergoing substantial change during the study, leading to a rapid shift in the amount of math being delivered in New York City pre-K classrooms.[14] Pre-K-as-usual classrooms engaged in 35 minutes of math instruction per morning (or approximately three hours per week) — substantially more than the average of 19 minutes per day (or an hour and a half a week) documented across earlier studies of Building Blocks.[15] In these prior studies of Building Blocks, where teachers conducted as little as 12 minutes of math each day, instruction probably focused on basic concepts like counting. In New York City pre-K classrooms, where teachers had increased their math teaching since 2013 to 35 minutes of math per morning in 2015, the additional math instruction was generally guided by a curriculum or framework like the Common Core; thus, it is likely to have included more extensive content. Forty-two percent of pre-K-as-usual sites were using structured math curricula, and over half were using some form of math computer games. While pre-K-as-usual classrooms in New York City spent more time in whole-group instruction (nearly one hour per morning) than small-group instruction (about 15 minutes per morning), the majority (over 60 percent) of the

[13]Mobility in these schools was similar between program and control groups, with 69 percent of children staying in the same school in the Making Pre-K Count public schools and 67.5 percent of children staying in the same school in the control group public schools. See Appendix C for more information about the analysis.

[14]Morris, Mattera, and Maier (2016).

[15]Clements and Sarama (2008) documented 16.4 minutes of math per morning; Sarama et al. (2008) documented 12.2 minutes of math per morning; Clements et al. (2011) documented 27.2 minutes of math per morning.

time in small-group activities — time that is often used as an opportunity to provide more individualized instruction — was spent on math.[16]

Although previous studies of the Building Blocks curriculum noted impacts of just 2 to 5 additional minutes of math in a day, those added minutes may have been particularly salient in a context where children typically experienced only an average of 19 minutes of math per morning. In contrast, while Making Pre-K Count teachers taught an additional 12 minutes of math per morning (or approximately 60 minutes more per week), those 12 minutes may have been less salient in an environment where students were already receiving 35 minutes of math instruction each morning. In other words, the large amount of math in the pre-K-as-usual condition may have contributed to the lack of pre-K impacts, despite the observed impact on the amount of instructional time in the Making Pre-K Count classrooms.

To test this hypothesis, a subgroup analysis was conducted to compare impacts at the end of pre-K according to the amount of math conducted in the pre-K-as-usual classrooms. As expected, impacts on pre-K math skills were larger in study blocks where there was less time spent on math instruction in pre-K-as-usual classrooms. In those blocks, there were positive, statistically significant impacts on children's math skills at the end of pre-K (as measured by the Woodcock-Johnson Applied Problems). (See Appendix A for more about the analysis.)[17]

The following year, children entered a kindergarten context that stood in contrast to the math instruction children were exposed to in pre-K and to what they would receive in the High 5s clubs.[18] The contrast was not in the amount of instruction; in fact, kindergarten teachers spent more time on math than in pre-K, as one might expect, averaging around 56 minutes per day. Instead, the contrast was in the way that the math instruction was delivered: Most classrooms had only one teacher, and over 80 percent of the math instruction involved whole-group

[16]Observations lasted three hours and included the main instructional time of the day. Observations were conducted using the Narrative Record (developed by Dale C. Farran and Carol Bilbrey at the Peabody Research Institute, Vanderbilt University, Nashville), an open-ended instrument for describing the types of activities and content of instruction in episodes of time.

[17]The analysis compared random assignment blocks in which control group classrooms spent fewer minutes of math each morning with blocks in which the control group classrooms spent more minutes on math. Blocks with fewer minutes of control group math had impacts on the Woodcock-Johnson in pre-K that were statistically significant and substantially larger (effect size = 0.18 standard deviations) than those with more control group math per morning (effect size = -0.02 standard deviations), as shown in Appendix Table A.1. The difference between the subgroups was almost statistically significant at the 10 percent level ($p = 0.11$). This exploratory analysis provides some support to the hypothesis that the large amount of math being conducted in New York City may have played a role in the lack of pre-K impacts, although there may be other differences between low-math and high-math sites that may also explain the different effects.

[18]Observations of the kindergarten math environment were conducted in 39 kindergarten classrooms in Making Pre-K Count public schools.

instruction or seat work. About two-thirds of the activities in which students engaged involved either no materials or only a workbook. The math clubs, in which one adult worked with a handful of children on focused math content, provided a unique opportunity for children to experience individualized math instruction.

Chapter 3

Impacts of Early Enhanced Math Instruction

This chapter presents the effects of Making Pre-K Count (MPC) and High 5s on children's math, language, and executive function skills at the end of kindergarten. *One year of math enrichment* in pre-K (Making Pre-K Count) had a small, positive, but not statistically significant impact on one of two measures of children's kindergarten math skills. Making Pre-K Count also led to positive effects on children's attitudes toward math and working memory in kindergarten, a year after they received the program. *Two years of math enrichment* in pre-K and kindergarten (MPC plus High 5s supplement) had a positive impact on one of two measures of children's assessed math skills in kindergarten compared with Making Pre-K Count alone in pre-K. However, the additional year of the program did not have a statistically significant impact on attitudes toward math or executive function beyond what students gained from participation in Making Pre-K Count alone. Together, *two years of math enrichment* across pre-K and kindergarten (MPC plus High 5s supplement) led to the largest impacts on children's kindergarten math skills when compared with *no math enrichment* in pre-K or kindergarten. Neither program had a statistically significant effect on children's assessed language or inhibitory control skills.

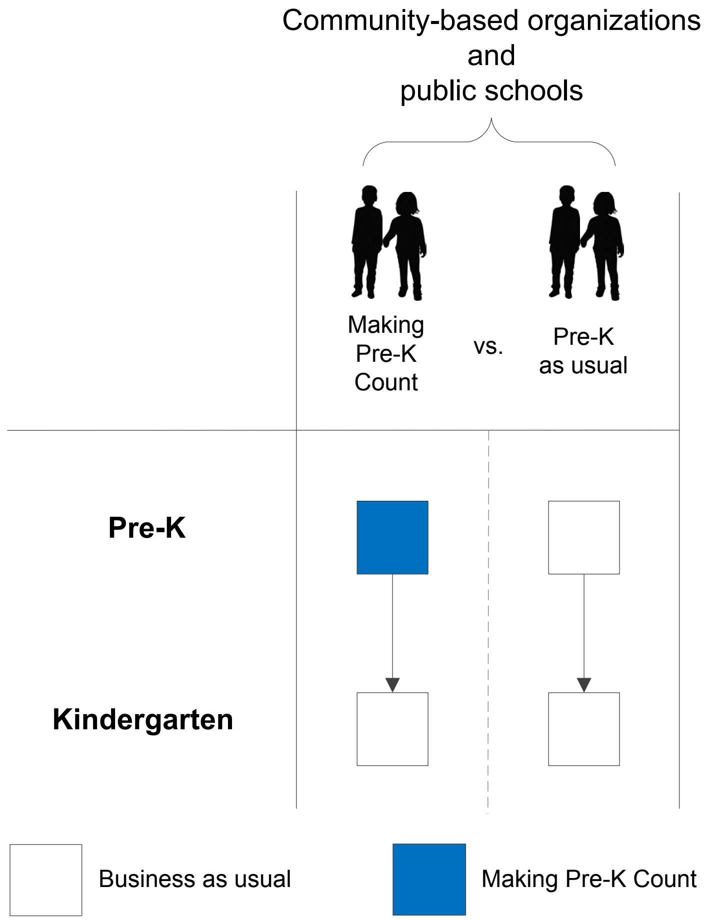
Impacts of Making Pre-K Count

The first set of findings compares kindergarten outcomes for children who received *one year of math enrichment* in pre-K via Making Pre-K Count to kindergarten outcomes for children who received *no math enrichment*. All the children in this analysis received kindergarten as usual, as shown in Figure 3.1. The findings for this analysis, which examine the unique effects of Making Pre-K Count on children's kindergarten outcomes, are presented in Table 3.1. (See Box 3.1 and Box 3.2 for descriptions of the measures used to assess children's kindergarten outcomes.)

- **A year after participation in the program, there was a small, positive, but not consistently statistically significant impact of Making Pre-K Count in pre-K on the more sensitive of two measures of children's kindergarten math skills, the Research-Based Early Math Assessment–Kindergarten (REMA-K). Scores on the other measure, the Woodcock-Johnson Applied Problems, did not differ among children in Making Pre-K Count and pre-K-as-usual sites. Impacts on the REMA-K were sensitive to how both the measure and the sample were specified, and some specifications yielded statistically significant impacts.**

Figure 3.1

Analytic Group for the Impact of Making Pre-K Count



At the end of kindergarten, impacts on REMA-K scores were positive but not statistically significant (effect size = 0.08, p = 0.11).[1] There were positive and statistically significant impacts on the REMA-K numeracy, geometry, and patterning subscales; see Appendix F. To test the robustness of the findings, a variety of sensitivity analyses were conducted using different scoring and sample specifications. The estimated effect of Making Pre-K Count on REMA-K scores was small but positive across all five measurement and model specifications and was statistically significant in three out of five of the models (see Appendix E), suggesting

[1]The effect size is calculated by dividing the estimated effect of the program (the difference between means for the program group and the control group) by the standard deviation for the control group. An effect size of 0.08 here represents an improvement in REMA-K scores equal to 8 percent of the standard deviation.

Table 3.1

Impacts of Making Pre-K Count in the Spring of the Kindergarten Year

Outcome Score	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Effect Size[a]
Math						
REMA-K[b]	38.35	37.61	0.73	0.11		0.08
Woodcock-Johnson Applied Problems[c]	104.28	104.24	0.04	0.96		0.00
Math attitudes						
Children's attitudes toward math (1-5)	3.60	3.43	0.17	0.04	**	0.11
Language						
Receptive vocabulary[c]	98.01	97.43	0.57	0.60		0.04
Executive function						
Inhibitory control (0-1)	0.69	0.68	0.01	0.43		0.05
Working memory[d]	2.38	2.22	0.16	0.06	*	0.11
Sample size						
Blocks	16	16				
Sites	35	34				
Children	641	684				

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K with corresponding outcomes for the pre-K-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

that there may have been a positive (but small) impact on children's REMA-K scores. Effect sizes in these analyses ranged from 0.08 to 0.12 standard deviations, which translates to between 1.1 and 1.2 months of growth.[2] These effects are smaller than those seen on the REMA-K in a previous study of Building Blocks that took place in a similar context, which found an effect size of 0.19 at the end of kindergarten.[3]

There was no difference in outcomes between the groups on the Woodcock-Johnson Applied Problems subtest.

- **Making Pre-K Count led to improvements in children's math attitudes and about two months' worth of growth in children's working memory skills at the end of kindergarten.**

Children were asked to indicate how happy or unhappy math made them feel on a scale from 1 to 5, with 5 indicating the most happy. Children in the program group rated their attitudes toward math as more positive (a score of 3.60 on a 5-point scale) than children in the pre-K-as-usual group (a score of 3.43). This translates to an effect size of 0.11 standard deviations.

Math has been closely connected with children's executive function — and specifically their ability to start and stop (inhibition) and to manipulate small amounts of information mentally (working memory). Children in the Making Pre-K Count group had statistically significantly higher working memory scores (2.38) than children in the pre-K-as-usual group (2.22). This difference translates into an effect size of 0.11 standard deviations, or two months of growth.[4] There was no statistically significant difference in children's inhibition skills as measured by the Hearts and Flowers computer task (effect size = 0.05).

---

[2]Growth was determined using the change in the score for the control sample of a prior study of Building Blocks, the TRIAD study (Hofer, Lipsey, Dong, and Farran, 2013), in which a version of the REMA adapted for kindergarten was used. The T score (a standardized score) at the end of pre-K was subtracted from the T score at the end of kindergarten; the result can be interpreted as the expected trajectory of growth over the course of the kindergarten year with no intervention (7.6 REMA T score points per year or 0.63 REMA T score points per month). To calculate growth per month, the REMA-K impact in this study was divided by the monthly growth in the TRIAD control sample. This calculation was used for all REMA-K growth estimates in this report.

[3]Clements et al., "Effects on Mathematics" (2016).

[4]Growth was determined using the Corsi Blocks backward span score (see Box 3.2) for the children in this study's control group. The change in the backward span score from the end of pre-K to the end of kindergarten in the control group can be interpreted as the expected trajectory of growth over the course of a year with no intervention (0.94 additional blocks per year or 0.08 additional blocks per month). In order to explain growth for the program group in this report beyond what would be expected for the control group, the impact on the Corsi Blocks backward span for the program group was divided by the control group difference in score between pre-K and kindergarten.

Box 3.1

Child Math Measures Used in Making Pre-K Count and High 5s Studies

The **Research-Based Early Math Assessment–Kindergarten** (REMA-K), an adaptation of the Research-Based Early Math Assessment,* measures thinking and learning along research-based developmental progressions for math topics. The REMA-K uses a subset of REMA items that were selected to represent a full range of early mathematics competencies related to numbers and operations, measurement, patterning, and shapes that are applicable within the prekindergarten, kindergarten, and first-grade years.

Woodcock-Johnson Applied Problems is a subscale of the Woodcock-Johnson III Tests of Achievement.† It is a valid standardized assessment of mathematical thinking for ages 2 through 90; early items are suitable for assessing simple math functions relevant at young ages. The Woodcock-Johnson Applied Problems test is a less detailed, more global measure of children's math ability than the REMA-K. The assessment is widely used in studies of early childhood curricula and has been nationally normed for ease of interpretability.

The REMA-K may be a more sensitive measure than the Woodcock-Johnson Applied Problems. A number of factors may play a role in the measure's sensitivity (described in more detail in Appendix D).

While the REMA-K was created by the developers of Building Blocks, analyses suggest that the REMA-K is not closely aligned with the Building Blocks and High 5s programs. Ten percent of the items on the REMA-K were similar in approach to the types of activities to which children were exposed in Building Blocks. (High 5s activities were developed separately from Building Blocks and were not directly based on Building Blocks activities, although they were designed to build on their content.) The magnitude and statistical significance of the estimated effects described in this report did not change when the more aligned items were removed from the score.‡ The results of this content analysis are consistent with a prior psychometric examination of the REMA.§ That analysis concluded that the REMA was a valid measure of children's math skills and that specific items were not more sensitive to the experience of being in the Building Blocks program. Instead, the following key differences between the two tests may have contributed to the sensitivity of the REMA assessment, relative to the Woodcock-Johnson.

1. The REMA-K is a lengthier test that includes more items with the intention of obtaining granular information about children's abilities in each mathematical skill at each level (Appendix Table D.1). For example, the REMA-K includes nine questions about geometry or geometric measurement, while the Woodcock-Johnson includes only three items in this domain. This is true of the other subdomains of math assessed, as well. Statistically significant impacts on the REMA-K were found across numeracy, patterning, and geometry subscales, suggesting more sensitivity in all three sub-content areas.

(continued)

Box 3.1 (continued)

2. Each of the REMA-K items assesses very precise learning about a specific skill in math. Most items tap only one specific skill at a time and assess different levels of that skill. For example, the REMA-K items assessing children's numeracy ability ask a child to solve "four plus three" and then "seven plus eight," identifying whether the child can add numbers below five and numbers above five, two separate levels of children's learning. By contrast, many of the Woodcock-Johnson items assess more than one math skill within a single item. For example, one item on the Woodcock-Johnson may ask children both to add and to understand coin value.

3. The REMA-K items vary in how they assess children's skills, assessing children's multiple approaches to solving math. Some of the REMA-K items use manipulatives and others use mental math (asking them to provide answers without objects). All of the Woodcock-Johnson items use picture visuals to support the questions; none rely on manipulatives and none rely on purely "mental math."

4. The Woodcock-Johnson items test math in the context of language, using word problems, for example, that require vocabulary knowledge to correctly solve the math problem.

See Appendix Box D.1 for an illustrative example of REMA-K and Woodcock-Johnson items.

---

*Clements, Sarama, and Liu (2008).

†Woodcock, McGrew, and Mather (2001).

‡For each sample, a sensitivity analysis was conducted in which impacts on the REMA-K raw score (the total number of items on the assessment that a child answered correctly) were compared with impacts on a version of the REMA-K raw score excluding the five items determined to be aligned with Building Blocks. (For more information about alignment, see Appendix D.) Results show that effect sizes in this sensitivity analysis were similar to those with the original REMA-K raw score. One year of math enrichment (Making Pre-K Count) compared with no math enrichment (pre-K as usual) had statistically significant impacts on both the REMA-K raw score (effect size = 0.12; $p = 0.04$) and the sensitivity measure (effect size = 0.12; $p = 0.04$). Two years of math enrichment (Making Pre-K Count and High 5s) compared with one year of enhanced math (Making Pre-K Count only) had positive impacts on children's assessed math skills on the REMA-K raw score (effect size = 0.22; $p < 0.01$) and the sensitivity measure (effect size = 0.23; $p < 0.01$). Two years of math enrichment (Making Pre-K Count and High 5s) compared with no math enrichment (pre-K and kindergarten as usual) had positive impacts on both the REMA-K raw score (effect size = 0.43; $p < 0.01$) and the sensitivity measure (effect size = 0.47; $p < 0.01$).

§Weiland and Yoshikawa (2013).

Box 3.2

Other Child Outcome Measures Used in Making Pre-K Count and High 5s Studies

Language

The **Receptive One-Word Picture Vocabulary Test** (ROWPVT-4) assesses children's receptive vocabulary, or their ability to understand spoken language, by asking them to match a word the assessor says out loud to a picture of an object, an action, or a concept.*

Math attitudes

The research team at MDRC created an assessment to measure children's attitudes toward math. Children were shown a card that displayed a range of five faces (sad to smiling) to describe how happy or unhappy school and math made them feel. A rating of 1 indicates that they felt very unhappy and a rating of 5 indicates that they felt very happy.

Executive function

Hearts and Flowers is a computerized task that measures inhibitory control. During this task, a child is asked to select the button on the same side of the screen if a heart appears on the screen and on the opposite side of the screen if a flower appears on the screen.†

The **Corsi Blocks** task assesses short-term working memory. During this task, a child is asked to repeat a sequence of blocks tapped by an assessor, tapping the blocks in reverse order. The child begins with a sequence of two blocks and more blocks are added to the sequence.‡

---

*Martin and Brownell (2011).
†Wright and Diamond (2014).
‡Corsi (1972); Lezak (1983).

- **There was no effect of Making Pre-K Count, positive or negative, on children's assessed language in the spring of kindergarten.**

The Building Blocks curriculum focuses on language development and helping children learn to talk about math. Therefore, it was hypothesized that the program might have a positive effect on children's language skills. Conversely, there was also a concern that a focus on math would reduce teachers' time supporting children's early literacy. There was no difference between children in Making Pre-K Count classrooms and those in control group classrooms in children's language skills as measured by the Receptive One-Word Picture Vocabulary Test (ROWPVT) at the end of kindergarten.

- **Making Pre-K Count had positive, statistically significant impacts on math and math attitudes for children who attended pre-K programs based in *public schools*. Impacts on working memory were positive and statistically significant for children in *community-based* pre-K programs (see Table G.1).**

Because High 5s was implemented only in public school sites, impacts were examined separately for pre-K programs based in public schools and community-based pre-K programs. Impacts on both the REMA-K and math attitudes assessments were positive and statistically significant for the public school sample. REMA-K impacts translate to nearly two additional months of growth (effect size = 0.13). However, these impacts were not statistically significantly larger than the impacts for children in community-based pre-K programs. By contrast, impacts on children's working memory skills as assessed by the Corsi Blocks task were statistically significant in community-based sites, where they translate to over four additional months of growth, but not in public school sites when examined separately. Again, there was no statistically significant difference between the impacts for children in the two kinds of venues.

The next set of analyses examine the effects of the High 5s math clubs. Again, these clubs took place in *public school program group sites only* and therefore on top of a base of positive impacts on children's math skills and math attitudes from the pre-K math program.

Impacts of the High 5s Kindergarten Math Supplement

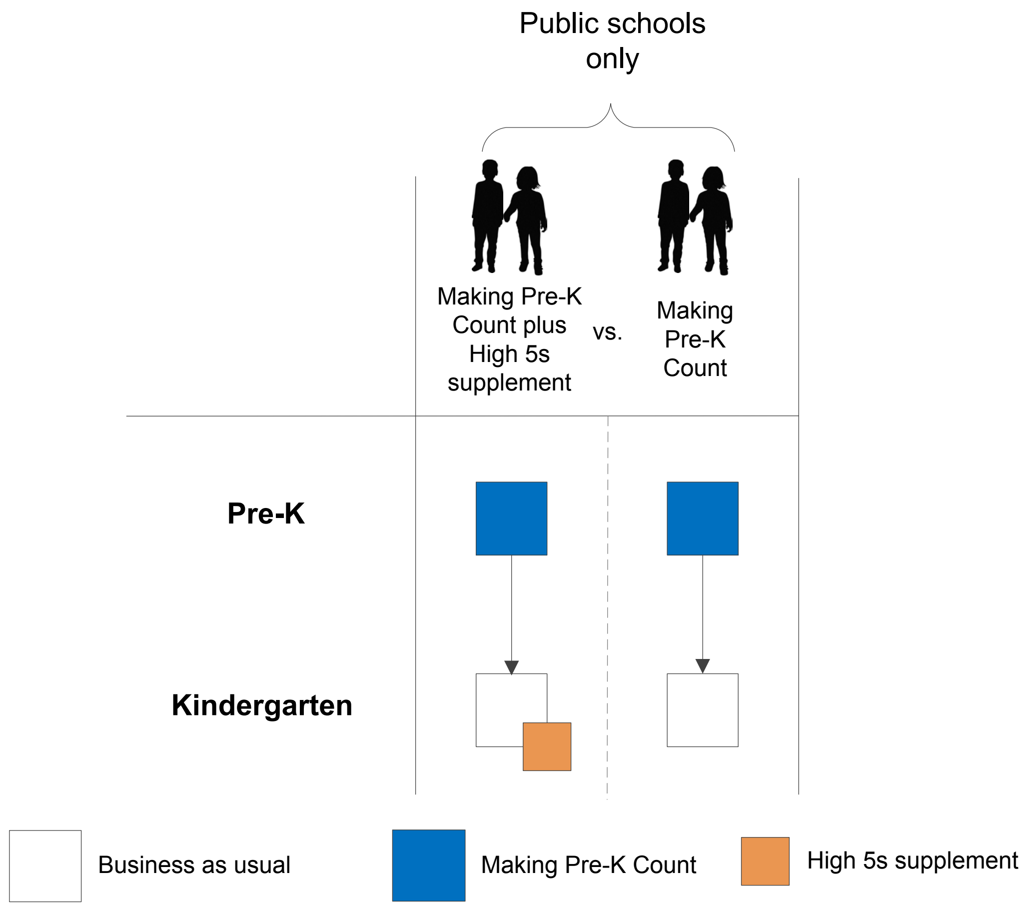
The next set of findings compares kindergarten outcomes for children assigned to *two years of math enrichment* (MPC plus High 5s supplement) with outcomes for children assigned to *one year of math enrichment* (Making Pre-K Count), as shown in Figure 3.2. Table 3.2 presents the impacts of MPC plus the High 5s supplement compared with only Making Pre-K Count on children's kindergarten outcomes.[5]

---

[5]Impacts of the High 5s kindergarten supplement were also estimated for the original 655 consenters who were randomly assigned (not taking into account those who did not consent). Impacts were similar regardless of how the groups were specified; see Appendix H.

Figure 3.2

Analytic Group for the Impact of the High 5s Supplement



- **The addition of the High 5s kindergarten supplement had a positive effect equivalent to two and a half months of growth on children's kindergarten math skills as measured by the REMA-K. Differences in children's scores on the Woodcock-Johnson assessment were small and positive, but not statistically significant.**

There was a positive impact on children's REMA-K scores (effect size = 0.19). That is, children who received both Making Pre-K Count plus the High 5s kindergarten supplement had a score of 39.47, and children who received only Making Pre-K Count had a score of 37.90. This difference is equivalent to approximately two and a half additional months of growth on the REMA-K. Said differently, this effect is equivalent to closing almost one-fifth of the

Table 3.2

Impacts of the High 5s Supplement in the Spring of the Kindergarten Year

Outcome Score	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Effect Size[a]
Math						
REMA-K[b]	39.47	37.90	1.57	0.01	**	0.19
Woodcock-Johnson Applied Problems[c]	104.17	102.98	1.19	0.18		0.09
Math attitudes						
Children's attitudes toward math (1-5)	3.51	3.45	0.07	0.58		0.04
Language						
Receptive vocabulary[c]	97.89	97.03	0.85	0.46		0.06
Executive function						
Inhibitory control (0-1)	0.68	0.68	-0.01	0.69		-0.03
Working memory[d]	2.29	2.22	0.07	0.51		0.05
Sample size						
Sites	24	24				
Children	303	310				

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group received only MPC in pre-K and participated in kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the group assigned to MPC in pre-K and kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for pre-K sites.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

achievement gap between low-income children and their higher-income peers.[6] These findings were robust to differences in scoring. There was a positive but not statistically significant impact on the Woodcock-Johnson Applied Problems (effect size = 0.09).

- **The addition of the High 5s kindergarten supplement did not lead to statistically significant differences in children's outcomes on math attitudes, language, or executive function.**

The High 5s program did not have an effect on children's math attitudes, language, and executive function outcomes above and beyond the effect of Making Pre-K Count.

Impacts of Making Pre-K Count Plus the High 5s Kindergarten Supplement

Kindergarten outcomes for children assigned to *two years of math enrichment* (MPC plus High 5s supplement) were also compared with outcomes for children with *no math enrichment* (pre-K and kindergarten as usual) in public school sites only, as shown in Figure 3.3. Table 3.3 presents the results.

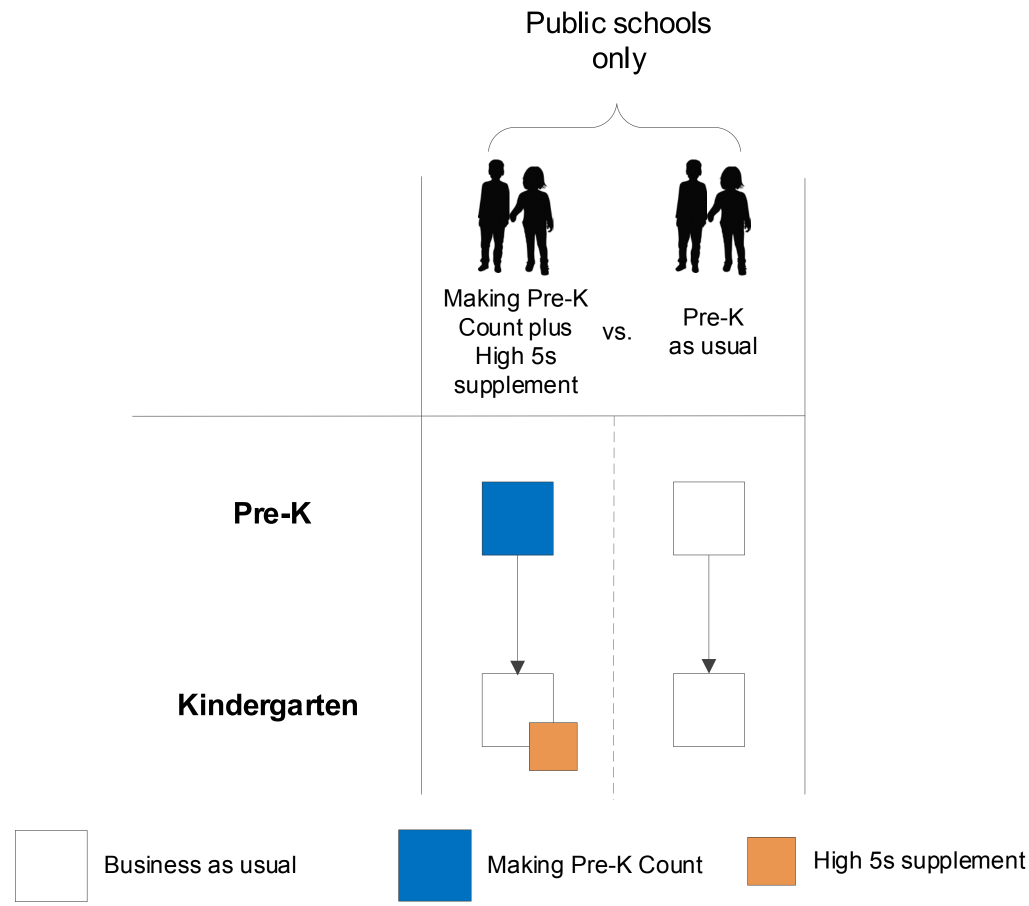
- **Two years of enhanced math instruction (MPC plus High 5s kindergarten supplement) had a positive and statistically significant impact equivalent to 4.2 months of additional growth on children's math skills as measured by the REMA-K, relative to no math enhancements. The difference in children's Woodcock-Johnson scores was small and positive but not statistically significant.**

The impact of two years of math enrichment on children's REMA-K scores was positive and statistically significant (effect size = 0.30). This effect translates to 4.2 months of additional growth in math skills as measured by the REMA-K, or approximately 29 percent of

[6]The effect sizes in this study are standardized measures of the difference in outcomes at the end of the kindergarten year for the control and program groups. Here, they are compared with standardized measures of the difference in outcomes at the end of kindergarten for children in the 90th income percentile and children in the 10th income percentile in the Early Childhood Longitudinal Study, Kindergarten Class (ECLS-K) conducted in the 2010-2011 year (as described in Reardon and Portilla, 2016). In spring 2011, the income achievement gap was equivalent to 1.046 standardized units. The effect size of the High 5s supplement in this study (0.19) is equivalent to 18 percent of that gap. The income achievement gap can also be calculated by comparing low-income children with their middle-income peers (equivalent to 0.556 at kindergarten entry in 2010). The effect found here is equivalent to closing one-third of the low-income–middle-income achievement gap.

Figure 3.3

Analytic Group for the
Impact of Making Pre-K Count Plus High 5s Supplement



the income achievement gap.[7] The effect of two years of enhanced math experiences on the REMA-K was larger when compared with the scores of students who received pre-K and kindergarten as usual in New York City than when compared with the scores of children who received Making Pre-K Count in pre-K but kindergarten as usual. These impacts are similar to those seen in a previous study of Building Blocks that included a kindergarten follow-through component (0.34 standard deviations).[8] There was a positive but not statistically significant impact on the Woodcock-Johnson Applied Problems subtest (effect size = 0.09).

---

[7]Or, using the alternative method described in note 6, this effect is equivalent to closing over half of the low-income–middle-income achievement gap.

[8]Sarama, Clements, Wolfe, and Spitler (2012).

Table 3.3

Impacts of Making Pre-K Count Plus High 5s Supplement in the Spring of the Kindergarten Year

Outcome Score	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Effect Size[a]
Math						
REMA-K[b]	39.12	36.50	2.63	0.00	***	0.30
Woodcock-Johnson Applied Problems[c]	103.89	102.83	1.06	0.40		0.09
Math attitudes						
Children's attitudes toward math (1-5)	3.51	3.29	0.22	0.10	*	0.14
Language						
Receptive vocabulary[c]	97.51	95.79	1.72	0.28		0.11
Executive function						
Inhibitory control (0-1)	0.68	0.66	0.02	0.40		0.08
Working memory[d]	2.26	2.14	0.12	0.41		0.08
Sample size						
Blocks	11	10				
Sites	24	22				
Children	303	345				

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group did not receive math enrichment and participated in pre-K and kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the pre-K-and-kindergarten-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

- **Making Pre-K Count plus High 5s had a positive effect on children's math attitudes in kindergarten.**

Children who received two years of enhanced math opportunities had a more positive attitude toward math than children who received typical pre-K and kindergarten math instruction (effect size = 0.14).

- **There were no statistically significant impacts on children's assessed language or executive function skills in kindergarten.**

Impacts on these outcomes were positive but not statistically significant compared with the outcomes for students who experienced business as usual in both pre-K and kindergarten. Effect sizes ranged from 0.08 to 0.11 standard deviations for these outcome measures.

Effects for Subgroups

Subgroups of Children

Subgroup analyses were conducted to understand whether the Making Pre-K Count and High 5s programs may have had differential impacts on child outcomes based on a predetermined set of school and child characteristics. Impacts on children's kindergarten outcomes were examined, first, for two comparisons: (a) children who received Making Pre-K Count only compared with children who received pre-K as usual and (b) children with two years of math enrichment (Making Pre-K Count plus High 5s) compared with children with one year of math enrichment (Making Pre-K Count only). For each of the two comparisons, four subgroup analyses were conducted, according to children's baseline (a) self-regulation skills, (b) receptive language ability, (c) gender, and (d) Spanish-speaking ability.[9]

Two subgroup analyses were also conducted for a third comparison: children who received two years of math enrichment compared with children who received no math enrichment. These analyses were based on (a) gender and (b) Spanish-speaking ability.

Appendix G presents the subgroup findings. As discussed in that appendix, there were statistically significant impacts on some outcomes for some subgroups of children. However, in

[9]For the first comparison pair (children who received Making Pre-K Count only and children who received pre-K as usual), baseline data were collected in fall of the pre-K year. For the second comparison pair (children with two years of math enrichment and children with one year of math enrichment), baseline data were collected in the spring of the pre-K year. Spanish-speaking ability was reported by the parent on the informed consent form at study entry via the following question: "Does your child speak and understand Spanish?"

most instances, impacts for these subgroups of children were not statistically larger than impacts for other children, and there is no consistent group of children who benefited across all outcomes. Therefore, it is not possible to make any definitive conclusions about which children are more likely to benefit from enhanced pre-K or kindergarten math experiences.

Subgroups by Classroom Time Spent on Math

A subgroup analysis was also conducted based on the amount of time spent on mathematics in kindergarten classrooms. This analysis only compared children assigned to two years of math enrichment with children with one year of math enrichment.

- **Impacts of High 5s on children's math attitudes varied depending on the amount of math instruction they were receiving in their kindergarten classrooms. Students who spent less time on math in their classrooms had larger positive impacts on attitudes toward math.**

In comparing pre-K outcomes, it was hypothesized that the amount of math instruction in the business-as-usual environment might have mattered for program impacts, given the large amount of math children were already receiving in absence of Making Pre-K Count. Similarly, it was hypothesized that the amount of math children received in their kindergarten classrooms might also matter for the degree to which High 5s affected their outcomes: The lower the amount of math provided in kindergarten classrooms, the greater the differential created by High 5s. On average, kindergarten classrooms in the 24 public schools in which High 5s was implemented conducted approximately 280 minutes of math a week. There were no statistically significant differences in children's math scores across schools with more or less math instruction. Interestingly, High 5s had a larger impact on children's math attitudes in schools with lower than average time spent on math (effect size = 0.29) than in schools with more time spent on math (effect size = -0.18), suggesting that the more time children spent in High 5s relative to classroom math instruction, the more likely they were to report positive attitudes toward math (Appendix Table G.8). This implies that while participation in High 5s had a positive impact on students' attitudes toward math, the impact may be especially positive when students were not spending a lot of time on mathematics in their kindergarten classrooms.

Conclusion

At the end of kindergarten, one year of enriched pre-K math led to a positive, small, but not consistently statistically significant impact on one of two measures of children's assessed math skills and had positive effects on children's math attitudes and working memory. An additional year of enriched math in kindergarten had a positive impact on children's math skills as measured on the REMA-K above and beyond any pre-K effects. Compared with no enrichment, two years of aligned, enriched math experiences had an even larger positive effect on children's math skills as measured by the REMA-K and a positive effect on children's math attitudes. Neither program had statistically significant effects on children's assessed language or inhibitory control skills. These observed impacts are in contrast to the findings at the end of pre-K, in which there were no observed impacts on math skills. The next chapter explores the various factors that may have contributed to the pattern of findings in this study.

Chapter 4

Contributing Factors and Implications

Making Pre-K Count provided more, somewhat higher-quality, and more in-depth math instruction to children in prekindergarten. These impacts on teachers' math instruction were on top of a base of substantial math instruction already being conducted in pre-K as usual. At the end of pre-K, perhaps partially because of the large amount of math already in place, there were no consistent impacts on children's skills from the Making Pre-K Count program. By the end of the kindergarten year, using a different measure of children's math skills from the assessment used in pre-K, there was a small, positive, but not consistently statistically significant impact of the Making Pre-K Count program on students' math skills, and there were positive impacts on students' attitudes toward math and their executive function skills. Some of these positive effects at the end of kindergarten are in contrast to the lack of effects detected at the end of pre-K.

Two years of enhanced math instruction (Making Pre-K Count plus the High 5s kindergarten supplement) led to larger impacts on math skills than Making Pre-K Count alone. The impacts of two years of enhanced math were especially large compared with outcomes for a group of children who received no enhanced math instruction in either pre-K or kindergarten. These impacts on children's math skills appear to be driven primarily by participation in High 5s math clubs. However, all students who participated in High 5s had a solid base of pre-K math through participation in Making Pre-K Count, and this study did not test what the impact of High 5s would have been *without* the foundation that Making Pre-K Count provided. (As noted earlier, High 5s math clubs took place only in public school sites, where the effect of Making Pre-K Count on both math skills and math attitudes at the end of kindergarten was positive and statistically significant.)

Some of the findings from this newest wave of analyses are consistent with expectations about the benefits of two years of enhanced math on children's math learning and attitudes. Findings from the enhanced pre-K math instruction, however, are inconsistent across pre-K and kindergarten, and the pattern of findings does not match expectations or the findings from most of the previous studies of the Building Blocks program.[1] Despite prior research that demonstrated the efficacy of the Building Blocks curriculum, strong implementation of the program in New York City, and a moderate impact on teacher math practices in pre-K classrooms, impacts of the pre-K program alone on children's math skills (as measured by the Early Childhood Longitudinal Study-Birth Cohort [ECLS-B] or Woodcock-Johnson standardized

[1]For example, Clements and Sarama (2007, 2008) and Hofer, Lipsey, Dong, and Farran (2013).

tests) were small and not statistically significant. The pre-K program did have a positive effect on children's math attitudes and children's executive function in kindergarten. At the same time, there were strong impacts on math skills as measured by the Research-Based Early Math Assessment–Kindergarten (REMA-K) from two years of math enhancement in both pre-K and kindergarten. Analyses suggest that the REMA-K is more sensitive than the other measures of children's math skills but is not overly aligned with either of the math programs. A number of factors may help to explain why the pattern of findings differed from pre-K to kindergarten and across Building Blocks studies — particularly, the math environment children experienced in the absence of the two math programs and differences in the sensitivity of the measures used to assess children's math skills at different points in time.

- **The small impacts that resulted from the pre-K program may in part be due to the relatively high amounts of math that children in the pre-K-as-usual condition received.**

As described in Chapter 2, in early studies of Building Blocks, substantially less math instruction took place in pre-K as usual — with amounts of math averaging 19 minutes per morning across prior studies. In Making Pre-K Count, teachers in pre-K-as-usual classrooms engaged in 35 minutes of math instruction per morning, on average, and many used a published math curriculum or math computer software. In existing studies of the Building Blocks programs that calculated the amount of time spent on math in study classrooms, having more minutes of math in pre-K-as-usual classrooms was associated with smaller effects on children's math outcomes in pre-K.[2] The relationship is not statistically significant but may be constrained by the limited variation across the studies. That said, the cross-study comparison suggests one possible explanation for the small impacts on children's math learning in pre-K. The current study provides similar evidence. A subgroup analysis, described in Chapter 2, demonstrated that impacts on pre-K math skills were larger where there was less time spent on math instruction in pre-K-as-usual schools.

- **In contrast to pre-K, math instruction in the kindergarten-as-usual condition differed substantially from the instruction that students received in the High 5s supplemental program.**

As described in the companion report, in High 5s clubs all work was conducted in small groups, and activities involved interactive games and the heavy use of manipulatives.[3] During regular kindergarten classroom time, instruction was primarily provided in a whole-class

[2]Clements and Sarama (2008); Clements et al. (2011); Clements et al., "Effects on Mathematics" (2016); Hofer, Lipsey, Dong, and Farran (2013); Sarama et al. (2008); Morris, Mattera, and Maier (2016).

[3]Jacob, Erickson, and Mattera (2018).

format. Kindergarten classrooms also covered a narrower range of content than the High 5s clubs. Kindergarten math instruction focused primarily on numbers and operations, while High 5s included measurement, geometry, and some patterning activities. In addition, High 5s covered somewhat more advanced mathematical topics and at an earlier point in the year than did the kindergarten classrooms. To be clear, all children in High 5s also received kindergarten instruction; unlike the analysis in pre-K, the High 5s program did not replace the kindergarten experience, but supplemented it.

The greater contrast between what students in High 5s experienced and the nature of instruction in kindergarten classrooms may have contributed to the stronger impacts of the High 5s program. Kindergarten classroom instruction involved almost no small-group work and was primarily centered on activities with no materials or completing workbook pages.

- **The sensitivity of the math measures used at each time point may also have played a role in the pattern of findings.**

Three different math outcome measures were used in this study. As described in Chapter 3 (see Box 3.1), the REMA-K is a sensitive, detailed direct assessment used to measure children's math skills. Although the REMA-K was created by the developers of Building Blocks, analyses demonstrate that the items in the REMA-K are not closely aligned with either math program (and, in fact, removing the small number of aligned items does not substantively change the magnitude or statistical significance of the estimated impacts). It was used in this study to assess outcomes at the end of kindergarten, and it was used to assess math outcomes in pre-K and kindergarten in prior Building Blocks studies.[4] The Woodcock-Johnson Applied Problems was used in both pre-K and kindergarten in this study and is a more distal, standardized assessment of children's problem solving. The ECLS-B is a third measure that was used only at the end of the pre-K period in this study; it is a nationally normed measure of children's math skills designed for a longitudinal study of early childhood.

Effects on children's math were more robust as measured by the REMA-K than by the Woodcock-Johnson Applied Problems or ECLS-B assessments in Making Pre-K Count. As described in Chapter 3, the REMA-K proved to be more sensitive to treatment effects than the Woodcock-Johnson Applied Problems. The Woodcock-Johnson items tended to rely more heavily on language ability and to assess multiple aspects of math within one item, and the Woodcock-Johnson assessment had fewer items in specific math sub-content areas such as geometry. The REMA-K also seemed to be more sensitive than the ECLS-B. An examination of the difficulty of the items on the ECLS-B and the REMA-K and how they corresponded to the ability levels of this sample showed that the REMA-K had more items that were designed to

[4]Clements et al., "Effects of TRIAD" (2016); Hofer, Lipsey, Dong, and Farran (2013).

assess children with higher math ability levels, while the ECLS-B had more items designed to assess children with lower math ability levels. Prior subgroup findings in pre-K demonstrated stronger and more positive impacts of Making Pre-K Count in the pre-K year for children at the top of the distribution — suggesting that the REMA-K was more sensitive for the portion of the sample distribution where impacts from this math program may have clustered. The lack of sensitivity in pre-K measures may also help explain why pre-K findings from Making Pre-K Count do not match pre-K findings from prior studies.

An unpublished study of Building Blocks in San Diego suggests that some combination of both measurement and the math environment probably played a role in the patterns of results in Making Pre-K Count.[5] The San Diego study collected the REMA *and* the ECLS-B measures in both pre-K and kindergarten.[6] The study found a lack of math impacts in pre-K on both the ECLS-B and the REMA, followed by a small, positive impact in kindergarten on the REMA only. As in the Making Pre-K Count study, San Diego also had large amounts of math in the pre-K-as-usual condition.

- **The Making Pre-K Count study also differed from prior studies of Building Blocks in the characteristics of the child sample, although there is little evidence of differential impacts between children who entered the study predominantly speaking Spanish and those whose primary language was English.**

Compared with prior Building Blocks studies, the Making Pre-K Count sample was more heavily Hispanic and entered pre-K with higher cognitive scores, on average.[7] While there is little evidence that either of these sample characteristics directly led to differentially small impacts, it is still possible that some combination of sample characteristics played a role in how the findings from this study compare with prior work.

Summary and Implications

Making Pre-K Count and High 5s set out to explore whether providing young children with enhanced math instruction in early childhood could lead to learning gains that would translate into a persistent achievement boost across a variety of domains. To date, the findings from these studies suggest a number of takeaways:

---

[5]Clements et al., "Effects on Mathematics" (2016).

[6]The study used a version of the REMA designed for pre-K through fifth grade, known as the TEAM.

[7]Clements et al. (2011); Moiduddin et al. (2012); Peisner-Feinberg et al. (2014).

- **Sustained, aligned instruction in mathematics over multiple years is potentially important for children's math development.**

Children who received two years of aligned math enhancement had larger, more robust effects on math skills than students who received only one year or no enhanced instruction. While it is possible that the High 5s program would have been successful in raising students' math scores on its own, High 5s was implemented in combination with an enhanced pre-K experience, and thus the combined impacts may have resulted from a synergistic effect of the two programs together. Such a synergistic effect is consistent with findings from a prior study of Building Blocks that also included a supplemental kindergarten enrichment program added to the pre-K program and found sustained positive effects through first grade.[8]

The alignment between pre-K and kindergarten math instruction may have been one factor that contributed to the impacts of the High 5s program. Small-group enrichment programs, like High 5s, may be one short-term approach to making early elementary school more similar to what many children experience in pre-K. Many pre-K programs are centered around small-group instruction and involve hands-on activities. Small-group enrichment (delivered in a classroom as a "push-in" service or by pulling children out of class) is one way to recreate this environment in early elementary school settings. Further, these small-group experiences could be tailored to build on content that was covered in preschools.

- **Math clubs that deliver engaging small-group instruction can be successfully implemented in schools and have the potential to have an impact on children's math skills.**

High 5s was a new math program that has not been previously tested. At the outset of the study, it was not clear that the program could be successfully implemented and that it could make a difference for children's outcomes. As described in the companion report, the program was implemented well, with strong attendance and adherence to the curricular materials, and the number and content of club sessions were delivered as intended by the model.[9] The findings described here show that implementation of the math clubs led to positive impacts on children's math skills in kindergarten.

While there is an extensive body of research on the effectiveness of reading instruction in small groups[10] and on math tutoring for older students that has produced positive impacts,[11] at the time the High 5s program was implemented, there was much less evidence regarding the

[8]Clements et al., "Effects of TRIAD" (2016).

[9]Jacob, Erickson, and Mattera (2018).

[10]Slavin, Lake, Davis, and Madden (2010).

[11]Fryer (2011); Smith et al. (2013).

effectiveness of small-group tutoring for boosting math achievement in the early grades. This study adds to the growing body of evidence that suggests that small-group instruction, even when delivered by staff members at the paraprofessional level, can be an effective way to boost math skills.

The companion report lays out a number of mechanisms through which the program may have worked, including adding more math instructional time to children's school week, providing children with more individualized math instruction through the small-group format of the clubs, introducing children to somewhat more advanced mathematical content, and using a more interactive and less didactic instructional format. The program's positive effects on math demonstrate that some combination of these math instructional practices (more dosage, advanced content, instructional climate, and individualized instruction) may be important for changing children's math outcomes and may provide a guide for ways to improve kindergarten math instruction more generally.

- **As hypothesized, early math enrichment may have positive implications in domains beyond math.**

At the outset of this study, short-term improvements in math were hypothesized to spill over into other domains of children's learning, such as executive function. Empirical work in this area is quite limited, but there are a few studies from pre-K math interventions,[12] as well as from correlational research,[13] suggesting associations between math learning and executive function skills. Findings from this study demonstrate that a math intervention can have effects on executive function skills. These effects on executive function are noteworthy because it has been hypothesized that executive function may be important for supporting longer-term effects and reducing fade-out of early childhood programs.[14] It remains to be seen whether these impacts will persist over time and whether effects in other domains might emerge later for the children who participated in the Making Pre-K Count and High 5s programs.

- **Programs that target early math may bring the greatest gains in environments where there is a limited amount of math already occurring.**

Despite taking place in a shifting pre-K environment with growing amounts of math across years, Making Pre-K Count led to an additional 12 minutes of math instruction per morning. These 12 minutes of math, however, did not seem to operate in this study in the same way that they have in prior studies of Building Blocks. That is, more math instruction did not

[12]For example, Clements et al., "Effects on Mathematics" (2016).
[13]For example, Blair and Razza (2007).
[14]Gathercole, Pickering, Knight, and Stegmann (2004); McClelland et al. (2013); Ludwig (2011).

lead to better outcomes for children at the end of pre-K. Analyses suggest that one of the reasons for the limited impacts observed at the end of pre-K may have been the relatively high amount of math taking place in pre-K-as-usual classrooms. Indeed, where less math was occurring in pre-K-as-usual classrooms, the impact of Making Pre-K Count on math scores was substantially larger. This suggests that more high-quality math may matter most in an environment with very little math to begin with.

What's Next

Many questions remain about which impacts will be sustained into later elementary school for the children who participated in these two programs. One study of Building Blocks has tracked children's math skills into later elementary school and found a reemergence of impacts as children move further into elementary school, while other studies have shown a fade-out of math effects.[15] There is little information about what would lead to the different pattern of impacts across time in different studies. Making Pre-K Count and High 5s will uniquely be able to track the trajectory of effects across time and across domains to understand whether other processes in noncognitive domains may explain whether long-term impacts fade. The studies will continue to follow children into third grade to better understand the long-term effects of these early enhanced math programs. Data from these studies will also help inform the field about the longitudinal relationships between early math measures and later child outcomes.

The current study also raises questions about the relative benefits of providing enhanced math curricula in pre-K versus kindergarten and whether children benefit differentially from exposure in one year or the other. It is possible that it does not matter whether children learn basic math skills in the pre-K year or the kindergarten year; this would account for the fade-out effect often seen in pre-K studies as children's peers catch up once they learn the same skills in kindergarten. Alternatively, some have suggested that deep early math learning is a necessary stepping-stone on which children can build in subsequent years to develop more complex math understanding and skills.[16] In this study, the High 5s program was implemented with students who were first exposed to a high-quality pre-K math program. It is not clear what the effect of High 5s would have been in a context that did not provide the same aligned math opportunities.

The Making Pre-K Count study also examined additional factors — specifically, the business-as-usual environment and sample characteristics — that may play a role in explaining the study's findings. Evidence from a recent reanalysis of the federal Head Start Impact Study similarly shows that the business-as-usual environment had an influence on the magnitude of

[15]Clements et al., "Effects of TRIAD" (2016); Hofer et al. (2013).

[16]Aunola, Leskinen, Lerkkanen, and Nurmi (2004).

impacts from the Head Start intervention offered to children.[17] Further analyses will explore how these factors may have influenced the pattern of findings from prior Building Blocks studies or in other early childhood interventions.

Given the impact of the High 5s program on math skills, it is worth considering whether High 5s could be replicated in other contexts. It is not clear how the math enrichment in High 5s led to positive effects for children. Several possibilities exist, each of which would suggest a different way of using the program. High 5s was originally created as a program that occurred outside of regular instructional time and, as a result, added more math instructional time to a child's day (75 additional minutes of math a week). The program also focused on small-group, hands-on experiences characterized by instruction that differentiates according to the needs and skill levels of individual children. If the program works through enhancing the quality of math instruction that children receive, the High 5s math activities could be disseminated to kindergarten classrooms for teachers to use to improve the quality of math instruction, using either a "push-in" or a "pull-out" method to provide more opportunities for small-group instruction.

If High 5s were to operate as a pull-out or push-in program, using paraprofessional or other staff members already in the school building would offer greater flexibility in scheduling and would eliminate the need for program facilitators to commute between schools, which was challenging for many of the facilitators in this study. Such a model might have other benefits as well. Paraprofessionals in the school might have more experience working with small groups of children and might be better equipped to mitigate behavioral challenges that were faced by the less experienced facilitators in this study. As described in a companion report, different implementation models would have different implications for the cost of the program. In addition to the facilitators' time, schools would need to take into account the costs of materials, supervision, training, and administrative support when considering the costs of replicating the program.[18] Regardless of how the program is adapted for different schools' needs, High 5s presents a potential model for how to create more individualized, hands-on experiences for children during kindergarten math instruction.

---

[17]Feller, Grindal, Miratrix, and Page (2016); McCoy et al. (2016).

[18]Jacob, Erickson, and Mattera (2018).

Appendix A

The Role of the Amount of Math in Pre-K-as-Usual Classrooms in the Making Pre-K Count Findings

As described in the report, the amount of math conducted in New York City pre-K programs grew rapidly between 2013 and 2015. By 2015, the year that Making Pre-K Count was implemented and impacts were assessed, pre-K-as-usual classrooms in the study were conducting 35 minutes of math each morning — a much larger amount of math than the average of 19 minutes seen in other studies of Building Blocks. Despite this large amount, the program led to an additional 12 minutes of math (for a total of 47 minutes of math in a morning in Building Blocks classrooms). However, these large impacts on the amount of pre-K math instruction did not lead to concurrent effects on children's assessed math skills at the end of pre-K.

This pattern of findings — with no effect on math scores at the end of pre-K — does not match prior Building Blocks studies, in which impacts on the amount of math instruction led to positive impacts on children's pre-K math skills. There are a number of differences between the Making Pre-K Count study and prior Building Blocks studies. As noted, one difference is that previous studies had much less math instruction in place in control group classrooms. Could the large amount of math already taking place have mattered for how the program worked in New York City?

This appendix describes the results from exploratory analyses conducted to examine whether impacts of Making Pre-K Count in the pre-K year are larger in places where there was less math already taking place and, perhaps, where the program could have made the biggest difference for children's math experience. Subgroup analyses were conducted to compare outcomes for children in random assignment blocks in which pre-K-as-usual classrooms spent fewer than 32 minutes on math each morning with outcomes in blocks in which the pre-K-as-usual classrooms spent more than 32 minutes on math. The analytic sample comprises 698 program group children and 691 control group children (320 and 323 of whom, respectively, are in the "low control group math minutes" blocks).

The findings in Appendix Table A.1 are consistent with the hypothesis. In blocks with *less* math instruction in pre-K-as-usual classrooms, there were positive, statistically significant impacts (effect size = 0.18) on the Woodcock-Johnson Applied Problems subtest of children's math skills at the end of pre-K. In contrast, in blocks with *more* math instruction in pre-K-as-usual classrooms, impacts on the Woodcock-Johnson subtest at the end of pre-K were slightly negative and not statistically significant (effect size = -0.02). The difference between the subgroups is marginally significant ($p = 0.11$). The impacts on the ECLS-B math measure in both the low and high groups was slightly positive but not statistically significant. It is important to note that these exploratory findings are suggestive but not conclusive. This pattern may be caused by other factors that are also associated with the amount of control group math but are not included in this analysis — for example, blocks with less math may also have less rigorous instructional practices overall.

Appendix Table A.1

Math Impacts in the Spring of the Pre-K Year, by Amount of Math in Control Group Classrooms

	Low Control Group Math Minutes[a]				High Control Group Math Minutes[b]						
Outcome	Control Group Mean	Difference (Impact)	P-Value	Effect Size[c]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[c]	Difference Between Subgroups	P-Value	Sig.
ECLS-B scale score[d]	26.53	0.42	0.46	0.07	26.74	0.20	0.75	0.03	0.22	0.79	
Woodcock-Johnson Applied Problems[e]	101.02	2.28	0.02 **	0.18	101.39	-0.32	0.80	-0.02	2.59	0.11	
Sample size											
Blocks	7				9						
Sites	14				20						
Children	323				368						

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2015.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

Rounding may cause slight discrepancies in sums and differences.

[a]"Low control group math minutes" comprises blocks where the average amount of math recorded during a morning observation in pre-K control group classrooms was less than 31.5 minutes.

[b]"High control group math minutes" comprises blocks where the average amount of math recorded during a morning observation in pre-K control group classrooms was greater than or equal to 31.5 minutes.

[c]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[d]Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) math assessment (Najarian et al., 2010). The potential score range is from 0 to 44.

[e]Woodcock-Johnson Applied Problems is a subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and

Appendix B

Sampling and Baseline Equivalence of Children Across Program and Control Groups

This appendix provides a summary of the selection, recruitment, and random assignment for the samples in Making Pre-K Count (MPC) and High 5s. The pre-K sample selection was described in more detail in a previous report.[1]

Sixty-nine pre-K programs housed in New York City public schools and community-based organizations throughout Brooklyn, the Bronx, Manhattan, and Queens were selected to participate in Making Pre-K Count. Sites were selected to reflect the geographical, racial, and ethnic diversity of New York City's low-income population,[2] although the sample was not designed to be a statistically representative sample. Sites had to serve a low-income population of 4-year-old children and offer full-day programs. Thirty-five sites were randomly assigned to receive Building Blocks and extensive professional development (the program group), while the remaining 34 were assigned to continue their typical pre-K programming (the "pre-K-as-usual" or control group).

There were 2,702 eligible children in study classrooms. Of those children, 96 percent were successfully tracked from pre-K to kindergarten, with similar child attrition across groups (3.9 percent attrition in the Making Pre-K Count program group and 4.9 percent attrition in the control group). Out of the total sample of eligible children, a subset of children were randomly chosen for data collection in kindergarten. Comparing outcomes for children in pre-K sites that implemented Making Pre-K Count with children in pre-K sites that continued business as usual allows for an investigation of the effects of *one year of math enrichment* (Making Pre-K Count) compared with *no math enrichment*.

As a supplement to the Making Pre-K Count program, High 5s provided a second year of math enrichment in kindergarten for children in the public school sites that had implemented Building Blocks in pre-K. All 24 Making Pre-K Count *program group* sites that were in public schools implemented High 5s. Children were eligible for High 5s if their parents consented to their participation in the program and if they stayed in the same public school from pre-K to kindergarten. In contrast to Making Pre-K Count, where *sites* were randomly assigned to a program group or control group, for High 5s *individual children* within Making Pre-K Count program group public schools were randomly assigned to the kindergarten program or kindergarten as usual.

In the spring of pre-K, 655 parents gave consent for their children to participate in the High 5s program; the children were randomly assigned when they got to kindergarten in the fall. This random assignment process created two groups in kindergarten within the 24 public

[1]Morris, Mattera, and Maier (2016).

[2]A "low-income population" was defined as at least 70 percent of children being eligible for free or reduced-price lunch.

schools that implemented Making Pre-K Count in pre-K: children who received Making Pre-K Count in pre-K and High 5s in kindergarten (MPC plus High 5s supplement, n = 320) and children who received Making Pre-K Count in pre-K and kindergarten as usual (Making Pre-K Count only, n = 335). Of those children, 8.5 percent left the school by spring. Children who stayed in the same school and did not consent to participate in the High 5s program (n = 18) remained in the original Making Pre-K Count sample and were selected for kindergarten data collection.[3] Although these children did not attend High 5s, they were randomly assigned in a second phase to either the MPC plus High 5s supplement group or the Making Pre-K Count group and were included in the impact analysis to maintain the internal validity of the random assignment design when comparisons were made with the no-math-enrichment group (described below).

Children in the MPC plus High 5s supplement group can also be compared with children in a third group: children in the 23 *control group* public schools who stayed in the same school from pre-K to kindergarten. The two-stage sequential random assignment design created three experimental groups, which can be used to rigorously investigate the effects of *two years of math enrichment* (MPC plus High 5s supplement) compared with *one year of math enrichment* (Making Pre-K Count) and the effects of *two years of math enrichment* (MPC plus High 5s supplement) compared with *no math enrichment* (the pre-K-and-kindergarten-as-usual control condition). These comparisons focus on the sample of 47 *public schools* in the Making Pre-K Count design.

Making Pre-K Count in Pre-K Versus Pre-K as Usual (n = 1,325)

For this analysis, kindergarten outcomes were compared for children who attended the 35 pre-K sites that implemented Making Pre-K Count versus children who attended the 34 pre-K sites that continued business as usual. The analysis estimates the effect that one year of math enrichment in pre-K had on kindergarten child outcomes compared with no math enrichment in pre-K. Because High 5s took place in some of the Making Pre-K Count sites (the public schools in the program group), a subset of children for the pre-K analytical sample received the High 5s program as well (n = 189). To estimate the unique effect of only Making Pre-K Count in pre-K, this analysis removed those children from the analytical sample. Simply removing children, however, could lead to an unbalanced sample (with program group public school children underrepresented in the sample). To account for this, children who did not receive High 5s were oversampled in these Making Pre-K Count public school program group sites (n = 132). A Wald test of joint significance was used to test whether those 189 High 5s program group

[3]The consent rate was fairly high, with 95 percent of eligible children who stayed in the same school consenting to participate in the High 5s intervention.

children differed from the 132 children who were part of Making Pre-K Count only. The test result (see Appendix Table B.1) indicates that the aforementioned subsamples were not systematically different along parent and child demographic characteristics.

This analysis therefore comprises children who received *only* the Making Pre-K Count program and children who received no math enrichment in pre-K or kindergarten (n = 1,325). The program and control groups are balanced across parent and child demographic characteristics. A Wald test of joint significance was used to determine whether parent and child baseline demographic characteristics were predictive of a child being in the pre-K program group or the pre-K-as-usual control group. The result of this test (Table B.2) indicates that program group children and control group children did not systematically differ along baseline demographic characteristics.

MPC Plus High 5s Supplement Versus Making Pre-K Count (n = 613)

This analysis compared kindergarten outcomes for children who received *two years of math enrichment* (MPC plus High 5s supplement) with outcomes for children who received *one year of math enrichment* (Making Pre-K Count). This analysis builds on the individual-level random assignment design and includes only children from the 24 Making Pre-K Count public school program sites. Children are included in this sample if their parents gave positive consent to participate in High 5s and if they intended to and did stay in the same school from pre-K to kindergarten. Eighteen children did not consent to participate in the High 5s program but were still part of the larger Making Pre-K Count study and thus had their data collected in kindergarten. These children were randomly assigned in a second phase to the MPC plus High 5s supplement group or the Making Pre-K Count group (using the same random assignment ratio as the children with consent) and were included in the impact analysis as nonattenders to maintain the internal validity of the random assignment design when making comparisons with the no-math-enrichment group. (That analysis is described below.) The final analytic sample included only children who stayed in the same public school, for a total of 303 children in the MPC plus High 5s supplement group and 310 children in the Making Pre-K Count group.

A Wald test of joint significance (see Table B.3) indicated that the two groups of children were not systematically different along baseline demographic characteristics and pre-K assessment performance.

Appendix Table B.1

Baseline Equivalence: High 5s Recipients Removed from Making Pre-K Count Analytic Sample and Their Replacements

Characteristic	Children Removed from Sample	Replacement Group
Parent demographics		
Race and ethnicity (%)		
Hispanic	48.7	51.5
Non-Hispanic white	4.9	9.2
Non-Hispanic black	41.6	30.8
Other/multiracial[a]	4.9	8.5
Highest level of education (%)		
At least high school/GED	76.8	73.9
Child demographics		
Average age in fall 2014 (years)[b]	4.17	4.16
Female (%)	52.9	50.8
English-speaking (%)[c]	89.8	84.7
Joint test of difference between groups[d]	(F-value = 0.77)	
Sample size	189	132

SOURCE: MDRC calculations from parents' reports on demographics on the informed consent form (ICF) collected in fall 2014.

NOTES: GED = General Educational Development certificate.
Rounding may cause slight discrepancies in sums and differences.
[a]Other includes Asian, Native Hawaiian/Pacific Islander, and American Indian/Alaska Native.
[b]This is the age at the beginning of pre-K as of September 1, 2014.
[c]Child speaks and understands English as reported on the ICF.
[d]A Wald test was used to determine whether there was a systematic difference between the two samples based on the characteristics included in this table.

Appendix Table B.2

Baseline Equivalence: Making Pre-K Count Versus Pre-K as Usual

Characteristic	Program Group	Control Group
Parent demographics		
Race and ethnicity (%)		
Hispanic	55.9	58.7
Non-Hispanic white	6.0	1.2
Non-Hispanic black	34.0	35.9
Other/multiracial[a]	4.1	4.2
Highest level of education (%)		
At least high school/GED	74.6	69.3
Child demographics		
Average age in fall 2014 (years)[b]	4.16	4.17
Female (%)	50.4	52.6
English-speaking (%)[c]	90.5	87.7
Joint test of difference between groups[d]	(F-value = 0.03)	
Sample size	641	684

SOURCE: MDRC calculations from parents' reports on demographics on the informed consent form (ICF) collected in fall 2014.

NOTES: GED = General Educational Development certificate.

The program group received Making Pre-K Count in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Rounding may cause slight discrepancies in sums and differences.

[a]Other includes Asian, Native Hawaiian/Pacific Islander, and American Indian/Alaska Native.

[b]This is the age at the beginning of pre-K as of September 1, 2014.

[c]Child speaks and understands English as reported on the ICF.

[d]A Wald test was used to determine whether there was a systematic difference between the two samples based on the characteristics included in this table.

Appendix Table B.3

Baseline Equivalence: MPC Plus High 5s Supplement Versus Making Pre-K Count

Characteristic	Program Group	Control Group
Parent demographics		
Race and ethnicity (%)		
Hispanic	51.3	55.5
Non-Hispanic white	4.7	6.3
Non-Hispanic black	37.9	33.3
Other/multiracial[a]	6.0	5.0
Highest level of education (%)		
At least high school/GED	77.0	73.6
Child demographics		
Average age in fall 2014 (years)[b]	4.17	4.17
Female (%)	55.8	49.4
English-speaking (%)[c]	89.7	86.7
Child skills at the end of pre-K (mean)		
Math		
ECLS-B math score (0-44)[d]	27.33	26.99
Woodcock-Johnson Applied Problems Standard Score[e]	102.70	101.89
Language		
ROWPVT Standard Score[f]	97.04	96.50
Executive function		
Pencil Tap: proportion correct (0-1)[g]	0.76	0.74
Arrows Mixed: proportion correct (0-1)[h]	0.83	0.79
Corsi Blocks forward: number correct[i]	2.96	2.99
PSRA Attention and Inhibition Score (0-3)[j]	2.68	2.67
Joint test of difference between groups[k]	(F-value = 0.71)	
Sample size	303	310

(continued)

Appendix Table B.3 (continued)

SOURCES: MDRC calculations from parents' reports on demographics on the informed consent form (ICF) collected in fall 2014 and the direct child assessments administered in spring 2015.

NOTES: GED = General Educational Development certificate.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group received only Making Pre-K Count in pre-K.

Rounding may cause slight discrepancies in sums and differences.

[a]Other includes Asian, Native Hawaiian/Pacific Islander, and American Indian/Alaska Native.

[b]This is the age at the beginning of pre-K as of September 1, 2014.

[c]Child speaks and understands English as reported on the ICF.

[d]Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) math assessment (Najarian et al., 2010). The potential score range is from 0 to 44.

[e]Woodcock-Johnson Applied Problems is a child math assessment included in the battery of tests in the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001). The score is age normalized to 100, with a standard deviation of 15.

[f]Receptive One-Word Picture Vocabulary Test (Martin and Brownell, 2011). The score is age normalized to 100, with a standard deviation of 15.

[g]Pencil Tap task (Luria, 1966; Diamond and Taylor, 1996). The score reports the total number of trials (out of 16) that a child got correct.

[h]Spatial Conflict Arrows task (Willoughby, Wirth, Blair, and Family Life Project Investigators, 2012). This score is calculated by dividing the number of correct responses for "mixed" trials where arrows were depicted either laterally (with left-pointing arrows appearing on the left side of the tablet screen and right-pointing arrows appearing on the right side) or contralaterally (with left-pointing arrows appearing on the right side of the tablet screen and right-pointing arrows appearing on the left side) by the total number of mixed lateral and contralateral trials.

[i]Corsi Blocks (Corsi, 1972; Lezak, 1983). The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

[j]Children's self-regulation skills were measured using the Preschool Self-Regulation Assessment (PSRA; Smith-Donald, Raver, Hayes, and Richardson, 2007).

[k]A Wald test was used to determine whether there was a systematic difference between the two samples based on the characteristics and measures included in this table.

MPC Plus High 5s Supplement Versus Pre-K and Kindergarten as Usual (n = 648)

Kindergarten outcomes for children who received *two years of math enrichment* (MPC plus High 5s supplement) were also compared with outcomes for children who received *no math enrichment* (pre-K-and-kindergarten-as-usual control condition). This analysis builds on the cluster-level random assignment design and compares children from the 24 Making Pre-K Count program group public schools with children from the 23 Making Pre-K Count control group public schools.[4]

As with the previous analysis, public school children were eligible if they stayed in the same school from pre-K to kindergarten. Mobility rates between the two groups were similar.

[4]During spring 2016 data collection, no assessments were completed for children in one control group public school because of issues with locating children. The sample of assessed children represents 22 control group public school sites.

Of the 890 Making Pre-K Count control group public school children, 522 children, or 58.7 percent, intended to and did stay at the same school from pre-K to kindergarten. Of the 940 Making Pre-K Count program group public school children, 622 children, or 66.2 percent, intended to and did stay at the same school from pre-K to kindergarten. In the control group, 345 children were randomly selected to participate in kindergarten data collection. The MPC plus High 5s supplement group, as described above, contains 303 children (including both High 5s consenters and nonconsenters) who were randomly assigned to one of two groups from Making Pre-K Count program group public schools.

A Wald test of joint significance (see Table B.4) indicated that the two groups were not systematically different along baseline demographic characteristics.

Appendix Table B.4

Baseline Equivalence: MPC Plus High 5s Supplement Versus Pre-K and Kindergarten as Usual

Characteristic	Program Group	Control Group
Parent demographics		
Race and ethnicity (%)		
Hispanic	51.3	60.8
Non-Hispanic white	4.7	0.9
Non-Hispanic black	37.9	34.4
Other/multiracial[a]	6.0	3.9
Highest level of education (%)		
At least high school/GED	77.0	67.8
Child demographics		
Average age in fall 2014 (years)[b]	4.17	4.16
Female (%)	55.8	51.0
English-speaking (%)[c]	89.7	86.6
Joint test of difference between groups[d]	(F-value = 1.40)	
Sample size	303	345

SOURCE: MDRC calculations from parents' reports on demographics on the informed consent form (ICF) collected in fall 2014.

NOTES: GED = General Educational Development certificate.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group did not receive math enrichment and participated in pre-K and kindergarten as usual.

Rounding may cause slight discrepancies in sums and differences.

[a]Other includes Asian, Native Hawaiian/Pacific Islander, and American Indian/Alaska Native.

[b]This is the age at the beginning of pre-K as of September 1, 2014.

[c]Child speaks and understands English as reported on the ICF.

[d]A Wald test was used to determine whether there was a systematic difference between the two samples based on the characteristics included in this table.

Appendix C

Kindergarten Impacts Analysis Models

Appendix C describes the analytic decisions and empirical methods used for the three main kindergarten analyses:

1. Making Pre-K Count in pre-K compared with no math enrichment
2. Making Pre-K Count in pre-K and High 5s supplement in kindergarten compared with Making Pre-K Count only in pre-K
3. Making Pre-K Count in pre-K and High 5s supplement in kindergarten compared with no math enrichment

As in previous study analyses, multiple imputation (n = 50) was used to address issues of missing covariate data, using the PROC MI procedure in SAS. Baseline data on children (demographic and outcome variables) and blocks were used to estimate values for missing covariates across 50 data sets. Using PROC MI ANALYZE, impact estimates were created for each of the 50 data sets and then combined, using a standard computation to correct for added uncertainty in the standard errors. Data were not imputed for missing outcome variables. Sensitivity tests using unimputed and site-level mean imputed covariates yielded similar impact estimates.

Making Pre-K Count in Pre-K Versus Pre-K as Usual

To estimate the effect of one year of math enrichment in pre-K on kindergarten outcomes, this analysis compares the kindergarten outcomes for children who attended the 35 pre-K programs that implemented Making Pre-K Count with outcomes for children who attended 34 pre-K programs that continued business as usual. For detailed information on the sample refer to Appendix B.

Program impacts were estimated by comparing mean outcomes for the group assigned to Making Pre-K Count in pre-K with corresponding means for the control group with an adjustment for selected background characteristics and dummy variables for the random assignment blocks. Multilevel modeling was used to account for the nested structure of the data. Although children in this analysis were in their kindergarten year and the pre-K site no longer accounted for a large portion of shared variance, random assignment for this portion of the study occurred at the pre-K *site* level within random assignment blocks; therefore, this analysis is clustered at the pre-K site level. Children are nested within pre-K sites and sites are nested within random assignment blocks. A set of dummy variables representing each random assignment block was included at the site level in the impact analysis.

The analysis included a standard set of covariates used in previous analyses to improve the precision of the impact estimates, thereby increasing the capability to detect true impacts and reducing the likelihood that any differences between the program and control groups are due to random variation in the sample. Covariates for this analysis include parental education (a dummy variable for whether the parent had a high school diploma or equivalent or a higher degree), child age at time of kindergarten assessment, and baseline measures for English language proficiency (assessed by pre-LAS),[1] executive function abilities (inhibition and cognitive flexibility),[2] attention and inhibition,[3] and a pretest if available. Two additional covariates were used for math outcomes — baseline levels of receptive language (ROWPVT) and a measure of the child's baseline executive function (Corsi Blocks forward score).[4]

The following two-level model was used for kindergarten child outcomes:

Level 1: Children within sites

$$Y_{sc} = \alpha_{0c} + \sum_{i>0} \alpha_i X_{isc} + \varepsilon_{sc}$$

Level 2: Sites within blocks

$$\alpha_{0c} = \sum_{b=1}^{16} \gamma_b Z_{bc} + \Pi\, T_c + \upsilon_c$$

where:

Y_{sc} = the outcome for student *s* from site *c*

X_{isc} = baseline characteristic *i* for student *s* from site *c*

Z_{bc} = an indicator variable for random assignment block *b*, which is equal to one if site *c* is in random assignment block *b* and zero otherwise

T_c = the treatment indicator, which is equal to one if site *c* was randomized to treatment (MPC) and zero if it was randomized to control status

ε_{sc} = a random error for student *s* from site *c* that is independently and identically distributed across students in sites

---

[1]Pre-Language Assessment Scales (Duncan and De Avila, 1998).

[2]Spatial Conflict Arrows task (Willoughby, Wirth, Blair, and Family Life Project Investigators, 2012).

[3]Preschool Self-Regulation Assessment (Smith-Donald, Raver, Hayes, and Richardson, 2007).

[4]Receptive One-Word Picture Vocabulary Test (ROWPVT; Martin and Brownell, 2011); Corsi Blocks (Corsi, 1972).

υ_c = a random error for site *c* that is independently and identically distributed across sites

Making Pre-K Count (MPC) Plus High 5s Supplement Versus Making Pre-K Count

This analysis compares the kindergarten outcomes for children assigned to *two years of math enrichment* (MPC plus High 5s supplement) with outcomes for children assigned to *one year of math enrichment* (Making Pre-K Count). This analysis builds on the individual-level random assignment design and includes only children from the 24 Making Pre-K Count public school intervention sites. For detailed information on the sample refer to Appendix B.

Program impacts were estimated by comparing mean outcomes for children assigned to High 5s with corresponding means for the control group with an adjustment for selected background characteristics and a dummy variable for each pre-K school. Random assignment for High 5s took place in the fall of the kindergarten year; therefore, baseline covariates for children in this analysis come from the spring of the pre-K year. Covariates for this analysis include parental education (a dummy variable for whether the parent had a high school diploma or equivalent or a higher degree), child age at time of kindergarten assessment, child's English language proficiency in the fall of pre-K (assessed by pre-LAS),[5] and spring of pre-K (baseline) measures for levels of receptive language,[6] math skills,[7] and executive function abilities (attention, inhibition, cognitive flexibility, and working memory).[8]

In High 5s, *individual children* within the 24 Making Pre-K Count program group public schools were randomly assigned to the kindergarten intervention or kindergarten as usual. The model employs fixed effects for pre-K site. The following single-level model for child outcomes is used:

$$Y_s = \alpha_0 + \sum_{i=1}^{10} \alpha_i X_{is} + \pi T_s + \sum_{c=1}^{24} \alpha_c Z_{cs} + \varepsilon_s$$

where:

[5]Pre-Language Assessment Scales (Duncan and De Avila, 1998).

[6]Receptive One-Word Picture Vocabulary Test (ROWPVT; Martin and Brownell, 2011).

[7]ECLS-B (Najarian, Snow, Lennon, and Kinsey, 2010); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001).

[8]Pencil Tap task (Luria, 1966; Diamond and Taylor, 1996); Preschool Self-Regulation Assessment (Smith-Donald, Raver, Hayes, and Richardson, 2007); Corsi Blocks (Corsi, 1972; Lezak, 1983); Spatial Conflict Arrows task (Willoughby, Wirth, Blair, and Family Life Project Investigators, 2012).

Y_s = the outcome for student *s*

X_{is} = baseline characteristic *i* for student *s*

Z_{cs} = an indicator variable for school *c* for student *s*

T_s = the treatment indicator, which equals one if student *s* was randomized to treatment (High 5s) and zero if the student was randomized to control status

ε_s = a random error for student *s* that is independently and identically distributed across students in classrooms

MPC Plus High 5s Supplement Versus Pre-K and Kindergarten as Usual

Kindergarten outcomes for children assigned to *two years of math enrichment* (MPC plus High 5s Supplement) were also compared with outcomes for children who received *no math enrichment* (pre-K-and-kindergarten-as-usual control condition). This analysis builds on the cluster-level random assignment design and compares children from the 24 Making Pre-K Count program group public schools with children from the 23 Making Pre-K Count control group public schools.[9] For detailed information on the sample refer to Appendix B.

Program impacts were estimated by comparing mean outcomes for the group assigned to Making Pre-K Count in pre-K and High 5s in kindergarten with corresponding means for the control group with an adjustment for selected background characteristics and dummy variables for the random assignment blocks. Multilevel modeling was used to account for the nested structure of the data. Although children in this analysis were in their kindergarten year and the pre-K site no longer accounted for a large portion of shared variance, random assignment for this portion of the study occurred at the pre-K *site* level within random assignment blocks; therefore, this analysis is clustered at the pre-K site level. Children are nested within pre-K sites and sites are nested within random assignment blocks. A set of dummy variables representing each random assignment block was included at the site level in the impact analysis.

The analysis included a standard set of covariates used in previous analyses to improve the precision of the impact estimates, thereby increasing its capability to detect true impacts and reducing the likelihood that any differences between the program and control groups are due to

[9]During spring 2016 data collection, no assessments were completed for children in one control group public school due to issues with locating children. The sample of assessed children represents 22 control group public school sites.

random variation in the sample. Covariates for this analysis include parental education (a dummy variable for whether the parent had a high school diploma or equivalent or a higher degree), child age at time of kindergarten assessment, and baseline measures for English language proficiency (assessed by pre-LAS),[10] executive function abilities (inhibition and cognitive flexibility),[11] attention and inhibition,[12] and a pretest if available. Two additional covariates were used for math outcomes — baseline levels of receptive language (ROWPVT) and a measure of the child's baseline executive function (Corsi Blocks forward score).[13]

This analysis used the same analytic model as the analysis comparing one year of pre-K math enrichment (Making Pre-K Count in pre-K) with no math enrichment (pre-K as usual).

[10]Pre-Language Assessment Scales (Duncan and De Avila, 1998).

[11]Spatial Conflict Arrows task (Willoughby, Wirth, Blair, and Family Life Project Investigators, 2012).

[12]Preschool Self-Regulation Assessment (Smith-Donald, Raver, Hayes, and Richardson, 2007).

[13]Receptive One-Word Picture Vocabulary Test (ROWPVT; Martin and Brownell, 2011); Corsi Blocks (Corsi, 1972).

Appendix D

Measurements: REMA-K and Woodcock-Johnson

As described in Chapter 3, two measures were used to assess children's kindergarten math skills in this study: the Research-Based Early Math Assessment–Kindergarten (REMA-K) and the Woodcock-Johnson Applied Problems subscale.[1]

The REMA-K is a direct assessment that measures thinking and learning along research-based developmental progressions for math topics. This study used an adaptation of the Research-Based Early Math Assessment created by Douglas Clements, Julie Sarama, and Xiufeng Liu. Item selection represents the full range of early mathematics competencies applicable within the prekindergarten, kindergarten, and first-grade years.

The Woodcock-Johnson Applied Problems is a subscale of the Woodcock-Johnson III Tests of Achievement. It is a valid standardized assessment of mathematical thinking for ages 2 through 90; early items are suitable for assessing simple math functions relevant at young ages (such as identifying the number when more objects are added to a picture). The Woodcock-Johnson Applied Problems test is a less detailed, more global measure of children's math ability than the REMA-K.

In this study, the REMA-K was a more sensitive measure of children's math skills than the Woodcock-Johnson Applied Problems. This appendix lays out a number of factors that may have played a role in the measure's sensitivity.

Alignment

The REMA-K was created by the developers of Building Blocks; therefore, it was possible that any sensitivity in the REMA-K was caused by the developers having created both the program and the measure. To examine overalignment, MDRC-based operational staff members formally trained on the Building Blocks programs and Building Blocks training reviewed each REMA-K item for alignment with Building Blocks activities. These analyses suggest the REMA-K is not closely aligned with the Building Blocks and High 5s programs. Only 10 percent of the items on the REMA-K (5 out of 48) were similar in approach to the types of activities to which children were exposed in Building Blocks. For example, one question on the REMA-K asks the following, "Pretend I give you three candies and then two more. How many will you have altogether?" In the Building Blocks game Dinosaur Shop, the child needs to know that if there is a request for three blue dinosaurs and two red dinosaurs, the request is for a total of five dinosaurs. The activity approaches the same mathematical question in the same way. In a sensitivity analysis in which the more aligned items were removed from the score, the magnitude and statistical significance of the estimated effects described in this report did not change.

[1]Clements, Sarama, and Liu (2008); Woodcock, McGrew, and Mather (2001).

A prior study of Building Blocks in Boston also examined the appropriateness of a shortened version of the REMA for assessing program impacts.[2] The results from that psychometric analysis align with the content analysis conducted in this study. The Boston study found that the measure was valid for assessing children's math skills and that the REMA items did not function differentially for program and control group children.

A similarly small percentage of REMA-K items were aligned with High 5s activities, which were developed at the University of Michigan. Although the High 5s content was designed to be aligned with the content covered in the Building Blocks program, most activities were independently created by the study team.

Range of Content

The REMA-K is a lengthier test than the Woodcock-Johnson Applied Problems; it includes more items with the intention of obtaining granular information about children's abilities in each mathematical skill at each level. An analysis was conducted to categorize each item on the REMA-K and Woodcock-Johnson Applied Problems.

As shown in Appendix Table D.1, the REMA-K covers a wider range of math content areas. Both measures have the largest focus on numbers and operations, with numbers or operations making up 56 percent of items for the REMA-K and nearly all items (84 percent) on the Woodcock-Johnson Applied Problems. The REMA-K, however, also has 21 items devoted to geometry, geometric measurement, patterning, relational thinking or early algebra, and measurement. The Woodcock-Johnson includes only five items focused on geometry, geometric measurement, measurement, and spatial thinking. Building Blocks has a unique focus on geometry; the REMA-K includes nine questions about geometry or geometric measurement, while the Woodcock-Johnson includes only three items. Despite this unique focus on geometry, in this study there were impacts on the REMA-K across the numeracy, patterning, and geometry subscales.

Specificity of Content

Each assessment item is designed to provide precise information about a specific skill set. In the REMA-K, most items tap only one specific skill set at a time. Within that skill set, the REMA-K items are able to dig deeply into each skill. For example, the REMA-K is unique in that it has many items that ask children to do mental math (with questions such as "how much is four plus

[2]Weiland and Yoshikawa (2013).

Appendix Table D.1

Raw Items per Content Area

Primary Math Content Area	REMA-K	WJ III: Applied Problems
Numbers	7	9
Operations	20	17
Geometry	5	2
Geometric measurement	4	1
Patterning	4	0
Relational thinking/early algebra	4	0
Measurement	4	1
Spatial thinking	0	1
Total	48	31

SOURCES: Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008) and the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001).

three" or "how much is seven plus eight," rather than using objects). This type of question allows the REMA-K to assess not just a child's general ability in one area (for example, addition), but the child's skill level within that area (for example, addition with numbers below five or addition with numbers above five). In comparison, many of the Woodcock-Johnson items assess a mix of math content areas within one item. The Woodcock-Johnson assessment also draws upon children's language skills in addition to math.

Appendix Box D.1 presents an example of how a REMA-K item and a Woodcock-Johnson item each assess addition. The REMA-K item examines children's "pure" adding skills — the ability to add two quantities together. In the Woodcock-Johnson item, the child must be

Appendix Box D.1

Example of How the REMA-K and Woodcock-Johnson Tests Assess Addition

Research-Based Early Math Assessment–Kindergarten (REMA-K), Item 45: The assessor places three pennies under a cloth in full view of the child and puts six on the table. The assessor says, "Here are six pennies. There are three more under this cloth. How many are there in all?"

Woodcock-Johnson Applied Problems, Item 25: The assessor shows the child a picture of five pennies and a nickel and asks, "When added together, how much money is this?"

able to identify pennies and nickels, understand the value of each of the coins, and then add them together. For the answer to be correct, the word "cents" must follow the numeral. Therefore, the Woodcock-Johnson in this case is not a clear and sensitive assessment of a child's addition abilities; whether the child answers correctly relies not just on addition skills, but also language ability and knowledge of coin value.

Appendix E

Sensitivity Analyses for the Impacts of Making Pre-K Count on the REMA-K Math Measure

In the studies in this report, children's math skills in kindergarten were assessed using two measures: the Research-Based Early Math Assessment–Kindergarten (REMA-K)[1] and the Woodcock-Johnson Applied Problems subscale.[2] The REMA-K is a 15- to 20-minute, detailed direct assessment based on research about children's mathematical knowledge that measures mastery of numbers, operations, measurement, patterning, and geometry.

Sensitivity analyses were conducted to examine how robust the unique impacts of Making Pre-K Count were to different scoring specifications. Three different scores can be created from the REMA-K:

- The raw score is the number of items a child answered correctly.
- The logit score is calculated with a Rasch model, which takes into account item difficulty. It is intended to capture each child's ability level on the tested dimension or skill.
- The T score is a linear conversion of the logit score to a standardized score with a mean of 50 and standard deviation of 10, which improves interpretability. The T score has been used in previous studies of Building Blocks and is therefore the primary score used in this study.[3]

In addition, sensitivity analyses were conducted to examine how robust the impact findings were to different ways of specifying the sample for the analysis of the unique effect of Making Pre-K Count. In the primary analysis, 189 children who received the High 5s program in kindergarten in addition to the Making Pre-K Count program in pre-K were removed and were replaced by 132 children who received only Making Pre-K Count in pre-K and kindergarten as usual (see Appendix B for details).

- The first analysis (sensitivity sample 1) removed the group of 189 children who also received the High 5s supplement from the Making Pre-K Count kindergarten analytic sample. But instead of replacing these children, this analysis compensated for their removal by up-weighting the 183 Making Pre-K Count program group children already included in the analysis who did not receive the High 5s supplement.
- The second analysis (sensitivity sample 2) also removed the group of 189 children who received the High 5s supplement from the sample, and the re-

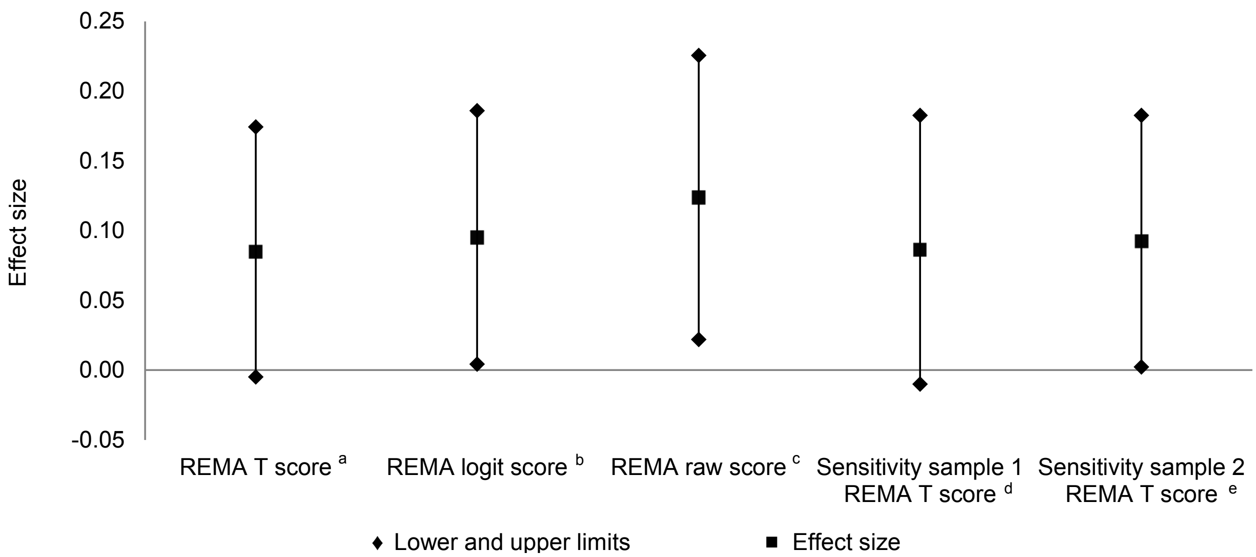
[1]Clements, Sarama, and Liu (2008).
[2]Woodcock, McGrew, and Mather (2001).
[3]Impacts presented in the body of the report are for this primary T score.

placement group of 132 Making Pre-K Count program group children who did not receive the High 5s supplement was added to the sample, as in the primary analysis. But in this analysis, the replacement group was weighted up to be equivalent to the size of the group that was removed.

The impact of Making Pre-K Count on the REMA-K is small but positive across five different measurement and model specifications and is statistically significant in three out of five of them. Appendix Figure E.1 displays the 90 percent confidence intervals for the impact estimates using each of three REMA-K scores and using the T scores for the two sensitivity samples.

Appendix Figure E.1

90 Percent Confidence Intervals for Making Pre-K Count Effect Sizes on the Research-Based Early Math Assessment–Kindergarten (REMA-K) for Variations in Scoring and Sample Specifications



SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: [a]The T score is a linear conversion of the logit score to a standardized score with a mean of 50 and standard deviation of 10, which improves interpretability. The T score has been used in previous studies of Building Blocks and is therefore the primary score used in this study.

[b]The logit score is calculated with a Rasch model, which takes into account item difficulty. It is intended to capture each child's ability level on the tested dimension or skill.

[c]The raw score is the number of items a child answered correctly.

[d]The first sensitivity analysis (sensitivity sample 1) removed a group of 189 children who received the High 5s supplement from the Making Pre-K Count kindergarten analytic sample. Instead of these children being replaced as in the primary analysis, the 183 Making Pre-K Count program group children already included in this analysis who did not receive the High 5s supplement were up-weighted to balance the 189 children who were removed.

[e]The second sensitivity analysis (sensitivity sample 2) also removed the group of 189 children who received the High 5s supplement from the sample. They were replaced by a group of 132 Making Pre-K Count program group children who did not receive the High 5s supplement, and the replacement group was weighted up to be equivalent to the size of the group that was removed.

Compared with pre-K as usual, Making Pre-K Count did not have a statistically significant effect on the analysis using the primary REMA-K T score. When using the logit or raw scores, there are small, statistically significant impacts of Making Pre-K Count on the REMA-K (effect sizes = 0.10 and 0.12). When specifying the Making Pre-K Count program group sample differently, the effect remained small (effect size = 0.09) and was statistically significant in one analysis but not the other.

Results are sensitive to how the REMA-K score is created, with some scores yielding impacts that are statistically significant and some not. Regardless of how the measure is scored, the effect of Making Pre-K Count on children's kindergarten math skills seems to be modest, with effect sizes ranging from 0.08 to 0.12. These effects are smaller than that seen on the REMA in a previous study of Building Blocks in a similar context, which found an effect size of 0.19.[4]

[4]Clements et al., "Effects on Mathematics" (2016). The study used the TEAM, a version of the REMA designed for pre-K through fifth grade.

Appendix F

REMA-K Subscales

While many pre-K curricula focus primarily on numeracy, Building Blocks is unique in its additional focus on geometry (and, to a lesser extent, measurement) in the pre-K year. Appendix F examines the effects of Making Pre-K Count on children's responses to items on the Research-Based Math Assessment–Kindergarten (REMA-K) in the following specific content areas: numbering, patterning, measurement, and geometry. As shown in Appendix D, while the REMA-K and the Woodcock-Johnson Applied Problems have similar numbers of numeracy items, the REMA-K has substantially more items devoted to other math content areas: geometry (nine REMA-K items versus three Woodcock-Johnson items), patterning (eight versus zero), and measurement (four versus one).

Due to the nature of the REMA-K, it is not possible to separate an item's difficulty from its type of content. The REMA-K presents items in an order of increasing difficulty, regardless of content area. For example, if geometry questions are more difficult than other questions, children may not get far enough in the test to receive many geometry items and receive a high score on the geometry subscale. This is true even if the easier items they were answering incorrectly were actually numeracy items. Therefore, a low geometry score is likely to indicate some combination of lower geometry skills and lower math skills overall.[1]

Making Pre-K Count had positive, statistically significant impacts on children's numbering, patterning, and geometry skills, as shown in Appendix Table F.1. As hypothesized, the largest effect was on children's geometry skills, with children in the program group answering on average 0.81 geometry questions correctly and children in the control group answering 0.64 geometry questions correctly (effect size = 0.19).[2]

---

[1]All children received at least one numeracy item and one patterning item.

[2]While it is impossible to parse apart the effect of Making Pre-K Count on getting far enough in the test to receive a geometry question and its effect on geometry skills, additional analyses (not shown) indicate that at least some of the effect on geometry was due to a child getting far enough to receive a geometry question (effect size = 0.11).

Appendix Table F.1

Impacts of Making Pre-K Count in the Spring of the Kindergarten Year, REMA-K Math Subscales

Math Subscale (Raw Scores)	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]
Numbering	6.29	5.91	0.37	0.10 *	0.09
Patterning	1.53	1.43	0.10	0.03 **	0.13
Measurement	0.40	0.34	0.06	0.28	0.08
Geometry	0.81	0.64	0.17	0.00 ***	0.19
Sample size					
Blocks	16	16			
Sites	35	34			
Children	641	684			

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

The program group received Making Pre-K Count in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

The Research-Based Early Math Assessment–Kindergarten (REMA-K) is an adaptation of the REMA (Clements, Sarama, and Liu, 2008). Item selection represents the full range of early mathematics competencies applicable within the pre-K, kindergarten, and first-grade years.

The REMA-K presents items in an increasingly difficult order, regardless of content area. Therefore, it is not possible to parse apart an item's difficulty from its content area.

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

Appendix G

Subgroup Analyses for Making Pre-K Count and High 5s Programs

Appendix G describes the results from subgroup analyses conducted to understand whether the Making Pre-K Count (MPC) and High 5s programs may have had differential impacts on child outcomes based on a predetermined set of school and child characteristics. While there does seem to be variation in impacts on children's skills in kindergarten by child-level characteristics, there is no consistent group of children that benefited across all outcomes.

Subgroup impacts are presented for three comparisons:

- One year of math enrichment (Making Pre-K Count) compared with no math enrichment (pre-K-as-usual control group)
- Two years of math enrichment (MPC plus High 5s supplement) compared with one year of math enrichment (Making Pre-K Count only)
- Two years of math enrichment (MPC plus High 5s supplement) compared with no math enrichment (pre-K-and-kindergarten-as-usual control group)

Subgroup Impacts of One Year of Math Enrichment Compared with No Math Enrichment

Impacts on children's kindergarten outcomes were estimated comparing children who received Making Pre-K Count only with children who received pre-K as usual for subgroups based on five characteristics:

- venue (community-based organizations versus public schools)
- children's self-regulation skills at the *beginning* of pre-K
- children's receptive language ability at the *beginning* of pre-K
- gender
- Spanish-speaking ability[1]

There were no statistically significant differences in impacts on either measure of children's kindergarten math skills based on pre-K venue (Appendix Table G.1), self-regulation or receptive language skills at the start of pre-K (Appendix Table G.2), gender (Appendix Table G.3), or Spanish-speaking ability (Appendix Table G.4).

[1]Spanish-speaking ability was reported by the parent on the informed consent form at study entry via the following question: "Does your child speak and understand Spanish?"

There were statistically significant differences in impacts on measures of executive function for two subgroups. Making Pre-K Count had statistically significant impacts on inhibitory control for children coming in with low self-regulation skills in pre-K (effect size = 0.31) but not for children coming in with high self-regulation skills (Appendix Table G.2). For boys, Making Pre-K Count had no statistically significant impacts on executive function outcomes, but for girls, positive impacts on working memory were observed, with an effect size of 0.27 standard deviations (Appendix Table G.3). In addition, statistically significant positive impacts (effect size = 0.18) on working memory were observed for children who did not speak Spanish; no impacts were observed for Spanish-speaking children (Appendix Table G.4). The difference in impact was not statistically significant.

Subgroup Impacts of Two Years of Math Enrichment Compared with One Year of Math Enrichment

High 5s took place exclusively in public schools, so subgroup differences by venue were not conducted. Because High 5s randomly assigned children in the fall of kindergarten within Making Pre-K Count program sites, subgroup analyses based on children's skill levels used children's outcomes as assessed at the *end* of pre-K (instead of at the beginning of pre-K, as in other subgroup analyses for the Making Pre-K Count sample). Subgroup analyses based on demographic characteristics (gender and Spanish-speaking ability) draw on data from the *beginning* of pre-K. Subgroup analyses were conducted based on six characteristics:

- children's math skills at the *end* of pre-K
- children's inhibitory control at the *end* of pre-K
- children's receptive language at the *end* of pre-K
- gender
- Spanish-speaking ability
- math instructional time in kindergarten

Across the six subgroups, there were no consistent differential impacts of receipt of High 5s on top of Making Pre-K Count (Appendix Tables G.5 through G.8). As shown in Appendix Table G.5, High 5s on top of Making Pre-K Count appeared to have a statistically significant impact on the global math measure for children who had high inhibitory control skills at kindergarten entry (effect size = 0.17) but not for children who entered kindergarten with low inhibitory control. Additionally, there were statistically significant impacts of two years of math enrichment on the detailed math measure (effect size = 0.19) for children who entered kindergarten with stronger receptive language ability but not for children who entered with weaker receptive language ability. As shown in Appendix Table G.7, there was a statistically significant difference between the impacts on working memory for Spanish-speaking children and for those who did not speak Spanish. As shown in Appendix Table G.8, there were no statistically significant differences in children's math scores across schools with more or less math instructional time. Interestingly, High 5s had a larger impact on children's math attitudes in schools with lower than average time spent on math (effect size = 0.29) than in schools with more time spent on math (effect size = -0.18), suggesting that the more time children spent in High 5s relative to classroom math instruction, the more likely they were to report positive attitudes toward math.

Subgroup Impacts of Two Years of Math Enrichment Compared with No Math Enrichment

To assess whether there were differential impacts of MPC plus the High 5s supplement compared with receiving no math enrichment, subgroup analyses were conducted based on two characteristics:

- gender
- Spanish-speaking ability

There were no statistically significant subgroup differences by gender (Appendix Table G.9) or Spanish-speaking ability (Appendix Table G.10) for the sample of children who received Making Pre-K Count plus High 5s compared with children in public schools who had no math enrichment.

Conclusion

While there does seem to be some variation in impacts on children's skills in kindergarten by child-level characteristics, there is no consistent group of children that benefited across all outcomes. In general, subgroup analyses suggest that the program may have been most beneficial for students entering with the highest skills, but findings are not conclusive.

Appendix Table G.1

Impacts of Making Pre-K Count in the Spring of the Kindergarten Year, by Venue

Outcome Score	Community-Based Organization: Control Group Mean	Community-Based Organization: Difference (Impact)	Community-Based Organization: P-Value	Community-Based Organization: Effect Size[a]	Public School: Control Group Mean	Public School: Difference (Impact)	Public School: P-Value	Public School: Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Math											
REMA-K[b]	38.55	-0.24	0.78	-0.03	37.21	1.18	0.04 **	0.13	-1.42	0.17	
Woodcock-Johnson Applied Problems[c]	105.41	-0.60	0.64	-0.05	103.73	0.26	0.79	0.02	-0.85	0.59	
Math attitudes											
Children's attitudes toward math (1-5)	3.55	0.14	0.38	0.10	3.39	0.17	0.10 *	0.11	-0.03	0.88	
Language											
Receptive vocabulary[c]	97.12	2.16	0.24	0.14	97.58	-0.19	0.89	-0.01	2.35	0.31	
Executive function											
Inhibitory control (0-1)	0.68	0.02	0.46	0.08	0.69	0.01	0.75	0.02	0.01	0.67	
Working memory[d]	2.18	0.35	0.02 **	0.23	2.24	0.07	0.47	0.05	0.27	0.13	
Sample size											
Blocks	5				11						
Sites	11				23						
Children	197				487						

(continued)

Appendix Table G.1 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K with corresponding outcomes for the pre-K-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

Appendix Table G.2

Impacts of Making Pre-K Count in the Spring of the Kindergarten Year, by Skill Level at the Beginning of Pre-K

Outcome Score	Low Skills: Control Group Mean	Low Skills: Difference (Impact)	Low Skills: P-Value	Low Skills: Effect Size[a]	High Skills: Control Group Mean	High Skills: Difference (Impact)	High Skills: P-Value	High Skills: Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Self-regulation skills subgroups[b]											
Math											
REMA-K[c]	36.22	0.86	0.37	0.10	38.59	1.47	0.08 *	0.17	-0.61	0.64	
Woodcock-Johnson Applied Problems[d]	102.60	0.37	0.76	0.03	105.85	-0.10	0.93	-0.01	0.47	0.78	
Math attitudes											
Children's attitudes toward math (1-5)	3.25	0.40	0.04 **	0.26	3.45	0.18	0.33	0.12	0.22	0.40	
Language											
Receptive vocabulary[d]	97.31	-0.67	0.70	-0.04	99.52	1.74	0.27	0.12	-2.41	0.31	
Executive function											
Inhibitory control (0-1)	0.64	0.06	0.02 **	0.31	0.71	-0.01	0.66	-0.05	0.07	0.05	††
Working memory[e]	1.96	0.24	0.15	0.17	2.52	0.21	0.17	0.15	0.03	0.91	
Receptive language skills subgroups[f]											
Math											
REMA-K[c]	34.53	1.15	0.26	0.14	40.46	1.06	0.25	0.13	0.09	0.95	
Woodcock-Johnson Applied Problems[d]	99.88	-0.19	0.88	-0.02	107.86	0.58	0.62	0.05	-0.77	0.66	
Math attitudes											
Children's attitudes toward math (1-5)	3.44	0.31	0.10	0.20	3.35	0.35	0.05 *	0.22	-0.04	0.88	
Language											
Receptive vocabulary[d]	94.07	-2.61	0.10	-0.20	102.96	2.14	0.18	0.14	-4.75	0.04	††

(continued)

Appendix Table G.2 (continued)

Outcome Score	Low Skills: Control Group Mean	Low Skills: Difference (Impact)	Low Skills: P-Value	Low Skills: Effect Size[a]	High Skills: Control Group Mean	High Skills: Difference (Impact)	High Skills: P-Value	High Skills: Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Executive function											
Inhibitory control (0-1)	0.63	0.01	0.60	0.07	0.71	0.04	0.13	0.19	-0.03	0.45	
Working memory[e]	1.86	0.31	0.07 *	0.21	2.59	0.19	0.23	0.14	0.12	0.60	
Sample size											
Blocks											
Self-regulation skills subgroup	16				16						
Receptive language skills subgroup	16				16						
Sites											
Self-regulation skills subgroup[g]	33				34						
Receptive language skills subgroup[h]	32				34						
Children											
Self-regulation skills subgroup	183				190						
Receptive language skills subgroup	197				174						

(continued)

Appendix Table G.2 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K with corresponding outcomes for the pre-K-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]Children's self-regulation skills were measured using the Preschool Self-Regulation Assessment (PSRA; Smith-Donald, Raver, Hayes, and Richardson, 2007), administered at pre-K entry in the fall of 2014. Children scoring below the median constitute the low-skills group; children scoring at or above the median constitute the high-skills group.

[c]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[d]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[e]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

[f]Children's receptive language skills were measured using the ROWPVT, administered at pre-K entry in the fall of 2014. Children scoring below the median constitute the low-skills group; children scoring at or above the median constitute the high-skills group.

[g]At one site in the control group, all assessed children scored at or above the median of the PSRA Attention and Inhibition score; therefore, the sample size of sites for the control group in the "low" subgroup for entering self-regulation skills is 33.

[h]At two sites in the control group, all assessed children scored at or above the median of the total ROWPVT score; therefore, the sample size of sites for the control group in the "low" subgroup for entering language skills is 32.

Appendix Table G.3

Impacts of Making Pre-K Count in the Spring of the Kindergarten Year, by Gender

Outcome Score	Male: Control Group Mean	Male: Difference (Impact)	Male: P-Value	Male: Effect Size[a]	Female: Control Group Mean	Female: Difference (Impact)	Female: P-Value		Female: Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Math												
REMA-K[b]	37.77	0.32	0.67	0.04	37.44	1.13	0.09	*	0.13	-0.81	0.42	
Woodcock-Johnson Applied Problems[c]	104.60	-0.23	0.82	-0.02	103.92	0.36	0.73		0.03	-0.58	0.69	
Math attitudes												
Children's attitudes toward math (1-5)	3.50	0.20	0.10	0.13	3.36	0.13	0.26		0.09	0.07	0.70	
Language												
Receptive vocabulary[c]	96.25	1.07	0.44	0.07	98.84	-0.20	0.87		-0.01	1.26	0.50	
Executive function												
Inhibitory control (0-1)	0.68	0.03	0.15	0.12	0.69	0.00	0.89		-0.01	0.03	0.27	
Working memory[d]	2.34	-0.08	0.57	-0.05	2.09	0.39	0.00	***	0.27	-0.46	0.01	†††
Sample size												
Blocks	16				16							
Sites	34				34							
Children	324				360							

(continued)

Appendix Table G.3 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K with corresponding outcomes for the pre-K-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

Appendix Table G.4

Impacts of Making Pre-K Count in the Spring of the Kindergarten Year, by Spanish-Speaking Ability

Outcome Score	Not Spanish-Speaking[a] Control Group Mean	Not Spanish-Speaking[a] Difference (Impact)	Not Spanish-Speaking[a] P-Value		Not Spanish-Speaking[a] Effect Size[b]	Spanish-Speaking[a] Control Group Mean	Spanish-Speaking[a] Difference (Impact)	Spanish-Speaking[a] P-Value	Spanish-Speaking[a] Effect Size[b]	Difference Between Subgroups	P-Value	Sig.
Math												
REMA-K[c]	38.19	1.03	0.14		0.12	36.03	0.44	0.54	0.05	0.59	0.55	
Woodcock-Johnson Applied Problems[d]	105.41	0.58	0.59		0.05	101.91	-0.30	0.76	-0.03	0.88	0.55	
Math attitudes												
Children's attitudes toward math (1-5)	3.36	0.20	0.09	*	0.13	3.52	0.15	0.26	0.10	0.05	0.76	
Language												
Receptive vocabulary[d]	99.95	1.34	0.33		0.09	94.47	0.64	0.66	0.04	0.70	0.72	
Executive function												
Inhibitory control (0-1)	0.71	0.03	0.17		0.11	0.64	0.00	0.99	0.00	0.03	0.34	
Working memory[e]	2.27	0.25	0.02	**	0.18	2.08	0.01	0.94	0.01	0.24	0.16	
Sample size												
Blocks	16					16						
Sites[f]	34					32						
Children	364					307						

(continued)

Appendix Table G.4 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K with corresponding outcomes for the pre-K-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Children's language was reported by their parents on the informed consent form at study entry via the question, "Does your child speak and understand Spanish?"

[b]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[d]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[e]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

[f]At two sites in the control group, all assessed children did not speak Spanish; therefore, the sample size of sites for the control group in the Spanish-speaking subgroup is 32 instead of 34.

Appendix Table G.5

Impacts of the High 5s Supplement in the Spring of the Kindergarten Year, by Skill Level at the End of Pre-K

	Low Skills				High Skills						
Outcome Score	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Math skills subgroups[b]											
Math											
REMA-K[c]	34.17	0.85	0.43	0.12	42.55	1.14	0.15	0.20	-0.28	0.83	
Woodcock-Johnson Applied Problems[d]	98.40	0.38	0.80	0.03	107.62	2.04	0.14	0.17	-1.66	0.41	
Math attitudes											
Children's attitudes toward math (1-5)	3.73	-0.35	0.18	-0.24	3.47	0.14	0.57	0.09	-0.49	0.17	
Language											
Receptive vocabulary[d]	94.32	-1.59	0.49	-0.11	103.30	0.18	0.93	0.01	-1.77	0.57	
Executive function											
Inhibitory control (0-1)	0.62	-0.02	0.57	-0.10	0.73	0.01	0.82	0.04	-0.03	0.58	
Working memory[e]	2.03	-0.07	0.74	-0.05	2.71	-0.01	0.95	-0.01	-0.06	0.83	
Inhibitory skills subgroups[f]											
Math											
REMA-K[c]	35.06	1.26	0.37	0.15	41.84	0.63	0.34	0.11	0.63	0.68	
Woodcock-Johnson Applied Problems[d]	99.63	0.67	0.72	0.05	106.99	1.98	0.09 *	0.17	-1.31	0.55	
Math attitudes											
Children's attitudes toward math (1-5)	3.51	0.03	0.92	0.02	3.55	0.07	0.75	0.04	-0.04	0.92	
Language											
Receptive vocabulary[d]	97.49	-3.44	0.26	-0.24	102.51	0.24	0.88	0.02	-3.68	0.29	

(continued)

Appendix Table G.5 (continued)

	Low Skills				High Skills						
Outcome Score	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Executive function											
Inhibitory control (0-1)	0.63	0.01	0.82	0.05	0.71	0.00	0.97	0.00	0.01	0.88	
Working memory[e]	2.16	-0.17	0.49	-0.14	2.62	0.00	0.98	0.00	-0.16	0.58	
<u>Receptive language skills subgroups</u>[g]											
Math											
REMA-K[c]	36.18	0.60	0.57	0.08	40.24	1.36	0.06 *	0.19	-0.75	0.56	
Woodcock-Johnson Applied Problems[d]	99.52	0.79	0.62	0.06	106.07	2.01	0.11	0.17	-1.22	0.55	
Math attitudes											
Children's attitudes toward math (1-5)	3.49	-0.10	0.69	-0.07	3.53	0.14	0.54	0.09	-0.24	0.48	
Language											
Receptive vocabulary[d]	92.10	0.34	0.89	0.03	105.09	0.12	0.95	0.01	0.23	0.94	
Executive function											
Inhibitory control (0-1)	0.66	-0.04	0.35	-0.17	0.69	0.02	0.64	0.07	-0.05	0.32	
Working memory[e]	1.91	0.31	0.15	0.25	2.78	-0.20	0.22	-0.16	0.50	0.06	†
Sample size											
Sites											
Math skills subgroup	23				22						
Inhibitory skills subgroup	22				24						
Receptive language skills subgroup	23				22						
Children											
Math skills subgroup	78				85						
Inhibitory skills subgroup	57				99						
Receptive language skills subgroup	78				85						

(continued)

Appendix Table G.5 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group received only MPC in pre-K and participated in kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the group assigned to MPC in pre-K and kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for pre-K sites.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]Children's math skills were measured using the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) math assessment (Najarian et al., 2010), administered at the end of pre-K in spring 2015. Children scoring below the median constitute the low-skills group; children scoring at or above the median constitute the high-skills group.

[c]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[d]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[e]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

[f]Children's inhibitory skills were measured using the Pencil Tap task (Luria, 1966; Diamond and Taylor, 1996), which requires children to tap once immediately after the experimenter taps twice and vice versa. The task was designed to measure aspects of executive control including working memory, inhibitory control, and attention, and was administered at the end of pre-K in spring 2015. Children scoring below the median constitute the low-skills group; children scoring at or above the median constitute the high-skills group.

[g]Children's language skills were measured using the Receptive One-Word Picture Vocabulary Test (ROWPVT; Martin and Brownell, 2011), administered at the end of pre-K in spring 2015. Children scoring below the median constitute the low-skills group; children scoring at or above the median constitute the high-skills group.

Appendix Table G.6

Impacts of the High 5s Supplement in the Spring of the Kindergarten Year, by Gender

	Male				Female						
Outcome Score	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Math											
REMA-K[b]	38.17	1.61	0.09 *	0.17	37.95	1.56	0.06 *	0.22	0.05	0.97	
Woodcock-Johnson Applied Problems[c]	103.61	0.86	0.55	0.07	102.67	1.62	0.17	0.14	-0.76	0.69	
Math attitudes											
Children's attitudes toward math (1-5)	3.54	0.01	0.98	0.00	3.32	0.14	0.41	0.10	-0.13	0.60	
Language											
Receptive vocabulary[c]	95.53	1.33	0.42	0.09	98.11	0.58	0.74	0.04	0.75	0.76	
Executive function											
Inhibitory control (0-1)	0.69	-0.03	0.32	-0.12	0.68	0.02	0.50	0.08	-0.05	0.23	
Working memory[d]	2.23	0.01	0.96	0.01	2.22	0.13	0.36	0.10	-0.13	0.57	
Sample size											
Sites	24				23						
Children	157				153						

(continued)

Appendix Table G.6 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group received only MPC in pre-K and participated in kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the group assigned to MPC in pre-K and kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for pre-K sites.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

Appendix Table G.7

Impacts of the High 5s Supplement in the Spring of the Kindergarten Year, by Spanish-Speaking Ability

	Not Spanish-Speaking[a]				Spanish-Speaking[a]						
Outcome Score	Control Group Mean	Difference (Impact)	P-Value	Effect Size[b]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[b]	Difference Between Subgroups	P-Value	Sig.
Math											
REMA-K[c]	38.50	1.51	0.06 *	0.18	36.19	2.01	0.07 *	0.26	-0.50	0.71	
Woodcock-Johnson Applied Problems[d]	103.49	1.32	0.27	0.11	101.04	0.72	0.63	0.06	0.60	0.76	
Math attitudes											
Children's attitudes toward math (1-5)	3.41	0.17	0.28	0.11	3.50	-0.06	0.77	-0.04	0.23	0.37	
Language											
Receptive vocabulary[d]	98.85	1.90	0.22	0.13	93.24	-0.22	0.90	-0.02	2.12	0.38	
Executive function											
Inhibitory control (0-1)	0.73	-0.04	0.06 *	-0.19	0.60	0.02	0.45	0.11	-0.07	0.08	†
Working memory[e]	2.44	-0.20	0.15	-0.15	1.95	0.50	0.01 ***	0.37	-0.70	0.00	†††
Sample size											
Sites	23				19						
Children	189				117						

(continued)

Appendix Table G.7 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group received only MPC in pre-K and participated in kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the group assigned to MPC in pre-K and kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for pre-K sites.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Children's language was reported by their parents on the informed consent form at study entry via the question, "Does your child speak and understand Spanish?"

[b]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[d]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[e]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

Appendix Table G.8

Impacts of the High 5s Supplement in the Spring of the Kindergarten Year, by Math Instructional Time

	Low Kindergarten Math Instructional Time[a]				High Kindergarten Math Instructional Time[b]						
Outcome Score	Control Group Mean	Difference (Impact)	P-Value	Effect Size[c]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[c]	Difference Between Subgroups	P-Value	Sig.
Math											
REMA-K[d]	39.88	0.82	0.36	0.10	36.15	2.33	0.01 ***	0.29	-1.51	0.23	
Woodcock-Johnson Applied Problems[e]	103.68	1.92	0.17	0.14	102.33	0.70	0.56	0.06	1.22	0.51	
Math attitudes											
Children's attitudes toward math (1-5)	3.28	0.44	0.01 **	0.29	3.59	-0.25	0.13	-0.18	0.69	0.00	†††
Language											
Receptive vocabulary[e]	98.71	2.09	0.24	0.13	95.78	-0.23	0.88	-0.02	2.32	0.33	
Executive function											
Inhibitory control (0-1)	0.73	-0.03	0.26	-0.13	0.65	0.02	0.50	0.08	-0.05	0.20	
Working memory[f]	2.41	0.08	0.63	0.05	2.04	0.09	0.52	0.07	-0.02	0.94	
Sample size											
Sites	11				13						
Children	142				168						

(continued)

Appendix Table G.8 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group received only MPC in pre-K and participated in kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the group assigned to MPC in pre-K and kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for pre-K sites.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Children in schools where the average duration of math instructional time recorded during observations of kindergarten classrooms was less than 53 minutes constitute the low kindergarten math instructional time group.

[b]Children in schools where the average duration of math instructional time recorded during observations of kindergarten classrooms was greater than or equal to 53 minutes constitute the high kindergarten math instructional time group.

[c]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[d]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[e]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[f]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

Appendix Table G.9

Impacts of MPC Plus High 5s Supplement in the Spring of the Kindergarten Year, by Gender

	Male				Female						
Outcome Score	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[a]	Difference Between Subgroups	P-Value	Sig.
Math											
REMA-K[b]	36.45	3.03	0.01 ***	0.34	36.36	2.91	0.01 **	0.34	0.12	0.94	
Woodcock-Johnson Applied Problems[c]	102.11	2.07	0.26	0.17	102.96	1.03	0.54	0.09	1.04	0.68	
Math attitudes											
Children's attitudes toward math (1-5)	3.34	0.19	0.33	0.12	3.20	0.26	0.14	0.16	-0.07	0.79	
Language											
Receptive vocabulary[c]	92.66	3.86	0.07 *	0.24	98.51	0.16	0.93	0.01	3.71	0.19	
Executive function											
Inhibitory control (0-1)	0.66	0.00	0.98	0.00	0.66	0.03	0.36	0.13	-0.03	0.48	
Working memory[d]	2.28	-0.08	0.74	-0.05	1.98	0.35	0.07 *	0.25	-0.43	0.16	
Sample size											
Blocks	10				10						
Sites	22				22						
Children	169				176						

(continued)

Appendix Table G.9 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group did not receive math enrichment and participated in pre-K and kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the pre-K-and-kindergarten-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

Appendix Table G.10

Impacts of MPC Plus High 5s Supplement in the Spring of the Kindergarten Year, by Spanish-Speaking Ability

	Not Spanish-Speaking[a]				Spanish-Speaking[a]						
Outcome Score	Control Group Mean	Difference (Impact)	P-Value	Effect Size[b]	Control Group Mean	Difference (Impact)	P-Value	Effect Size[b]	Difference Between Subgroups	P-Value	Sig.
Math											
REMA-K[c]	36.41	3.32	0.01 ***	0.38	35.39	2.11	0.06 *	0.24	1.21	0.47	
Woodcock-Johnson Applied Problems[d]	102.90	1.44	0.43	0.12	101.86	0.27	0.86	0.02	1.17	0.63	
Math attitudes											
Children's attitudes toward math (1-5)	3.15	0.36	0.05 **	0.22	3.46	0.11	0.58	0.08	0.25	0.36	
Language											
Receptive vocabulary[d]	97.87	2.48	0.28	0.16	92.69	0.91	0.67	0.06	1.57	0.61	
Executive function											
Inhibitory control (0-1)	0.67	0.02	0.54	0.08	0.63	0.02	0.54	0.08	0.00	0.99	
Working memory[e]	2.14	0.03	0.87	0.03	2.12	0.21	0.28	0.14	-0.17	0.55	
Sample size											
Blocks	10				10						
Sites	22				20						
Children	188				154						

(continued)

Appendix Table G.10 (continued)

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Statistically significant differences in impact estimates across different subgroups are indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group did not receive math enrichment and participated in pre-K and kindergarten as usual.

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the pre-K-and-kindergarten-as-usual control group, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Children's language was reported by their parents on the informed consent form at study entry via the question, "Does your child speak and understand Spanish?"

[b]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[d]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[e]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

Appendix H

Impacts of Making Pre-K Count Plus High 5s Supplement Compared with Making Pre-K Count Only, Original Sample

As described in Chapter 2, the High 5s study was embedded within the larger Making Pre-K Count design. Within the 24 public schools that were assigned to receive Making Pre-K Count in the original Making Pre-K Count design, students whose parents gave consent and who remained in the same school between the spring of pre-K and the fall of kindergarten were randomly assigned (within the school) early in fall 2015 to either High 5s or a kindergarten-as-usual control condition. A total of 655 students were randomly assigned using a random number generator. In each school, the total number of High 5s clubs was determined and children with the lowest randomly generated number were assigned to the High 5s supplement until all the clubs were filled. The remaining children constituted the control group for the High 5s comparison: children who only received Making Pre-K Count in pre-K. A total of 320 students were assigned to the MPC plus High 5s supplement group, and 335 were assigned to the Making Pre-K Count only group.

The *analytic* sample differs slightly from that original sample of 655 children. Children in the primary analysis are included in the *analytic* sample only if they stayed in the same school from pre-K to kindergarten. In addition, the analytic sample includes 18 students whose parents did not give consent for High 5s participation (nonconsenters), who were randomized in a second phase to maintain the internal validity of the comparison with a control group of children who did not receive Making Pre-K Count in pre-K. The analysis presented in Appendix Table H.1 compares kindergarten outcomes for the *original* sample of 655 children (including children who did not stay in the same school) and does not include the nonconsenters. Five percent of the original 655 children were not available for assessment at follow-up.

As shown in Appendix Table H.1, the patterns of impact findings across the analytic and original samples are similar. The impact of two years of math enrichment on the detailed math measure is positive and statistically significant. No statistically significant impacts are found for the other outcomes.

Appendix Table H.1

Impacts of the High 5s Supplement in the Spring of the Kindergarten Year, Original Sample

Outcome Score	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Effect Size[a]
Math						
REMA-K[b]	39.55	37.83	1.72	0.01	***	0.20
Woodcock-Johnson Applied Problems[c]	104.28	103.12	1.16	0.20		0.09
Math attitudes						
Children's attitudes toward math (1-5)	3.51	3.49	0.03	0.81		0.02
Language						
Receptive vocabulary[c]	98.15	97.42	0.73	0.53		0.05
Executive function						
Inhibitory control (0-1)	0.68	0.68	-0.01	0.78		-0.02
Working memory[d]	2.27	2.23	0.05	0.68		0.03
Sample size						
Sites	24	24				
Children	307	315				

SOURCE: MDRC calculations based on the direct child assessments administered in spring 2016.

NOTES: Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

The program group received Making Pre-K Count (MPC) in pre-K and High 5s in kindergarten. The control group received only Making Pre-K Count in pre-K and participated in kindergarten as usual. The original sample consists of MPC program public school children who gave positive consent to participate in High 5s (including children who did not stay in the same school) and were assessed in the spring of their kindergarten year (spring 2016).

Impacts were estimated by comparing kindergarten outcomes for the group assigned to MPC in pre-K and High 5s in kindergarten with corresponding outcomes for the group assigned to MPC in pre-K and kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for pre-K sites.

Outcomes were measured by T scores from the Research-Based Early Math Assessment–Kindergarten (REMA-K; Clements, Sarama, and Liu, 2008); the Woodcock-Johnson Applied Problems subscale of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, and Mather, 2001); an MDRC-created assessment measuring children's attitudes toward math and school; the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011); the Hearts and Flowers (Wright and Diamond, 2014) computerized task; and the Corsi Blocks task (Corsi, 1972; Lezak, 1983).

Rounding may cause slight discrepancies in sums and differences.

[a]Effect size is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]The REMA-K has a mean score of 50 and a standard deviation of 10. The norm is based on a group of children between pre-K and third grade.

[c]This is a standardized measure with a mean score of 100 and a standard deviation of 15.

[d]The score reports the highest number of blocks the child was able to tap in correct order in two attempts.

References

Anthony, Karen, Dale C. Farran, and Kerry G. Hofer. 2013. "Improving Young Children's Math Learning Through Technology." Unpublished paper, PDF.

Aunola, Kaisa, Esko Leskinen, Marja-Kristiina Lerkkanen, and Jari-Erik Nurmi. 2004. "Developmental Dynamics of Math Performance from Preschool to Grade 2." *Journal of Educational Psychology* 96, 4: 699-713.

Blair, Clancy, and Rachel Peters Razza. 2007. "Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten." *Child Development* 78, 2: 647-663.

Bull, Rebecca, Kimberly Andrews Espy, and Sandra A. Wiebe. 2008. "Short-Term Memory, Working Memory, and Executive Functioning in Preschoolers: Longitudinal Predictors of Mathematical Achievement at Age 7 Years." *Developmental Neuropsychology* 33, 3: 205-228.

Clements, Douglas H., and Julie Sarama. 2004. "Learning Trajectories in Mathematics Education." *Mathematical Thinking and Learning* 6, 2: 81-89.

Clements, Douglas H., and Julie Sarama. 2007. "Effects of a Preschool Mathematics Curriculum: Summative Research on the Building Blocks Project." *Journal for Research in Mathematics Education* 38, 2: 136-163.

Clements, Douglas H., and Julie Sarama. 2008. "Experimental Evaluation of the Effects of a Research-Based Preschool Mathematics Curriculum." *American Educational Research Journal* 45, 2: 443-493.

Clements, Douglas H., and Julie Sarama. 2013. *Building Blocks: Teacher's Edition.* Columbus, OH: McGraw-Hill.

Clements, Douglas H., and Julie Sarama. 2014. *Learning and Teaching Early Math: The Learning Trajectories Approach*. New York: Routledge.

Clements, Douglas H., Julie Sarama, Carolyn Layzer, Fatih Unlu, Carrie Germeroth, and Lily Fesler. 2016. "Effects on Mathematics and Executive Function Learning of an Early Mathematics Curriculum Synthesized with Scaffolded Play Designed to Promote Self-Regulation Versus the Mathematics Curriculum Alone." Unpublished paper, PDF.

Clements, Douglas H., Julie Sarama, Carolyn Layzer, Fatih Unlu, Christopher B. Wolfe, Mary Elaine Spitler, and Daniel Weiss. 2016. "Effects of TRIAD on Mathematics Achievement: Long-Term Impacts." Paper presented at the Society for Research on Educational Effectiveness Spring Conference, Washington, DC, March 2-5.

Clements, Douglas H., Julie Sarama, and Xiufeng H. Liu. 2008. "Development of a Measure of Early Mathematics Achievement Using the Rasch Model: The Research-Based Early Maths Assessment." *Educational Psychology* 28, 4: 457-482.

Clements, Douglas H., Julie Sarama, Mary Elaine Spitler, Alissa A. Lange, and Christopher B. Wolfe. 2011. "Mathematics Learned by Young Children in an Intervention Based on Learning Trajectories: A Large-Scale Cluster Randomized Trial." *Journal for Research in Mathematics Education* 42, 2: 127-166.

Corsi, Philip Michael. 1972. "Human Memory and the Medial Temporal Region of the Brain." PhD dissertation. Montreal: McGill University.

Diamond, Adele, and Colleen Taylor. 1996. "Development of an Aspect of Executive Control: Development of the Abilities to Remember What I Said and to 'Do as I Say, Not as I Do.'" *Developmental Psychobiology* 29, 4: 315-334.

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, and Jeanne Brooks-Gunn. 2007. "School Readiness and Later Achievement." *Developmental Psychology* 43, 6: 1428-1446.

Duncan, Greg J., and Katherine Magnuson. 2009. "The Nature and Impact of Early Skills, Attention, and Behavior." Paper presented at the Russell Sage Foundation Conference on Social Inequality and Educational Outcomes, New York City.

Duncan, Sharon E., and Edward A. De Avila. 1998. *preLAS 2000*. Monterey, CA: CTB/ McGraw-Hill.

Feller, Avi, Todd Grindal, Luke W. Miratrix, and Lindsay C. Page. 2016. "Compared to What? Variation in the Impacts of Early Childhood Education by Alternative Care Type." *Annals of Applied Statistics* 10, 3: 1245-1285.

Fryer, Roland G. 2011. *Injecting Successful Charter School Strategies into Traditional Public Schools: Early Results from an Experiment in Houston*. Cambridge, MA: National Bureau of Economic Research.

Gathercole, Susan E., Susan J. Pickering, Camilla Knight, and Zoe Stegmann. 2004. "Working Memory Skills and Educational Attainment: Evidence from National Curriculum Assessments at 7 and 14 Years of Age." *Applied Cognitive Psychology* 18: 1-16.

Ginsburg, Herbert P., Joon Sun Lee, and Judi Stevenson Boyd. 2008. "Mathematics Education for Young Children: What It Is and How to Promote It." *Social Policy Report* 22, 1. Ann Arbor, MI: Society for Research in Child Development.

Hofer, Kerry G., Mark W. Lipsey, Nianbo Dong, and Dale C. Farran. 2013. "Results of the Early Math Project: Scale-Up Cross-Site Results." Working paper. Nashville: Peabody Research Institute, Vanderbilt University.

Jacob, Robin, Anna Erickson, and Shira K. Mattera. 2018. *Launching Kindergarten Math Clubs: The Implementation of High 5s in New York City*. New York: MDRC.

Klein, Alice, Prentice Starkey, Douglas H. Clements, Julie Sarama, and Roopa Iyer. 2008. "Effects of a Pre-Kindergarten Mathematics Intervention: A Randomized Experiment." *Journal of Research on Educational Effectiveness* 1, 3: 155-178.

Lee, Valerie E., and Susanna Loeb. 1995. "Where Do Head Start Attendees End Up? One Reason Why Preschool Effects Fade Out." *Educational Evaluation and Policy Analysis* 17, 1: 62-82.

Lezak, Muriel Deutsch. 1983. *Neuropsychological Assessment*. New York: Oxford University Press.

Ludwig, Jens. 2011. "On What Should We Focus Our Early Interventions to Maximize Benefit-Cost Ratios?" Paper presented at Robin Hood Early Childhood Institute planning meeting, New York City.

Luria, Aleksandr Romanovich. 1966. *Higher Cortical Functions in Man*. New York: Basic Books.

Martin, Nancy A., and Rick Brownell. 2011. *Receptive One-Word Picture Vocabulary Test*, 4th ed. Novato, CA: Academic Therapy Publications.

Mattera, Shira, and Pamela Morris. 2017. *Counting on Early Math Skills: Preliminary Kindergarten Impacts of the Making Pre-K Count and High 5s Programs*. New York: MDRC.

McClelland, Megan M., Alan C. Acock, Andrea Piccinin, Sally Ann Rhea, and Michael C. Stallings. 2013. "Relations Between Preschool Attention Span-Persistence and Age 25 Educational Outcomes." *Early Childhood Research Quarterly* 28, 2: 314-324.

McCoy, Dana Charles, Pamela A. Morris, Maia C. Connors, Celia J. Gomez, and Hirokazu Yoshikawa. 2016. "Differential Effectiveness of Head Start in Urban and Rural Communities." *Journal of Applied Developmental Psychology* 43: 29-42.

Moiduddin, Emily, Nikki Aikens, Louisa Tarullo, Jerry West, and Yange Xue. 2012. *Child Outcomes and Classroom Quality in FACES 2009*. OPRE Report 2012-37a. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Morris, Pamela A., Shira K. Mattera, and Michelle F. Maier. 2016. *Making Pre-K Count: Improving Math Instruction in New York City*. New York: MDRC.

Najarian, Michelle, Kyle Snow, Jean Lennon, Susan Kinsey, and Gail Mulligan. 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.

Peisner-Feinberg, Ellen S., Doré R. LaForett, Jennifer M. Schaaf, Lisa M. Hildebrandt, John Sideris, and Yi Pan. 2014. *Children's Outcomes and Program Quality in the North Carolina Pre-Kindergarten Program: 2012-2013 Statewide Evaluation.* Chapel Hill: Frank Porter Graham Child Development Institute, University of North Carolina.

Reardon, Sean F., and Ximena A. Portilla. 2016. "Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry." *AERA Open* 2, 3.

Sarama, Julie, Douglas H. Clements, Prentice Starkey, Alice Klein, and Ann Wakeley. 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.

Sarama, Julie, Douglas H. Clements, Christopher B. Wolfe, and Mary Elaine Spitler. 2012. "Longitudinal Evaluation of a Scale-Up Model for Teaching Mathematics with Trajectories and Technologies." *Journal of Research on Educational Effectiveness* 5, 2: 105-135.

Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores. 2005. *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*. Ypsilanti, MI: High/Scope Educational Research Foundation.

Slavin, Robert E., Cynthia Lake, Susan Davis, and Nancy A. Madden. 2010. *Identifying What Works for Struggling Readers: Educator's Guide*. Best Evidence Encyclopedia. Center for Data-Driven Reform in Education, Johns Hopkins University. Website: www.bestevidence.org.

Smith, Thomas M., Paul Cobb, Dale C. Farran, David S. Cordray, and Charles Munter. 2013. "Evaluating Math Recovery: Assessing the Causal Impact of a Diagnostic Tutoring Program on Student Achievement." *American Educational Research Journal* 50, 2: 397-428.

Smith-Donald, Radiah, Cybele Raver, Tiffany Hayes, and Breeze Richardson. 2007. "Preliminary Construct and Concurrent Validity of the Preschool Self-Regulation Assessment (PSRA) for Field-Based Research." *Early Childhood Research Quarterly* 22, 2: 173-187.

Weikart, David P., Dennis J. Deloria, Sarah A. Lawser, and Ronald Wiegerink. 1970. *Longitudinal Results of the Ypsilanti Perry Preschool Project.* Ypsilanti, MI: High/Scope Educational Research Foundation.

Weiland, Christina, and Hirokazu Yoshikawa. 2013. "Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills." *Child Development* 84, 6: 2112-2130.

Willoughby, Michael T., R. J. Wirth, Clancy B. Blair, and Family Life Project Investigators. 2012. "Executive Function in Early Childhood: Longitudinal Measurement Invariance and Developmental Change." *Psychological Assessment* 24, 2: 418.

Woodcock, Richard W., Kevin S. McGrew, and Nancy Mather. 2001. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.

Wright, Andy, and Adele Diamond. 2014. "An Effect of Inhibitory Load in Children While Keeping Working Memory Load Constant." *Frontiers in Psychology* 5, Article 213.

Other MDRC Publications on Making Pre-K Count and High 5s

Launching Kindergarten Math Clubs
The Implementation of High 5s in New York City
2018. Robin Jacob, Anna Erickson, Shira K. Mattera

Counting on Early Math Skills
Preliminary Kindergarten Impacts of the Making Pre-K Count and High 5s Programs
2017. Shira Mattera, Pamela Morris

Making Pre-K Count
Improving Math Instruction in New York City
2016. Pamela A. Morris, Shira K. Mattera, Michelle F. Maier

NOTE: All the publications listed above are available for free download at www.mdrc.org.

About MDRC

MDRC is a nonprofit, nonpartisan social and education policy research organization dedicated to learning what works to improve the well-being of low-income people. Through its research and the active communication of its findings, MDRC seeks to enhance the effectiveness of social and education policies and programs.

Founded in 1974 and located in New York; Oakland, California; Washington, DC; and Los Angeles, MDRC is best known for mounting rigorous, large-scale, real-world tests of new and existing policies and programs. Its projects are a mix of demonstrations (field tests of promising new program approaches) and evaluations of ongoing government and community initiatives. MDRC's staff members bring an unusual combination of research and organizational experience to their work, providing expertise on the latest in qualitative and quantitative methods and on program design, development, implementation, and management. MDRC seeks to learn not just whether a program is effective but also how and why the program's effects occur. In addition, it tries to place each project's findings in the broader context of related research — in order to build knowledge about what works across the social and education policy fields. MDRC's findings, lessons, and best practices are shared with a broad audience in the policy and practitioner community as well as with the general public and the media.

Over the years, MDRC has brought its unique approach to an ever-growing range of policy areas and target populations. Once known primarily for evaluations of state welfare-to-work programs, today MDRC is also studying public school reforms, employment programs for ex-prisoners, and programs to help low-income students succeed in college. MDRC's projects are organized into five areas:

- Promoting Family Well-Being and Children's Development
- Improving Public Education
- Raising Academic Achievement and Persistence in College
- Supporting Low-Wage Workers and Communities
- Overcoming Barriers to Employment

Working in almost every state, all of the nation's largest cities, and Canada and the United Kingdom, MDRC conducts its projects in partnership with national, state, and local governments, public school systems, community organizations, and numerous private philanthropies.