LONG-TERM EFFECTS OF ENHANCED EARLY CHILDHOOD MATH INSTRUCTION

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

MDRC
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

Shira K. Mattera
Robin Jacob
Cullen MacDowell
Pamela A. Morris

December 2021

Long-Term Effects of Enhanced Early Childhood Math Instruction

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

Shira K. Mattera
Robin Jacob
Cullen MacDowell
Pamela A. Morris

December 2021

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY

Funder

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 organizations and individuals 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, Arnold Ventures, 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, and Sandler 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

Studies have shown that math skills in early childhood are uniquely and strongly predictive of later outcomes across a range of domains and well into adulthood, including the likelihood of graduating from high school and college completion. The Making Pre-K Count and High 5s studies were designed to rigorously test the short- and long-term effects of improving children's math experiences in prekindergarten (pre-K) and kindergarten.

Making Pre-K Count provided pre-K teachers in New York City with a high-quality, evidenced-based math curriculum (Building Blocks) and ongoing teacher training and coaching. The Making Pre-K Count study compared students who were exposed to this curriculum with their peers in pre-K as usual in public school and community-based sites. The High 5s math program was developed to offer children who had received Making Pre-K Count in pre-K in public schools hands-on, supplemental math enrichment in small groups, or clubs, outside of regular instructional time in kindergarten. The High 5s study compared students assigned to Making Pre-K Count in pre-K and High 5s in kindergarten with children assigned to Making Pre-K Count in pre-K and kindergarten as usual. The studies also compared two years of math enrichment with no math enrichment.

The studies used random assignment and tracked children through third grade to test the effects of these math enrichment programs. The confirmatory outcome examined was children's third-grade math scores.

KEY FINDINGS

- **Making Pre-K Count:** Though not statistically significant, Making Pre-K Count had small, positive, longer-term impacts on children's third-grade math test scores, compared with pre-K as usual in public school and community-based sites.
- **High 5s:** The impact of High 5s on children's third-grade math test scores in public schools, over and above the effect of Making Pre-K Count alone, was close to zero and not statistically significant.
- **Making Pre-K Count plus High 5s:** Making Pre-K Count and High 5s together had moderate, statistically significant impacts on children's math test scores, compared with pre-K and kindergarten as usual in public schools.

The study team also explored the impact of these two math interventions on children's third-grade literacy test scores, chronic absenteeism, retention in a grade, and placement in special education. These exploratory analyses suggest that Making Pre-K Count alone and the two years of math enrichment together reduced chronic absenteeism and improved children's literacy test scores, though findings were not always statistically significant for literacy test scores.

Taken together, the Making Pre-K Count and High 5s studies present new evidence about the long-term effects of early math interventions on children's later outcomes. Early math enrichment experiences can lead to lasting gains for children across a variety of outcome domains, even years later. The findings suggest that high-quality early math instructional practices could make a difference, particularly for children with the greatest need.

Contents

Exhibits

TABLE

FIGURE

Acknowledgments

The Making Pre-K Count and High 5s studies were a 10-year endeavor, supported by many contributors and collaborators over the years. First and foremost, we thank the children, families, teachers, and administrators who gave so generously of their time and cooperation, and without whom the study would not have been possible.

High 5s and Making Pre-K Count benefited greatly from the support of the coaches, facilitators, and supervisors at Bank Street College, and particularly our ongoing partnership with Katherine Baldwin. 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 have continued to work with us throughout the study. We also appreciate the dedication of and thoughtful consultation with the developers of Building Blocks, Drs. Doug Clements and Julie Sarama, as well as Anna Erickson and Kristi Hanby at the University of Michigan, who helped conceptualize the High 5s intervention and training.

The execution of these studies was made possible by our research partners at RTI International, including Jean Lennon and Jennifer Keeney and the many dedicated data collectors, and by James Kemple at the Research Alliance for New York City Schools. We also extend our appreciation to the steering committees for Making Pre-K Count and High 5s and our academic partners Greg Duncan of the University of California–Irvine and Dale Farran of Vanderbilt University.

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 JoAnn Hsueh, Sharon Huang, John Hutchins, Ginger Knox, and Marie-Andrée Somers for their guidance and review of the report. We thank the operations and technical assistance team for supporting implementation of the programs. We thank the data team that made the many analyses possible, particularly Desiree Alderson and Lyndsay McDonough for their work on data management, as well as Sam Xia, Kelly Saunders, Mirjana Pralica, and Rubin Danberg-Biggs for their assistance with programming. Thanks to Harrison Taylor and Jane Choi for providing research assistance and coordinating the production of the report; Christopher Boland, who edited the report; and Carolyn Thomas, who prepared it for publication.

The Authors

Executive Summary

Studies have found that math skills in early childhood are uniquely and strongly associated with outcomes later in life. Strong early math skills are correlated with not only later math achievement, but also with better reading skills and executive functioning.[1] Further, studies have shown that early math competencies predict outcomes well into adulthood, including the likelihood of graduating from high school and college completion.[2] The Making Pre-K Count and High 5s studies were designed to test the impact of early math enrichment interventions on children's short- and longer-term outcomes.

The Making Pre-K Count study was designed to rigorously assess the short- and long-term effects of improving children's math experiences in prekindergarten (pre-K). Making Pre-K Count operated in community-based and public school pre-K classrooms in New York City that served mostly children from families with low incomes. Making Pre-K Count provided teachers with a high-quality math curriculum (Building Blocks) and ongoing teacher training and coaching.[3] In the Making Pre-K Count study, whole pre-K sites—community-based organizations and public schools—were randomly assigned to receive either the evidence-based math curriculum plus coaching and training (n = 35) or continue with pre-K-as-usual (n = 34). During the time when the program was implemented, there was a growing emphasis on early math instruction in all New York City schools.[4] Children in the control group therefore received more math instruction than had previously been typical in prior studies of early math education programs.[5]

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 grouped three to four children with one facilitator for math clubs that met three times a week for 30 minutes each session, outside of regular classroom instruction. Children who were in public schools that implemented Making Pre-K Count and stayed in the same public school were eligible for High 5s. In those Making Pre-K Count program public schools, individual children were randomly assigned within a school to either two years of math enrichment (Making Pre-K Count in pre-K plus

---

[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, 6 (2007): 1,428–1,446; Douglas H. Clements, Julie Sarama, and Carrie Germeroth, "Learning Executive Function and Early Mathematics: Directions of Casual Relations," *Early Childhood Research Quarterly* 36 (2016): 79–90.
[2]Greg J. Duncan and Katherine Magnuson, "Investing in Preschool Programs," *The Journal of Economic Perspectives* 27, 2 (2013): 109–132; Greg J. Duncan and Katherine Magnuson, "The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems," pages 47–69 in Greg J. Duncan and Richard J. Murnane (eds.), *Whither Opportunity: Rising Inequality, Schools, and Children's Life Chances* (New York: Russell Sage, 2011).
[3]Douglas H. Clements and Julie Sarama, *Building Blocks: Teacher's Edition* (Columbus, OH: McGraw-Hill Companies, Inc., 2013)
[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]Julie Sarama, Douglas H. Clements, Prentice Starkey, Alice Klein, and Ann Wakeley, "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 (2008): 89–119; 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, 2 (2011): 127–166.

High 5s in kindergarten, n = 320) or one year of math enrichment (Making Pre-K Count in pre-K and kindergarten as usual, n = 335).

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. 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. Its flagship projects, Making Pre-K Count and High 5s, were 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 report is the fifth report based on these studies.

A key feature of the Making Pre-K Count and High 5s studies was a focus on developing the math competencies of children enrolled in pre-K as a pathway to improving a broader set of children's outcomes into elementary school. Third grade is considered a particularly important moment in a child's educational experience. Literacy skill levels in third grade predict rates of high school completion.[6] While third grade may be a critical time for ensuring children's future success, few studies have tracked the effects of pre-K programs in the longer term, and the evidence on whether gains from pre-K interventions are sustained into early elementary school and beyond from those that have is mixed.[7]

The design of the Making Pre-K Count and High 5s studies makes it possible to rigorously assess the impact on children's outcomes from one year of math enrichment in pre-K (Making Pre-K Count compared with pre-K as usual), an additional year of math enrichment in kindergarten (Making Pre-K Count plus High 5s in kindergarten compared with Making Pre-K Count only), and two years of math enrichment (Making Pre-K Count plus High 5s in kindergarten compared with pre-K and kindergarten as usual). The samples of sites and children used in these analyses do not perfectly overlap, therefore the findings cannot be directly compared with one another. However, considered together, these analyses provide useful insights about the longer-term effects of early math enrichment interventions.

Earlier reports on these studies examined the effects of math enrichment at the end of pre-K and at the end of kindergarten.[8] The pre-K math program had small but not statistically significant effects on children's math skills by the end of kindergarten, and statistically significant effects on children's

---

[6]Duncan and Magnuson (2011); Catherine E. Snow, Susan M. Burns, and Peg Griffin, *Preventing Reading Difficulties in Young Children* (Washington, DC: National Academy Press, 1998); Donald J. Hernandez, *Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation* (Baltimore, MD: Annie E. Casey Foundation, 2011).
[7]Janet Currie and Duncan Thomas, "Does Head Start Make a Difference?" *The American Economic Review* 85, 3 (1995): 341–364; Eliana Garces, Duncan Thomas, and Janet Currie, "Longer-Term Effects of Head Start," *The American Economic Review* 92, 4 (2002): 999–1,012; James J. Heckman, Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Non-cognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics* 24, 3 (2006): 411–482; Jens Ludwig and Douglas L. Miller, "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *The Quarterly Journal of Economics* 122, 1 (2007): 159–208; David Deming, "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics* 1, 3 (2009): 111–134; Lawrence J. Schweinhart, "Long-Term Follow-Up of a Preschool Experiment," *Journal of Experimental Criminology* 9, 4 (2013): 389–409.
[8]Morris, Mattera, and Maier (2016); Shira K. Mattera, Robin Jacob, and Pamela A. Morris, *Strengthening Children's Math Skills with Enhanced Instruction: The Impacts of Making Pre-K Count and High 5s on Kindergarten Outcomes* (New York: MDRC, 2018);

math attitudes and working memory. The kindergarten math clubs had positive effects equivalent to an additional 2.5 months of math learning on one of two math measures at the end of kindergarten. The two programs jointly had a positive effect on one of two measures of children's math skills by the end of kindergarten, equivalent to over four months of additional math learning.

The current report presents the longer-term impacts on third-grade outcomes. The confirmatory outcome for these studies is children's third-grade math scores, since math skills are the direct target of the Making Pre-K Count and High 5s programs. The key confirmatory findings at the end of third grade are the following:

- **One year of math enrichment in pre-K:** Though not statistically significant, Making Pre-K Count had a small, positive, longer-term impact on children's third-grade math test scores (ES = 0.10), compared with pre-K as usual in control sites.

- **An additional year of math enrichment in kindergarten:** The impact of High 5s on children's third-grade math test scores in public schools, over and above the effect of Making Pre-K Count alone, was close to zero and not statistically significant (ES = 0.02).

- **Two years of math enrichment (pre-K and kindergarten):** Making Pre-K Count and High 5s together had moderate, statistically significant impacts on children's math test scores, compared with pre-K and kindergarten as usual in public schools (ES = 0.34).

The finding that two years of math enrichment (Making Pre-K Count plus High 5s) had moderate effects seems counter-intuitive given the small effects of each of the two interventions separately. This pattern of results is likely due to differences among the samples of children used in each analysis. Exploratory subgroup analyses suggest that early math enrichment may have been particularly beneficial for children with the most room to grow. Making Pre-K Count's impacts on third-grade math scores were fairly large—ranging from one-quarter to over a third of a standard deviation—for those children entering pre-K with the weakest language and attention skills. It appears that children with the lowest scores on the third-grade tests were more prevalent in the sample used to estimate the impact of two years of early math enrichment, and this difference may have contributed to the larger impacts observed in the sample.

The Making Pre-K Count and High 5s studies were also designed to test whether early math enrichment could have effects on outcomes beyond math skills. These outcomes are not the explicit focus of the programs, and empirical evidence that early math programming can have an impact on these outcomes is more limited.

Exploration of these outcomes suggest that Making Pre-K Count alone and when supplemented with High 5s may reduce chronic absenteeism and improve children's literacy test scores in third grade, though the findings are not always statistically significant. Making Pre-K Count alone, and when combined with an additional year of early math enrichment, led to a statistically significant reduction in children's chronic absenteeism in third grade, equivalent to about 9 percentage points or 28 percent. The effects of the programs on children's third-grade literacy test scores were similar in magnitude to the effects on third-grade math scores. None of the early math enrichment programs had an effect, positive or negative, on children's retention in a grade or placement in special education.

IMPLICATIONS

The Making Pre-K Count and High 5s studies rigorously tested the potential of early math enrichment interventions to both improve children's short-term outcomes and sustain these effects into elementary school.

- **These findings contribute to growing evidence about the longer-term importance of high-quality early math instruction for children, particularly those with the most room to grow.**

Correlational studies have suggested that early math skills could be a powerful lever for improving children's later skills, in math and in other domains. These studies hypothesize that early math learning may help children develop other skills, such as language skills and executive functioning, which may set the stage for effects on a wider range of longer-term outcomes. However, few studies have examined the long-term effects of enriched early math instruction to see whether or not the gains are sustained into elementary school.

The Making Pre-K Count and High 5s studies were designed not only to test the effects of the programs on math skills, but also to test whether early math programs could affect outcomes in other domains as well. The Making Pre-K Count and High 5s studies add to the base of evidence by demonstrating that enriched early math instruction has the potential to improve children's skills, both in math and other domains, and to sustain those improvements for at least four years.

Prior findings indicate that Making Pre-K Count had small, positive effects on outcomes in pre-K and kindergarten across multiple domains, including math skills, executive functioning, and children's attitudes toward math. This report finds that the effects of Making Pre-K Count were sustained into third grade, with small effects on children's math and literacy scores and favorable effects on chronic absenteeism. The effects of Making Pre-K Count on math test scores are comparable to those of other similar curricula implemented at scale and translate to approximately 12 percent of the achievement gap in fourth grade between low-income children and their high-income peers.[9] When children received two years of early math enrichment, the effects on tests are equivalent to approximately 40 percent of the achievement gap in fourth grade between low-income children and their high-income peers.

---

[9]Long-term effects from other interventions implemented at scale range from effects of 0.28 on third-grade literacy test scores from a social-emotional learning intervention to effects of 0.26 on fifth-grade math skills in a study of Building Blocks. Meghan P. McCormick, Robin Neuhaus, Erin E. O'Connor, Hope I. White, E. Parham Horn, Samantha Harding, Elise Cappella, and Sandee McClowry, "Long-Term Effects of Social-Emotional Learning on Academic Skills: Evidence from a Randomized Trial of INSIGHTS," *Journal of Research on Educational Effectiveness* 14, 1 (2021): 1–27; Tyler W. Watts, Greg J. Duncan, Douglas H. Clements, and Julie Sarama, "What Is the Long-Run Impact of Learning Mathematics During Preschool?" *Child Development* 89, 2 (2018): 539–555. Effect sizes in this study are standardized measures of the difference in outcomes at the end of third grade for the control and program groups. To contextualize these impacts, effect sizes are compared with other available standardized data on the difference in achievement between children who are eligible for free or reduced price lunch and those who are not eligible. Using National Assessment of Educational Progress data from 2,000 for children at the end of fourth grade, the achievement gap between those eligible for free or reduced price lunch and those not eligible was equivalent to 0.85 standardized units. Carolyn J. Hill, Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey, "Empirical Benchmarks for Interpreting Effect Sizes in Research," *Child Development Perspectives* 2, 3 (2008): 172–177. The effect of Making Pre-K Count on third-grade math scores (0.10) is equivalent to 12 percent of that difference.

The effects on chronic absenteeism are substantively meaningful. Rates of chronic absenteeism were approximately 33 percent among third-graders in the control group and 24 percent in the program group. Reducing absenteeism by 9 percentage points for third-graders citywide in New York City could lead to over 7,000 fewer chronically absent third-graders per year.[10] Chronic absenteeism is associated with lower achievement in reading and math and poor socioemotional outcomes, even after controlling for a wide range of background characteristics.[11]

The pattern of long-term effects, which suggests that impacts were largest for those with the most room to grow, supports the "academic risk hypothesis," which posits that effects of early childhood education may be the largest for children who need the most support.[12]

- **Well-designed math enrichment programs can have an effect even when layered on top of existing math instruction.**

The Making Pre-K Count program compared students who were exposed to a well-implemented, evidence-based early math enrichment program with their peers in other New York City pre-K programs. All students in the sample attended pre-K. During the time in which the program was implemented, there was a growing emphasis on early math instruction in New York City schools, and even children in the control group received more math instruction than had been typical in previous studies of early math enrichment interventions.[13] Thus, these long-term impacts reflect the added value of implementing high-quality math instruction in pre-K, above and beyond the impact of pre-K itself and of typical pre-K math instruction.

The Making Pre-K Count and High 5s studies contribute new evidence about the effects of early math enrichment experiences on children's later outcomes. Such experiences can lead to lasting gains for children, particularly for children with the greatest need.

[10]New York City had 78,141 third-graders in 2019-2020. New York State Education Department, "NYC Public Schools at a Glace 2019-20" (2020), website: www.data.nysed.gov/profile.php?instid= 7889678368. An analysis by New York University estimated that 22.8 percent of students were chronically absent in 2018. Research Alliance for New York City Schools, "How Has Attendance in NYC Schools Changed Over Time?" (2019), website: www.steinhardt.nyu.edu/research-alliance/research/spotlight-nyc-schools/how-has-attendance-nyc-schools-changed-over-time. According to those numbers, an estimated 17,816 third-graders would be chronically absent. After a reduction of chronic absenteeism by 9 percentage points, an estimated 10,784 third-graders would be chronically absent.

[11]Mariajosé Romero and Young-Sun Lee, *A National Portrait of Chronic Absenteeism in the Early Grades* (New York: National Center for Children in Poverty, 2007); Michael A. Gottfried, "Chronic Absenteeism and Its Effects on Students' Academic and Socioemotional Outcomes," *Journal of Education for Students Placed at Risk* 19, 2 (2014): 53–75.

[12]Bridget K. Hamre and Robert C. Pianta, "Can Instructional and Emotional Support in the First-Grade Classroom Make a Difference for Children at Risk of School Failure?" *Child Development* 76, 5 (2005): 949–967; Bridget K. Hamre and Robert C. Pianta, "Early Teacher-Child Relationships and the Trajectory of Children's School Outcomes through Eighth Grade" *Child Development* 72, 2 (2001): 625–638.

[13]Morris, Mattera, and Maier (2016); Sarama et al. (2008); Clements et al. (2011).

1

Introduction

Studies have found that math skills in early childhood are uniquely and strongly predictive of outcomes later in life. Strong early math skills are associated with not only later math achievement, but also better reading skills and executive functioning.[1] Further, studies have shown that early math competencies predict outcomes well into adulthood including the likelihood of graduating from high school and college completion.[2] While compelling, these studies are all based on correlational data, and few studies to date have tried to rigorously assess the impact of improving early math skills on later outcomes.

The Making Pre-K Count and High 5s studies were designed to rigorously assess the short- and long-term effects of improving children's math experiences in prekindergarten (pre-K) and kindergarten. Making Pre-K Count 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.[3] The Making Pre-K Count study compared students who were exposed to a well-implemented, evidence-based math program with their peers in other New York City pre-Ks.[4] Making Pre-K Count operated in community-based and public school pre-Ks that served mostly children from families with low incomes. During the time when the program was implemented, there was a growing emphasis on early math instruction in New York City schools, and children in the control group received more math instruction than had been observed in prior studies of early math education programs.[5]

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 grouped three to four children with one facilitator for math clubs that met three times a week for 30 minutes each session, outside of regular classroom instruction. The High 5s study was only conducted in the public school sites, where children could stay in the same school for pre-K and kindergarten. The community-based Pre-K sites were not included in the High 5s study because children who attended them dispersed to schools across the city for kindergarten. The High 5s study compared children who were offered Making Pre-K Count in pre-K and High 5s in kindergarten with children who were offered only Making Pre-K Count in pre-K.

[1]Duncan et al. (2007); Clements, Sarama, and Germeroth (2016).
[2]Duncan and Magnuson (2013); Duncan and Magnuson (2011).
[3]Clements and Sarama (2013).
[4]Morris, Mattera, and Maier (2016).
[5]Morris, Mattera, and Maier (2016); Sarama et al., (2008); Clements et al. (2011).

The design of the Making Pre-K Count and High 5s studies also makes it possible to evaluate the effect of two years of early math enrichment (Making Pre-K Count plus High 5s) compared with no early math enrichment. The analyses comparing these impacts are based on samples used in both Making Pre-K Count and High 5s studies and therefore only include children eligible for High 5s—that is, children in public schools sites who stayed in the same school for pre-K and kindergarten.

Previous reports examined the impact of Making Pre-K Count and High 5s on children's outcomes at the end of both pre-K and kindergarten.[6] By the end of kindergarten, Making Pre-K Count had small, positive, but not consistently statistically significant impacts on one of two measures of children's math skills, and statistically significant impacts on both attitudes toward math and working memory skills, compared with children who had not received math enrichment in pre-K. Making Pre-K Count did not have statistically significant impacts on children's language or inhibitory control skills. High 5s led to positive and statistically significant impacts on one of two measures of students' math skills, when compared with Making Pre-K Count alone. High 5s did not have statistically significant impacts on children's attitudes toward math, language skills, or executive functioning, when compared with students who received Making Pre-K Count only. The two years of aligned math enrichment (Making Pre-K Count plus High 5s) led to positive and statistically significant impacts on one of two measures of students' math skills and also led to more positive attitudes toward math among students, compared with those who had received no math enrichment in either pre-K or kindergarten. The two years of combined math enrichment programming did not have statistically significant impacts on children's language skills or executive functioning, or on the other, more global, measure of math skills.

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 research organization. Its flagship projects, Making Pre-K Count and High 5s, were 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 report presents longer-term effects of the two interventions on children's outcomes in third grade. It is the fifth report based on these studies.

WHY THIRD GRADE?

A key feature of the Making Pre-K Count and High 5s studies was a focus on developing pre-K children's math competencies as a pathway to improving a broader set of children's outcomes into elementary school. Third grade is considered a particularly important moment in a child's educational experience. Research has consistently found that third-grade reading outcomes strongly predict future academic challenges including dropping out of high school.[7] Similarly, strong and

[6]Morris, Mattera, and Maier (2016); Mattera, Jacob, and Morris (2018).
[7]Snow, Burns, and Griffin (1998); Hernandez (2011).

sustained math skills in elementary school predict higher rates of high school completion and college enrollment.[8]

Third grade may be a critical time for putting children on track for future success, but existing evidence on whether pre-K interventions are able to sustain any early gains into elementary school is mixed. While initial impacts on cognitive and achievement test scores tend to fade, some of these early childhood education programs nonetheless appear to have important long-term effects on high school completion, college attendance, earnings, healthy behaviors, and criminal involvement.[9]

Although there are few studies with which to test the hypothesis, Heckman, Stixrud, and Urzua did posit in a 2006 study that some of the longer-term effects of pre-K programs are a result of impacts on a set of "non-cognitive skills" that are frequently unmeasured, such as executive functioning and self-regulation, academic motivation or attitudes, and social-emotional skills.[10] Correlational findings suggest that math skills may have spillover effects into these non-cognitive domains, which may help to sustain longer-term impacts.[11] The Making Pre-K Count and High 5s studies were designed to examine this hypothesis by testing the short- and long-term effects of early math enrichment across both cognitive and non-cognitive domains.

Another hypothesis for the observed fading out of the effects on cognitive and achievement outcomes is that the instruction that children receive in pre-K, in terms of the instructional content or pedagogical approach, is not well aligned with the instruction they receive in kindergarten and beyond.[12] This "sustaining environments" hypothesis suggests that better aligning instructional experiences in pre-K with those in early elementary school could help sustain the impacts of programs implemented in pre-K.[13] The High 5s study was expressly designed to test whether an additional year of aligned math enrichment would help maintain the effects of early math enrichment into elementary school.

THIRD-GRADE FOLLOW-UP

This report presents the longer-term impacts of early math enrichment in pre-K (Making Pre-K Count) and in kindergarten (High 5s) on children's third-grade outcomes. The research team obtained data on children's third-grade test scores, chronic absenteeism, retention in a grade, and placement in special education from the New York City Department of Education administrative records. The confirmatory outcome for these studies is children's third-grade math scores, since math skills are the direct target of the Making Pre-K Count and High 5s programs. The Making Pre-K Count and High 5s studies were also designed to test whether early math enrichment could have effects on outcomes beyond math skills. These other outcomes are considered exploratory because they are not

[8]Duncan and Magnuson (2011).
[9]Currie and Thomas (1995); Garces, Thomas, and Currie (2002); Heckman, Stixrud, and Urzua (2006); Ludwig and Miller (2007); Deming (2009); Schweinhart (2013).
[10]Heckman, Stixrud, and Urzua (2006).
[11]Sarama, Lange, Clements, and Wolfe (2012); Blair, Knipe, and Gamson (2008).
[12]Engel, Claessens, and Finch (2013); Engel, Claessens, Watts, and Farkas (2015); Lee and Loeb (1995); Bailey, Jenkins, and Alvarez-Vargas (2020).
[13]Bailey, Jenkins, and Alvarez-Vargas (2020).

the explicit focus of the programs, and empirical evidence that early math programming can have an impact on these outcomes is more limited.

Though not statistically significant, Making Pre-K Count had a small, positive, longer-term impact on the studies' confirmatory outcome of third-grade math test scores, compared with the pre-K as usual in the control sites. Making Pre-K Count led to small, positive, not statistically significant impacts on third-grade literacy test scores and moderate, statistically significant reductions in rates of chronic absenteeism, both exploratory outcomes. The program did not have effects on children's retention in a grade or placement in special education.

While High 5s had effects on children's math skills in the year it was implemented, at the end of third grade, its impact on children's math test scores, over and above the effect of Making Pre-K Count alone, was close to zero and not statistically significant. High 5s was implemented in public schools only. The effects of High 5s on exploratory outcomes in third grade were also close to zero and not statistically significant.

At the end of third grade, Making Pre-K Count and High 5s together had moderate, statistically significant impacts on children's math test scores, compared with pre-K and kindergarten as usual. These two years of aligned early math enrichment also led to positive effects on children's third-grade literacy test scores and chronic absenteeism, both exploratory outcomes. They did not have effects on children's retention in a grade or placement in special education. This analysis only includes public schools and those students who remained in the same school for pre-K and kindergarten.

The finding that two years of early math enrichment (Making Pre-K Count plus High 5s) had moderate effects seems counter-intuitive given the small effects of each of the two interventions separately. This pattern of results is likely due to differences among the samples of children used in each analysis. Exploratory subgroup analyses suggest that early math enrichment may have been particularly beneficial for children with the most room to grow. For example, Making Pre-K Count's impacts on third-grade math scores were fairly large—ranging from one-quarter to over a third of a standard deviation—for those children entering pre-K with the weakest language and attention skills. It appears that children in the control group sample used to estimate the impact of two years of early math enrichment also had room to grow (having low third-grade test scores), and this difference may have contributed to the larger impacts observed in the sample.

This report explains the above findings in greater detail. Chapter 2 presents the research design, sample, and measures used in the studies. Chapter 3 describes the impacts of enhanced math experiences in pre-K and kindergarten on third-grade outcomes. Chapter 4 concludes with a discussion of the potential implications of these findings.

2

Design, Sample, and Measures

This chapter describes the design of the Making Pre-K Count and High 5s studies and the analysis of impacts on third-grade outcomes. The studies rigorously tested the effects of early math enrichment in prekindergarten (pre-K) and kindergarten using randomized controlled trials. Children were tracked from pre-K through third grade, and data on their outcomes were collected from the New York City Department of Education administrative records. The studies examine the effects of (1) one year of math enrichment in pre-K (Making Pre-K Count), compared with pre-K as usual, (2) a supplemental year of math enrichment in kindergarten (Making Pre-K Count plus High 5s), compared with math enrichment in pre-K only (Making Pre-K Count), and (3) two years of math enrichment (Making Pre-K Count plus High 5s), compared with pre-K and kindergarten as usual.

DESIGN

The research team tested Making Pre-K Count and High 5s using a rigorous two-stage random assignment design. Figure 2.1 illustrates this design.

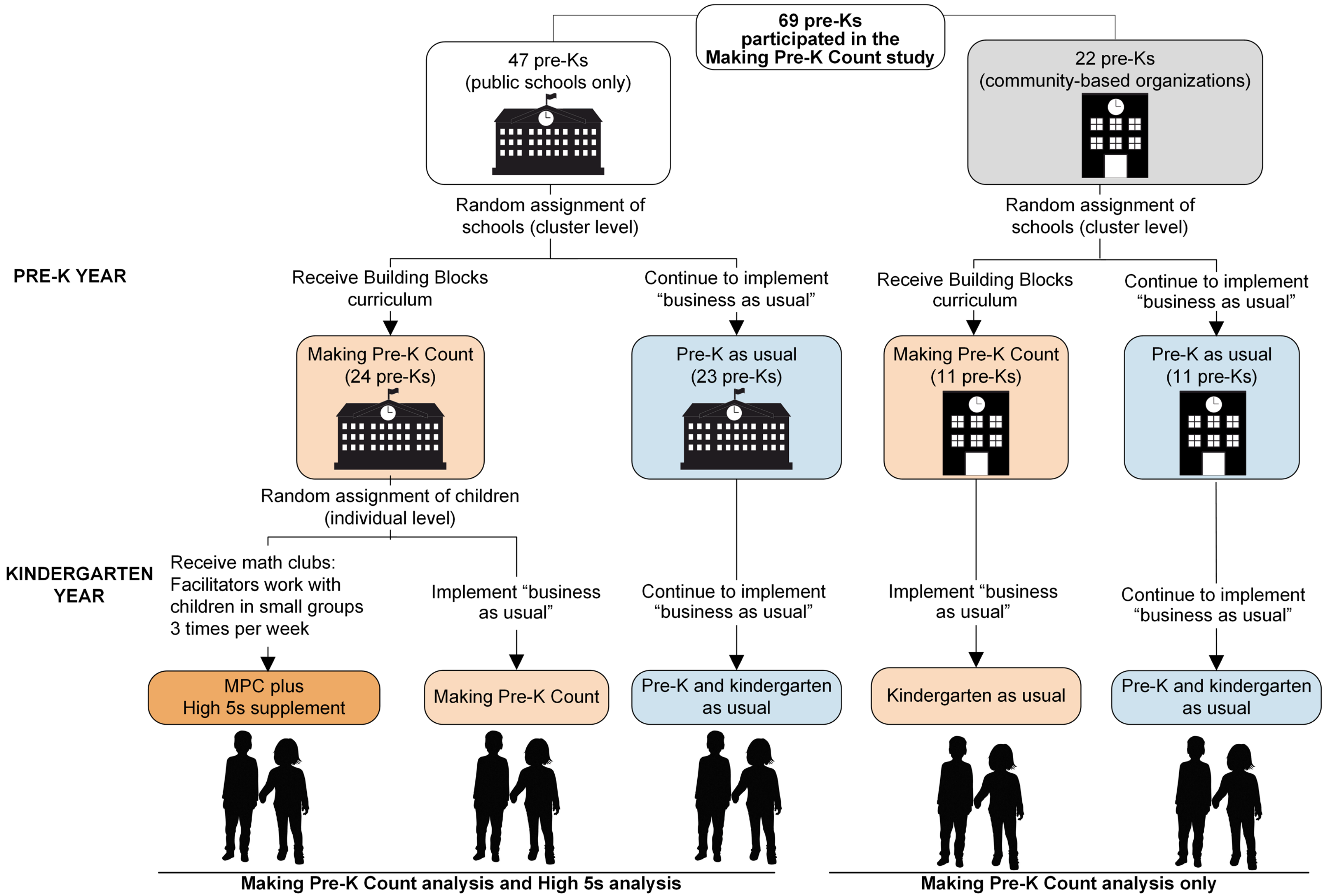
The Making Pre-K Count study tested the effects of an evidence-based pre-K math curriculum (Building Blocks), which was supported by two years of teacher training and in-classroom coaching.[1] In the study, the research team randomly assigned whole pre-K sites across New York City either to receive the evidence-based math curriculum and teacher training and coaching or to continue with pre-K as usual. The team blocked the sites by location (city borough), type (public school or community-based organization), and racial/ethnic composition (sites serving over 60 percent Hispanic children or sites serving children from other racial or ethnic backgrounds).[2] Groups of four to five sites were randomly assigned within blocks. The team estimated the effects of Making Pre-K Count by comparing the outcomes of children in the pre-K sites that implemented Making Pre-K Count with those of children in sites that continued with Pre-K as usual.

[1]Clements and Sarama (2013).

[2]Morris, Mattera, and Maier (2016); Mattera, Jacob, and Morris (2018). Sites were "blocked" into groups of four to five before randomization based on their borough, venue (community-based organizations versus school-based sites), and the racial/ethnic composition of the children (whether the sites served primarily Hispanic children or not). Blocking achieves two goals: First, it reduces the risk of a poor match between program and control groups by accident given the small number of units at the level of randomization; second, blocking in groups rather than pairs protects against the loss of sample sites between randomization and the study of program impact by allowing for the retention of all remaining sites if a single site drops out of the study.

Figure 2.1

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



The High 5s program was implemented in the year after children completed pre-K. Children who attended public schools that received Making Pre-K Count and who stayed in the same school for pre-K and kindergarten were eligible for High 5s. For the High 5s study, the research team randomly assigned individual eligible children within their public school to either the High 5s program group (Making Pre-K Count plus High 5s) or a kindergarten-as-usual group (Making Pre-K Count only group). Children who attended pre-K and kindergarten in the same pre-K-as-usual public school sites, which did not implement any early math enrichment, constituted the pre-K-and-kindergarten-as-usual control group.

This two-stage sequential random assignment design thus created three experimental groups at the Making Pre-K Count public school sites. The research team used these groups to investigate two additional comparisons: (1) the effects of two years of early math enrichment (Making Pre-K Count plus High 5s), compared with one year of early math enrichment (Making Pre-K Count) and (2) the effects of two years of early math enrichment (Making Pre-K Count plus High 5s), compared with no math enrichment (pre-K and kindergarten as usual). Table 2.1 summarizes the analytic samples for the three confirmatory study comparisons.

Table 2.1

Analytic Samples for Third-Grade Confirmatory Study Comparisons

	Making Pre-K Count (MPC) vs. Pre-K as Usual		MPC Plus High 5s vs. MPC and Kindergarten as Usual		MPC Plus High 5s vs. Pre-K and Kindergarten as Usual	
Analytic Sample	**Program Group**	**Control Group**	**Program Group**	**Control Group**	**Program Group**	**Control Group**
Blocks	16	16	-	-	11	10
Sites	35	34	-	-	24	22
Students[a]	1,165	1,112	274	282	274	313

NOTES: The second and third comparisons only include public schools and their students (no community-based organizations).

The program students in the second and third comparisons are the same (n = 274).

[a]These sample sizes refer to the analytic samples in the study. They are inclusive of any students with outcome data in third grade.

SAMPLE AND ANALYTIC STRATEGY

Making Pre-K Count

The Making Pre-K Count study was conducted in 69 pre-K sites, including 173 classrooms, across New York City. Thirty-five sites were randomly assigned to the program group and received the Building Blocks curriculum and teacher training and coaching. Thirty-four sites were randomly assigned to the control group and continued with their usual pre-K practices. The sites in the program group implemented Making Pre-K Count over two school years, the 2013-2014 academic year and the 2014-2015 academic year. The first year was a "soft start" and allowed teachers to become familiar with the curriculum and receive training. For this reason, children participating in the program in the second academic year have been the main focus of the Making Pre-K Count study to date and make up the confirmatory sample (full implementation year sample) for this analysis. The research team also estimated the impacts of Making Pre-K Count on the outcomes of the students in the soft start year sample and three other exploratory samples to check that the pattern of effects was consistent across different samples. Appendix A describes these exploratory samples in greater detail.

The analytic strategy for estimating the impacts of Making Pre-K Count on children's third-grade outcomes builds on the strategy for estimating program impacts in kindergarten. The research team used multilevel modeling to account for the data's nested structure, with children nested within pre-K sites and the sites nested within blocks. The team estimated the program's impacts by comparing mean outcomes of students in the Making Pre-K Count group with those of students in the pre-K-as-usual control group, applying a regression adjustment for selected background characteristics and dummy variables for random assignment blocks. See Appendix B for further details about the analysis.

High 5s

The High 5s study was embedded in the larger Making Pre-K Count study. It was conducted in the 2015-2016 academic year in the 24 public schools that implemented Making Pre-K Count in 2013-2015. Children who stayed in the same public school for pre-K and kindergarten were eligible for High 5s. The research team randomly assigned the eligible children individually within their school to either a program group that received High 5s or a control group that received kindergarten as usual. The team randomly assigned a total of 655 children, 320 to the Making Pre-K Count plus High 5s program group and 335 to the Making Pre-K Count only (kindergarten-as-usual) control group. These students make up the High 5s sample, which the research team used to estimate the effect of math enrichment in kindergarten over and above the impact of Making Pre-K Count alone. Because the High 5s study involved two stages of random assignment (one for the pre-K sites and the other for the individual kindergarteners), the Making Pre-K Count plus High 5s group could also be compared with a third group of students: those in public schools who received no math enrichment in either pre-K or kindergarten (the pre-K-and-kindergarten-as-usual control group). This two years of math sample thus consisted of the 320 students in the Making Pre-K Count plus High 5s program group and the 345 students in the pre-K-and-kindergarten-as-usual control group.

The analytic strategy for estimating the impacts of High 5s on children's third-grade outcomes also builds on the strategy used to estimate the program's impacts in kindergarten. The research team estimated the effect of High 5s (over and above the effect of Making Pre-K Count only) by comparing the outcomes of children in the Making Pre-K Count plus High 5s group with the outcomes of children in the Making Pre-K Count-only group, applying a regression adjustment for selected background characteristics and dummy variables for school. This analysis only included the 47 public school sites in the larger Making Pre-K Count study. See Appendix B for further details about the analysis.

The research team estimated the effects of two years of early math enrichment by comparing the mean outcomes of children in the Making Pre-K Count plus High 5s group with the mean outcomes of children in the pre-K-and-kindergarten-as-usual control group. Because this comparison builds off the study's cluster-level random assignment design, the team used multilevel modeling to account for the nested structure of the data. The team applied a regression adjustment to the analysis for selected background characteristics and dummy variables for random assignment blocks. This analysis also only included the 47 public school sites in the larger Making Pre-K Count study. See Appendix B for further details about the analysis.

Attrition

The What Works Clearinghouse (WWC) standards, which provide boundaries for acceptable levels of attrition for minimizing bias in randomized controlled trials, guided the calculations for overall and differential attrition by third grade.[3] The WWC provides specific guidelines for judging whether the combination of overall and differential individual-level attrition is high and in need of a baseline equivalence testing under an "optimistic" standard for early childhood education studies, reflecting the WWC's assumption that most attrition in studies of interventions results from exogenous factors. Under optimistic assumptions, overall attrition up to 40 percent is acceptable when paired with differential attrition levels below 6 percent. Under cautious assumptions, overall attrition up to 40 percent is acceptable when paired with differential attrition levels below 2.6 percent.

Specifically, of the 2,702 eligible students in the Making Pre-K Count sample (for the confirmatory sample and outcome), 1,844 students remained in the New York City Department of Education data system, indicating an overall attrition rate of 32 percent. Differential attrition was 1.5 percent, with the program group having an attrition rate of 33 percent and the control group having a rate of 31 percent.[4] The High 5s math sample had an overall attrition rate of 30 percent and differential attrition rate of 2.0 percent, with the program group having an attrition rate of 29 percent and the control group having a rate of 31 percent. Overall and differential attrition rates in Making Pre-K Count and High 5s fall below both optimistic and cautious thresholds for differential attrition (1.5 to 2.0 percent) and overall attrition (30 to 32 percent).

MEASURES

The research team obtained data on children's third-grade outcomes from the New York City Department of Education administrative records, via the Research Alliance for New York City Schools. Outcomes included retention in a grade, placement in special education, chronic absenteeism, and test scores. Demographic data used as covariates included students' age, gender, race, and primary language at home.

Estimating impacts on all available outcomes indiscriminately could lead to finding that some estimates were statistically significant due to chance alone. The team used a multi-tiered approach to reduce this likelihood, while preserving the power to identify "true" program impacts. Currently,

---

[3]What Works Clearinghouse (2020).

[4]Across the five outcomes for the Making Pre-K Count full implementation year sample, total attrition ranged from 16 percent to 34 percent and differential attrition ranged from 0 percent to 2 percent. What Works Clearinghouse labels these as low attrition rates.

there is little consensus in the field of statistics or evaluation on the most appropriate methods for adjusting statistical tests to account for multiple comparisons. Moreover, while these statistical adjustments may make it less likely to find false positives, it is not clear that it is worth the tradeoff in making it harder to identify true positives. Therefore, rather than correcting for multiple comparisons, the team (1) carefully limited the number of outcomes in its analysis and (2) grouped research questions into confirmatory and exploratory categories.[5]

The confirmatory outcome for these studies was children's third-grade math scores, since math skills are the direct target of the Making Pre-K Count and High 5s programs. The studies used the New York State third-grade standardized math test score as a measure of math skills.[6] The confirmatory questions for these studies were (1) to what extent does Making Pre-K Count affect children's third-grade math test scores; (2) to what extent does High 5s affect children's third-grade math test scores; and (3) to what extent do two years of math enrichment (Making Pre-K Count in pre-K and High 5s in kindergarten) affect children's third-grade math test scores.

Exploratory research questions focused on (1) outcomes other than math skills and (2) subgroup analyses. Other outcomes included those the research team theorized could be indirectly affected by the Building Blocks program based on educational and developmental theory (and some empirical evidence, albeit more limited than for confirmatory outcome). For example, there is a growing consensus that math can help build language skills as well as math understanding. Math concepts such as counting and shapes expand and enrich vocabulary, as children use language to express and justify mathematical thinking.[7] In one study, children in Building Blocks classrooms significantly outperformed children in control classrooms on a measure of oral language.[8] While these outcomes are not the explicit focus of the Making Pre-K Count and High 5s programs, the studies were designed not only to test the effects of the programs on math skills, but also to test whether early math enrichment interventions could affect outcomes in other domains.

The New York State third-grade standardized English language arts test score was used as a measure of reading skills. The research team assessed children's retention in a grade by whether the student was below the expected grade level four years after pre-K.[9] The team measured chronic absenteeism by whether a student was present 90 percent of the days or less (absent 10 percent or more of school days) in third grade.[10] Finally, the team measured placement in special education by whether the

[5]The research team preregistered confirmatory and exploratory measures as part of the Making Pre-K Count, High 5s, and two years of math analytic plans. The preregistered plans can be found at: https://osf.io/bm6va, https://osf.io/ujxnr, and https://osf.io/68yxg.

[6]These scores were normed to have a mean of 0 and standard deviation of 1 using all New York City test scores for a given test (that is, math or English language arts) during a given school year. Children in the soft start year were expected to be in third grade in the 2017-2018 academic year and children in the full implementation year were expected to be in third grade in the 2018-2019 academic year, so scores were for children taking the test in their expected third-grade year.

[7]Ginsburg, Lee, and Boyd (2008).

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

[9]Four years after Making Pre-K Count refers to the academic year when students were expected to be in third grade. It refers to the 2017-2018 academic year for the soft start year and the 2018-2019 academic year for the full implementation year.

[10]In a 178-day academic year, a student would be considered chronically absent for missing 18 days or more. If the child was previously retained in a grade, chronic absenteeism was collected in whatever grade they were attending four years after participating in Making Pre-K Count—the 2017-2018 academic year for the soft start year and the 2018-2019 academic year for the full implementation year.

student had an Independent Education Program documented by the New York City Department of Education four years after participating in Making Pre-K Count.

The research team classified all subgroup analyses as exploratory because, at the start of Making Pre-K Count, there was little available evidence that the impact of Building Blocks varied by baseline characteristics of children, teachers, or sites. Subgroup analyses examined whether impacts varied by site (community-based organization versus public school) and child characteristics (race or ethnicity, language at home, gender, entering skill level).

3

Third-Grade Impacts

This chapter presents the effects of the Making Pre-K Count prekindergarten (pre-K) math curriculum and High 5s kindergarten math clubs on children's third-grade math and literacy test scores, retention in a grade, chronic absenteeism, and placement in special education.

- **One year of math enrichment in Pre-K:** Though not statistically significant, Making Pre-K Count had a small, positive, longer-term impact on children's third-grade math test scores (ES = 0.10), compared with the pre-K as usual at the control group sites.

- **An additional year of math enrichment in kindergarten:** The impact of High 5s on children's third-grade math test scores in public schools, over and above the effect of Making Pre-K Count alone, was close to zero and not statistically significant (ES = 0.02).

- **Two years of math enrichment (in pre-K and kindergarten):** Making Pre-K Count and High 5s together had moderate, statistically significant impacts on children's math test scores, compared with pre-K and kindergarten as usual in public schools (ES = 0.34).

The finding that two years of enrichment (Making Pre-K Count plus High 5s) had moderate effects seems counter-intuitive given the small effects of each of the two interventions separately. This pattern of results is likely due to differences among the samples of children used in each analysis. Exploratory subgroup analyses suggest that early math enrichment may have been particularly beneficial for children with the most room to grow—that is, those children entering pre-K with the weakest skills or the lowest test scores.

The research team also estimated the impact of these two early math enrichment interventions on other, exploratory outcomes. These exploratory analyses suggest that Making Pre-K Count alone and Making Pre-K Count plus High 5s reduced chronic absenteeism and improved children's literacy test scores, though findings were not always statistically significant for literacy test scores. The early math interventions had an effect close to zero on children's retention in a grade or placement in special education.

IMPACTS OF MAKING PRE-K COUNT

This section presents findings comparing third-grade outcomes of children who received one year of math enrichment in pre-K (Making Pre-K Count) with outcomes of children who received pre-K as usual. Table 3.1 presents the pre-K program effects for both the confirmatory outcome—math skills—and the exploratory outcomes—literacy test scores, chronic absenteeism, retention in a grade, and placement in special education. Appendix C presents the program's effects on further exploratory samples.

Table 3.1

Impacts of Making Pre-K Count on Third-Grade Outcomes

Outcome Measure	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Standard Error	Effect Size[a]
Math[b]	**-0.02**	**-0.12**	**0.10**	**0.19**		**0.08**	**0.10**
Literacy[c]	0.02	-0.09	0.11	0.12		0.07	0.11
Chronic absenteeism (%)[d]	23.6	32.6	-9.0	0.00	***	3.0	-0.19
Retention (%)[e]	12.5	12.1	0.4	0.84		2.0	0.01
Special education (%)[f]	18.6	20.1	-1.5	0.40		1.7	-0.04
Sample size							
Blocks	16	16					
Sites	35	34					
Students[g]	945	899					

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the full implementation year sample.

Bolded outcome is confirmatory, all others are exploratory.

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.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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]Citywide standardized z-score for state third-grade math test.

[c]Citywide standardized z-score for state third-grade English language arts test.

[d]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.

[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

[g]The sample size refers to the number of students from the full implementation year sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the full implementation year analytic sample, 81 percent have data for math and at least 79 percent have data for all other outcomes in the table.

- **Making Pre-K Count had a positive but not statistically significant impact on children's third-grade math skills, the study's confirmatory outcome.**

At the end of third grade, the program impacts on math test scores were positive but not statistically significant (ES = 0.10, p = 0.19).[1] Students in control group pre-K sites scored 0.12 standard deviations below the citywide average, and students in Making Pre-K Count sites scored 0.02 standard deviations below the average, a difference of 0.10 standard deviations. This effect is comparable to those found from other curricula implemented at scale and translate to approximately 12 percent of the achievement gap in fourth grade between low-income children and their high-income peers.[2]

- **The effects of Making Pre-K Count on all exploratory outcomes were in a favorable direction. Making Pre-K Count had a positive and marginally statistically significant impact on children's third-grade literacy skills. The program had a favorable and statistically significant impact on children's chronic absenteeism in third grade. The effect on children's retention in a grade or placement in special education was close to zero.**

As with math, at the end of third grade, the program impacts on literacy test scores were positive and not statistically significant (ES = 0.11, p = 0.12). Students in control group pre-K sites scored 0.09 standard deviations below the citywide average in reading, while students in Making Pre-K Count sites scored 0.02 standard deviations above the citywide average. This effect on literacy test scores is similar in magnitude to the effect on math test scores.[3] A child's reading skills in the third grade have long been an important indicator of whether the child completes high school and attends college, and numerous policy initiatives around the country focus on improving third-graders' reading skills as a crucial policymaking lever for improving a child's academic trajectory and later outcomes.[4] The effect is comparable to those found from other curricula implemented at scale and translates to approximately 15 percent of the achievement gap in third grade between low-income children and their high-income peers.[5]

---

[1]The effect size is calculated by dividing the estimated effect of the program (the difference between the means of the program group and the control group) by the standard deviation for the control group. An effect size of 0.10 here represents an improvement in math test scores equal to one-tenth of the standard deviation.
[2]Long-term effects from other interventions implemented at scale range from effects of 0.28 on third-grade literacy test scores from a social-emotional learning intervention (McCormick et al., 2021) to effects of 0.04 on fourth-grade math skills and 0.26 on fifth-grade math skills in a study of Building Blocks (Watts, Duncan, Clements, and Sarama, 2018). Effect sizes in the Making Pre-K Count study are standardized measures of the difference in outcomes at the end of third grade for the control and program groups. To contextualize these impacts, the research team compared the effect sizes with other available standardized data on the difference in the achievement gap between children who are eligible for free or reduced price lunch and those who are not eligible. As described in Hill, Bloom, Black, and Lipsey (2008), using National Assessment of Educational Progress (NAEP) data from 2000 for children at the end of fourth grade, the achievement gap between low-income children and their high-income peers was equivalent to 0.85 standardized units for math. The effect of Making Pre-K Count on third-grade math scores (0.10) is equivalent to 12 percent of that difference.
[3]An evaluation of a social-emotional early childhood curriculum in New York City also found statistically significant effects on reading but not math test scores, despite the effects on reading and math being similar in magnitude. This suggests that the reading test used in New York City may be somewhat more sensitive to program impacts (McCormick et al., 2021).
[4]Snow, Burns, and Griffin (1998); Duncan and Magnuson (2011); Hernandez (2011); Rose and Schimke (2012).
[5]Effect sizes in this study are standardized measures of the difference in outcomes at the end of third grade for the control and program groups. To calculate the proportion of the achievement gap, the research team compared the effect sizes with standardized measures of the difference in outcomes at the end of fourth grade of children who are eligible for free or reduced price lunch and those who are not eligible. As described in Hill, Bloom, Black, and Lipsey (2008), using NAEP data

Chronic absenteeism is a nationwide problem. In the 2015-2016 academic year, the U.S. Department of Education estimated that roughly 16 percent of students nationwide were chronically absent, with rates of chronic absenteeism often considerably higher in cities.[6] In the Making Pre-K Count study, approximately 28 percent of third-graders were chronically absent. Among older students, absenteeism is a strong predictor of both high course failure rates and low graduation rates.[7] A 2007 study of graduation patterns in Chicago Public Schools found absenteeism was eight times more predictive of course failure than test scores.[8] For younger students, research has shown that chronic absenteeism is associated with lower achievement in reading and math and poor socioemotional outcomes, even after controlling for a wide range of background characteristics.[9]

Making Pre-K Count had a favorable, statistically significant impact of 9 percentage points on children's chronic absenteeism in third grade (ES = -0.19, $p < 0.01$). Thirty-three percent of students who received pre-K as usual were chronically absent in third grade compared with 24 percent of students who received math enrichment in pre-K, which translates to a 28 percent reduction in chronic absenteeism.[10]

Rigorous studies of interventions designed to reduce absenteeism are rare, and those that do exist measure outcomes in a variety of different ways. However, a recent randomized control trial of the Early Warning Intervention and Monitoring System (EWIMS) gives a sense of the general magnitude of the effects these interventions can have. The EWIMS program includes highly detailed and structured guidance for schools, along with a tool to help monitor student attendance and academic performance; the evaluation indicated that the program reduced chronic absenteeism rates from 14 to 10 percent (a 28 percent decrease) after one year.[11] This effect is comparable to the decrease due to Making Pre-K Count.

Making Pre-K Count did not have an effect, positive or negative, on children's retention in a grade or placement in special education. These outcomes are not specific targets of Making Pre-K Count; however, some pre-K programs have found emerging effects on retention in a grade and placement in special education in kindergarten, although those effects did not always persist.[12]

The research team considers the subgroup analyses to be exploratory, and the Making Pre-K Count study was not primarily designed or powered to detect differences in impacts between groups of children or sites. Table 3.2 presents Making Pre-K Count's impact on the confirmatory outcome (math skills) for Hispanic and non-Hispanic students, boys and girls, and students whose primary language at home is English and those whose primary language at home is another language. The team tested subgroup effects in these groups using the pooled sample of students—students in both the "soft

from 2000, the income achievement gap was equivalent to 0.74 standardized units for literacy. The effect of Making Pre-K Count on third-grade reading scores (0.11) is equivalent to 15 percent of that gap.

[6]U.S. Department of Education (2019); Civil Rights Data Collection (2016).

[7]Allensworth and Easton (2007); Baltimore Education Research Consortium (2011).

[8]Allensworth and Easton (2007).

[9]Romero and Lee (2007); Gottfried (2014).

[10]The effect of Making Pre-K Count on chronic absenteeism began early and continued as children moved through elementary school. (See Appendix E.)

[11]Faria et al. (2017)

[12]Morris et al. (2014); Lipsey et al. (2013).

Table 3.2

Impacts of Making Pre-K Count on Third-Grade Math, by Demographics

Subgroup	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size[a]	Difference Between Subgroups	P-Value
Race/ethnicity								
Hispanic	-0.04	-0.18	0.14	0.02 **	0.06	0.14	-0.05	0.65
Non-Hispanic	0.06	-0.03	0.09	0.31	0.09	0.09		
Gender								
Male	0.02	-0.11	0.12	0.09 *	0.07	0.12	0.05	0.62
Female	-0.03	-0.11	0.07	0.30	0.07	0.08		
Home language[b]								
Non-English	-0.03	-0.16	0.13	0.04 **	0.06	0.14	0.05	0.62
English	-0.01	-0.09	0.08	0.25	0.07	0.08		
Sample size								
Blocks	16	16						
Sites	35	34						
Students	1,952	1,894						

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the pooled sample (students from both the soft start year sample and full implementation year sample).

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.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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]This represents the primary language spoken in the child's home.

start" year and the full implementation year—in order to maximize power. Table 3.3 shows Making Pre-K Count's impacts on the confirmatory outcome for children entering pre-K with higher and lower relative skills. Data on children's baseline skills were only available for a subset of children in the second, full implementation year because only a random subsample of children were selected for baseline testing. Appendix D presents Making Pre-K Count's impacts on exploratory outcomes by subgroup.

Table 3.3

Impacts of Making Pre-K Count on Third-Grade Math, by Entering Skill Level

Subgroup	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Standard Error	Effect Size[a]	Difference Between Subgroups	P-Value
Entering language skill level[b]									
High	0.21	0.08	0.14	0.27		0.12	0.14	0.14	0.48
Low	-0.27	-0.55	0.28	0.08	*	0.16	0.28		
Entering self-regulation skill level[c]									
High	0.08	-0.06	0.14	0.28		0.13	0.14	0.22	0.24
Low	-0.04	-0.40	0.37	0.01	***	0.13	0.36		
Sample size									
Blocks	16	16							
Sites[d]	35	34							
Students[e]	298	295							

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the full implementation year sample.

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.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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 language skills were measured using the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011), administered at pre-K entry in the fall of 2014.

[c]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.

[d]As few as 33 program and control sites are represented for certain subgroups.

[e]There are 296 students in the program group and 294 students in the control group for the language subgroup.

- **The effects of Making Pre-K Count were similar in magnitude across different subgroups of children. The effects were similar across Hispanic and non-Hispanic students, across boys and girls, and across students whose primary language at home was English and those whose primary language at home was not English.**

A prior study found that the Building Blocks curriculum had larger impacts for Black students; however, those studies included a limited sample of Hispanic students (22 percent of study participants).[13] Making Pre-K Count included a larger sample of Hispanic children (54 percent of study participants), with the rest of the sample identifying primarily as non-Hispanic Black (37 percent). By third grade, Making Pre-K Count had positive and statistically significant effects on math skills (ES = 0.14, p = 0.02) for Hispanic students. The effect on math test scores is equivalent to 16 percent of the achievement gap between low-income children and their high-income peers. The impact on non-Hispanic children's math scores (ES = 0.09) was of a similar magnitude, although not statistically significant.[14] Making Pre-K Count had a positive but not statistically significant effect on math skills (ES = 0.10, p = 0.42) for Black boys.

Making Pre-K Count had a positive and statistically significant effect on boys' math scores (ES = 0.12, p = 0.09). The effect on girls' math scores was similar in magnitude but not statistically significant (ES = 0.08, p = 0.30). Making Pre-K Count had a positive and statistically significant effect on students whose primary language at home was not English (ES = 0.14, p = 0.04) and a positive but not statistically significant effect on students whose primary language at home was English (ES = 0.08, p = 0.25).

- **Making Pre-K Count had statistically significant impacts on third-grade math scores for children entering the study with lower skills, with large and positive effects for children entering pre-K with lower self-regulation or language ability.**

The research team assessed children's baseline pre-K skills for only a subset of children, and only in the second year of implementation. These analyses include only those children who were randomly selected for this baseline assessment.

Making Pre-K Count had large and statistically significant impacts on math test scores for those children rated as having lower self-regulation skills (being more impulsive) or weaker language skills when entering pre-K in the fall (ES = 0.36 and 0.28, respectively). The program had small, positive, and not statistically significant impacts for children with stronger skills in the fall of pre-K (ES = 0.14 and 0.14).

[13]Clements et al. (2011).

[14]Alternative ways of testing the effect of Making Pre-K Count for children in different racial or ethnic subgroups showed substantively similar results. Tests comparing random assignment blocks that served a majority of students who were Hispanic with random assignment blocks that served a majority of non-Hispanic students showed the same pattern of effects on math skills (ES = 0.15, p = 0.10 for Hispanic blocks and ES = 0.03, p = 0.74 for non-Hispanic blocks).

- **Effects of Making Pre-K Count on test scores were generally larger in magnitude for children in public schools.**

Making Pre-K Count was implemented in both public school and community-based sites. Findings from pre-K and kindergarten suggest that Making Pre-K Count had positive effects on math outcomes in public school sites and executive functioning outcomes in community-based sites. Table 3.4 presents Making Pre-K Count's impacts on third-grade outcomes for children who attended pre-K in public schools and for those who attended pre-K in community-based sites.[15]

For children who attended pre-K in public schools, Making Pre-K Count had positive marginally significant effects on third-grade math scores (ES = 0.16, p = 0.10) and statistically significant effects on third-grade reading scores (ES = 0.18, p = 0.03) and chronic absenteeism (ES = -0.18, p = 0.02). These test score effects are equivalent to 19 and 24 percent of the achievement gap between low-income children and their high-income peers.

For children who attended pre-K in community-based sites, Making Pre-K Count had small, negative, and not statistically significant effects on third-grade math and literacy scores. The program had favorable and statistically significant effects on third-grade chronic absenteeism (ES = -0.23, p = 0.06) and retention in a grade (ES = -0.15, p = 0.06) for children who attended pre-K in community-based sites. The program had a statistically significantly larger effect on retention in community-based sites than it did on the same outcome in public schools.

The High 5s kindergarten math club was implemented only in public school sites. Therefore, the effects of the High 5s program described below are *on top of* Making Pre-K Count's positive impacts on outcomes for children in public schools. Importantly, control group children who attended pre-K in public schools had lower third-grade math test scores (-0.20) than control group children who attended pre-K in community-based organizations (0.06), suggesting that the public school children may have also been a higher-risk group.

IMPACTS OF HIGH 5s

High 5s operated in the 24 public schools that had implemented the Making Pre-K Count program in pre-K. All children in the High 5s study stayed in the same school for pre-K and kindergarten and therefore had at least one year of math enrichment (Making Pre-K Count). The research team randomly assigned half of the children who received Making Pre-K Count to an additional, aligned math enrichment intervention (High 5s).

This section presents findings comparing third-grade outcomes for children assigned to two years of math enrichment (Making Pre-K Count plus High 5s) with third-grade outcomes for children assigned to one year of math enrichment (Making Pre-K Count only). The effects of High 5s are in addition to Making Pre-K Count's positive impacts on outcomes of children in public schools.

[15]The research team also estimated impacts in public schools and community-based sites across implementation years using the pooled sample. The results showed similar effects to those in the full implementation year.

Table 3.4

Impacts of Making Pre-K Count on Third-Grade Outcomes, by Venue (Community-Based Organization Versus Public School)

	Community Based-Organization				Public School					
Outcome Measure	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
Math[b]	0.06	-0.03	0.85	-0.03	-0.20	0.16	0.10	0.16	-0.18	0.28
Literacy[c]	0.09	-0.04	0.77	-0.04	-0.16	0.17	0.03 **	0.18	-0.21	0.17
Chronic absenteeism (%)[d]	26.7	-9.6	0.06 *	-0.23	35.2	-8.6	0.02 **	-0.18	-1.1	0.86
Retention (%)[e]	14.7	-5.4	0.06 *	-0.15	10.9	3.0	0.23	0.09	-8.4	0.03 ††
Special education (%)[f]	20.8	-0.1	0.98	0.00	19.7	-2.1	0.30	-0.05	2.1	0.59
Sample size										
Blocks	5				11					
Sites	11				23					
Students[g]	267				632					

(continued)

Table 3.4 (continued)

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the full implementation year sample.

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 in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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]Citywide standardized z-score for state third-grade math test.

[c]Citywide standardized z-score for state third-grade English language arts test.

[d]The outcome is defined as whether the student was chronically absent (attended <90 perecent of school days) in third grade.

[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

[g]The sample size refers to the number of students from the full implementation year sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the full implementation year analytic sample, 81 percent have data for math and at least 79 percent have data for all other outcomes in the table.

- **The impact of High 5s, which was implemented in public schools only, on children's third-grade outcomes, over and above the effect of Making Pre-K Count alone, was close to zero and not statistically significant. (See Table 3.5.)**

While High 5s had short-term effects on children's math skills at the end of kindergarten, its effects on children's third-grade math and literacy scores, absenteeism, retention in a grade, and placement in special education, above and beyond the effects of Making Pre-K Count, were close to zero and not statistically significant.

Table 3.5

Impacts of High 5s on Third-Grade Outcomes

Outcome Measure	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size[a]
Math[b]	**-0.04**	**-0.06**	**0.02**	**0.82**	**0.08**	**0.02**
Literacy[c]	-0.03	0.02	-0.05	0.54	0.08	-0.05
Chronic absenteeism (%)[d]	24.3	28.9	-4.5	0.23	3.8	-0.10
Retention (%)[e]	13.8	12.4	1.4	0.62	2.8	0.04
Special education (%)[f]	14.6	14.9	-0.4	0.90	3.0	-0.01
Sample size						
Sites	24	24				
Students[g]	226	230				

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Bolded outcome is confirmatory, all others are exploratory.
Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.
The program group received Making Pre-K Count in pre-K and the High 5s in kindergarten. The control group recieved Making Pre-K Count in pre-K but kindergarten as usual.
Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count in pre-K and High 5s in kindergarten with corresponding outcomes for the control group that did not recieve Making Pre-K Count and had kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.
Rounding may cause slight discrepencies 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]Citywide standardized z-score for state third-grade math test.
[c]Citywide standardized z-score for state third-grade English language arts test.
[d]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.
[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.
[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.
[g]The sample size refers to the number of students from the High 5s sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the High 5s analytic sample, 82 percent have data for math and at least 82 percent have data for all other outcomes in the table.

IMPACT OF TWO YEARS OF EARLY MATH ENRICHMENT

In addition to examining the impacts of each program separately, the Making Pre-K Count and High 5s studies were designed to allow for a test of the effect of two years of math enrichment in pre-K and kindergarten, compared with pre-K and kindergarten as usual. This section presents findings comparing third-grade outcomes for children who received two years of math enrichment (the Making Pre-K Count plus High 5s program group) with third-grade outcomes for children who received no math enrichment (the pre-K-and-kindergarten-as-usual control group). Because this comparison builds off the High 5s program, it includes only children who were eligible for High 5s—that is, children who attended pre-K and kindergarten in the same public school.

- **Two years of enriched math instruction in public schools led to positive and statistically significant effects on children's third-grade math scores (ES = 0.34), the confirmatory outcome for the study, when compared with pre-K and kindergarten as usual.**

The research team observed only small impacts of Making Pre-K Count on children's third-grade outcomes in public schools (for example, ES = 0.16 for math skills in public schools) and found that High 5s had no added benefit. Nevertheless, the effect of two years of early math enrichment on outcomes of students who attended the same public school for pre-K and kindergarten was substantially larger than those estimates combined, when compared with a comparable sample of public school students who received no math enrichment in pre-K or kindergarten. (See Table 3.6.)

Table 3.6

Impacts of Making Pre-K Count and High 5s on Third-Grade Outcomes

Outcome Measure	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size[a]
Math[b]	**-0.08**	**-0.42**	**0.34**	**0.01 ****	**0.14**	**0.34**
Literacy[c]	-0.07	-0.34	0.27	0.02 **	0.11	0.29
Chronic absenteeism (%)[d]	23.4	32.8	-9.5	0.04 **	4.5	-0.20
Retention (%)[e]	14.0	11.0	3.0	0.36	3.3	0.10
Special education (%)[f]	14.3	19.6	-5.3	0.22	4.3	-0.13
Sample size						
Blocks	11	10				
Sites	24	22				
Students[g]	226	255				

(continued)

Table 3.6 (continued)

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Bolded outcome is confirmatory, all others are exploratory.
Statistical significance levels are indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.
The program group received Making Pre-K Count in pre-K and the High 5s in kindergarten. The control group recieved pre-K and kindergarten as usual.
Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count in pre-K and High 5s in kindergarten with corresponding outcomes for the control group that did not recieve Making Pre-K Count and had kindergarten as usual, with an adjustment for selected background characteristics and dummy variables for the random assignment blocks.
Rounding may cause slight discrepencies 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]Citywide standardized z-score for state third-grade math test.
[c]Citywide standardized z-score for state third-grade English language arts test.
[d]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.
[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.
[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.
[g]The sample size refers to the number of students from the two years of math sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the two years of math analytic sample, 82 percent have data for math and at least 83 percent have data for all other outcomes in the table.

- **Two years of early math enrichment also led to a positive and statistically significant impact on literacy test scores (ES = 0.29) and an impact of approximately 9 percentage points on chronic absenteeism (ES = -0.20), both exploratory outcomes.**

The effect of two years of math enrichment on retention in a grade or placement in special education were small and not statistically significant.

The relatively large impact on both math and reading test scores at the end of third grade for children who were offered two years of early math enrichment does not clearly align with Making Pre-K Count's relatively modest impacts and High 5s' lack of a statistically significant impact on children's third-grade outcomes.

As described earlier, the impact of Making Pre-K Count appeared to be largest for students entering pre-K with weaker skills—possibly because they had the most room to grow. The research team examined the sample of children used in the analysis of two years of early math enrichment (students who stayed in the same public school for pre-K and kindergarten) and found that it may have included more low-performing children than the full study sample.

Table 3.7 presents the average third-grade standardized math test score (a score of zero represents the city average) of control group children in six different samples: (1) the full Making Pre-K Count study sample, (2) the sample of children who attended pre-K in community-based sites, (3) the

Table 3.7

Control Group Average Third-Grade Math Scores, by Sample

Sample (Number of Students)	Mean[a]
Full Making Pre-K Count sample control group (899)	-0.12
Full Making Pre-K Count community-based organization sample control group (267)	0.06
Full Making Pre-K Count public school sample control group (632)	-0.20
Making Pre-K Count plus High 5s sample control group (255)	-0.42
Low entering language skills sample control group[b] (154)	-0.55
Low entering self-regulation skills sample control group[c] (136)	-0.40

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: The sample sizes refer to the analytic sample for the study's confirmatory outcome, math.
[a]Citywide standardized z-score for state third-grade math test.
[b]Children's language skills were measured using the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011), administered at pre-K entry in the fall of 2014.
[c]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.

sample of children who attended pre-K in public schools, (4) the sample of children in the analysis of two years of early math enrichment, (5) the sample of children with low language skills, and (6) the sample of children with low self-regulation skills. Children in the control group represent how children in the program group would have behaved or performed had they not received Making Pre-K Count, High 5s, or both.

Table 3.7 further shows that control group children in the full study sample performed 0.12 standard deviations below the citywide average in math at the end of third grade. The students who had attended pre-K in community-based sites scored 0.06 standard deviations above the citywide average in third-grade math, and students who had attended pre-K in public schools scored -0.20 standard deviations below the citywide average in third-grade math—suggesting that children in the public school sample were lower performing than those in the community-based site sample. Across New York City's pre-K systems, public schools tend to serve a population of children from higher-income families. However, the Making Pre-K Count and High 5s pre-Ks study samples may not have reflected this trend because they were drawn specifically from community school districts serving children from families with low incomes.[16]

Importantly, however, the control group students who were eligible for the High 5s and thus two years of early math enrichment (that is, those who attended pre-K and kindergarten in the same public school) scored even lower on standardized math tests than the overall public school control group. The control group students in this sample scored -0.42 standard deviations below the citywide

[16]Reid, Melvin, Kagan, and Brooks-Gunn (2019).

average in third-grade math, which is similar to the third-grade performance of the children in both the low self-regulation skills and low language skills groups.

In other words, it appears that the children who attended pre-K in public schools and stayed in those same schools into kindergarten may have had the lowest test scores and the most room to grow. Therefore, the children assigned to the group that received two years of early math enrichment may have been poised to benefit most from the interventions.

It is possible that the skills of the students in the control group declined over time due to factors unrelated to the intervention. Nevertheless, the research team explored a number of different hypotheses to explain the larger-than-expected impacts for this group of children (presented in Appendix F) and did not find evidence for these alternative explanations. The children and elementary schools in the program and control groups did not differ from each other at baseline (shown in Appendix Tables A.1 and F.1).

4

Conclusions

The Making Pre-K Count and High 5s studies were designed to understand the potential of enhanced early math instruction to produce long-term impacts. To date, few studies have tried to rigorously assess the impact of improving early math skills on later outcomes. Findings from the few prior studies of the long-term effects of enhanced early math instruction have been mixed, with some finding positive and some null effects in the years after prekindergarten (pre-K).[1]

Four years after pre-K, the long-term effect of Making Pre-K Count on math, the main target of the program and the study's confirmatory outcome, was small and positive but not statistically significant in the main confirmatory sample, suggesting that the program was as effective in the long term as existing practice in pre-K classrooms in New York City in 2015, and potentially slightly better. This was true even though all students in the study attended pre-K and, at the time the program was implemented, there was a growing emphasis on early math instruction in New York City schools. Even children in the control group received more math instruction than had been typical in previous studies of early math interventions.

The effect of High 5s on third-grade outcomes was close to zero. Despite the moderate effects on children's math skills found in the year that the program was implemented, the program's effect on these skills was not sustained three years later. High 5s was a supplemental math enrichment program delivered in kindergarten outside of instructional time and designed to align with and extend children's experiences in pre-K. Building off the "sustaining environments" hypothesis, High 5s was intended to sustain children's math enrichment experiences in pre-K into kindergarten.[2] However, it was not expected to align closely with day-to-day classroom content in later grades. It is possible that this lack of connection to classroom instruction in future years was related to the lack of sustained impacts for the program alone.

High 5s also did not lead to impacts on other, non-math outcomes in kindergarten. It is possible that effects in these domains are needed to maintain the early intervention's impacts, as posited by Heckman, Stixrud, and Urzua.[3] Finally, exploratory analyses of Making Pre-K Count suggest that the early math enrichment interventions are most effective for children with the most room to grow. Because High 5s served only children who received Making Pre-K Count in pre-K, all children in the sample already had relatively strong math skills. Children in the High 5s control group (those who had

[1]Dumas, McNeish, Sarama, and Clements (2019); Rittle-Johnson, Fyfe, Hofer, and Farran (2016).
[2]Bailey, Jenkins, and Alvarez-Vargas (2020).
[3]Heckman, Stixrud, and Urzua (2006).

received Making Pre-K Count in pre-K but continued with kindergarten as usual) had third-grade test scores near to the citywide mean (mean = -0.06).

The effect of Making Pre-K Count in pre-K *plus* High 5s in kindergarten on math outcomes, compared with not receiving any math enrichment in pre-K or kindergarten, was moderate and statistically significant. When children received two years of math enrichment, the effects on test results were equivalent to approximately 40 percent of the achievement gap in fourth grade between children from families with low incomes and their peers from families with high incomes.[4] The finding that two years of enrichment (Making Pre-K Count plus High 5s) had moderate effects seems counter-intuitive given the small effects of each of the two interventions separately. This pattern of results is likely due to differences among the samples of children in each analysis. Exploratory subgroup analyses suggest that early math enrichment may have been particularly beneficial for children with the most room to grow, and these children were more prevalent in the combined Making Pre-K Count and High 5s sample.

Making Pre-K Count had fairly large impacts on third-grade math and literacy test scores—ranging from one-quarter to over a third of a standard deviation—for children entering pre-K with the lowest language and self-regulation scores on standardized assessments and ratings. For example, children at sites in the control group (those who received pre-K as usual) who were rated as having higher impulsivity at the start of pre-K by evaluators blinded to the child's treatment status had third-grade math test scores 0.40 standard deviations below the citywide mean. The Making Pre-K Count program raised these third-grade test scores up to the citywide average (-0.04) for a comparable group of children with similar high impulsivity ratings at the start of pre-K. This pattern of long-term effects supports the "academic risk" hypothesis, which posits that early interventions may have the largest effects for children with the most room to grow.[5]

Making Pre-K Count also had small, positive and not statistically significant effects on literacy test scores four years after the pre-K year, suggesting that the program was at least as effective as existing practice in pre-K classrooms in New York City in 2015. These findings align with earlier studies of the Building Blocks curriculum that found it had impacts on aspects of children's language ability and with Duncan and colleagues' correlational study showing that early math skills are a strong predictor of third-grade reading test scores.[6]

Making Pre-K Count alone and the two years of math enrichment also had effects on chronic absenteeism. The early math enrichment reduced chronic absenteeism by about 9 percentage points across public schools and community-based organizations. These effects are substantively meaningful: In the present studies, Making Pre-K Count reduced rates of chronic absenteeism from approximately 33 percent of third-graders in the control group to 24 percent in program group. For younger

[4]Effect sizes in this study are standardized measures of the difference in outcomes at the end of third grade for the control and program groups. To contextualize these impacts, the research team compared the effect sizes with other available standardized data on the difference in achievement between children who are eligible for free or reduced price lunch and those who are not eligible. As described in Hill, Bloom, Black, and Lipsey (2008), using National Assessment of Educational Progress data from 2000 for children at the end of fourth grade, the achievement gap between those eligible for free or reduced price lunch and those not eligible was equivalent to 0.85 standardized units for math. The effect of two years of math enrichment on third-grade math scores (0.34) is equivalent to 40 percent of that difference.

[5]Hamre and Pianta (2005); Hamre and Pianta (2001).

[6]Sarama, Lange, Clements, and Wolfe (2012); Duncan et al. (2007).

students, chronic absenteeism is associated with lower achievement in reading and math, and poor socioemotional outcomes, even after controlling for a wide range of background characteristics.[7] Reducing absenteeism by 9 percentage points for third-graders citywide in New York City could lead to over 7,000 fewer chronically absent third-graders per year.[8]

For young children, attendance is generally thought to be a function of the family's engagement with school. While it is unclear how a teacher-provided instructional program during the school day could change families' attitudes toward school, earlier findings demonstrated that Making Pre-K Count had a positive effect on children's attitudes toward math in kindergarten. Making Pre-K Count may thus have led families to either see school or their children's excitement about school more positively, leading to higher attendance and improved engagement with school. Regardless of the mechanism, chronic absenteeism is negatively associated with later academic achievement.[9] Students who are chronically absent in pre-K and kindergarten have lower test scores by third grade.[10] Chronic absenteeism can be considered a red flag for children at risk of performing poorly in school, and Making Pre-K Count's impacts on this high-risk group of children align with the larger pattern of results showing the largest effects for children with the greatest needs.

The sustained effects on chronic absenteeism, in combination with Making Pre-K Count's earlier impacts on children's executive functioning and attitudes toward math, suggest that early math interventions can have spillover effects into non-cognitive outcomes that are not usually assessed in long-term intervention studies. These spillover effects lend support to the hypothesis posited by Heckman, Stixrud, and Urzua that impacts on non-cognitive outcomes may help maintain longer-term gains in cognitive domains, although the present analyses do not directly test whether effects on non-cognitive outcomes mediate the relationship between the intervention and effects on test scores.[11] Recent findings from a study examining the long-term effects of the Boston pre-K program show that impacts in non-cognitive domains could have potential value for sustaining the effects of pre-K programming into adulthood.[12] The researchers found that the program had short-term effects on student behavior but not on test scores, and ultimately had long-term impacts on high school graduation and college enrollment.

The findings from the Making Pre-K Count and High 5s studies contribute to growing evidence about the importance of early math instruction. They suggest that a well-implemented, evidence-based early math enrichment program has the potential to improve academic achievement over the longer term as well as outcomes in other domains—not just math skills. The effects across a number of domains, including non-cognitive domains, also suggest that early math enrichment could potentially lead to even longer-term effects on students' outcomes as they move into middle and high school.

---

[7]Romero and Lee (2007); Gottfried (2014).
[8]New York City had 78,141 third-graders in 2019-2020 (New York State Education Department, 2020). An analysis by New York University estimated that 22.8 percent of students were chronically absent in 2018 (Research Alliance for New York City Schools, 2019). According to those numbers, an estimated 17,816 third-graders would be chronically absent. After a reduction of chronic absenteeism by 9 percentage points, an estimated 10,784 third-graders would be chronically absent.
[9]Romero and Lee (2007); Ansari and Purtell (2018); Ehrlich, Gwynne, and Allensworth (2018); Simon, Nylund-Gibson, Gottfried, and Mireles-Rios (2020).
[10] Connolly and Olson (2012).
[11]Heckman, Stixrud, and Urzua (2006).
[12]Gray-Lobe, Pathak, and Walters (2021).

Appendix A

Sample Descriptions and Baseline Equivalence of Children Across Program and Control Groups

This appendix describes confirmatory and exploratory samples for the Making Pre-K Count and High 5s studies. It also lays out the baseline equivalence of children for each sample. More information about the sample selection is available in previous reports.[1]

MAKING PRE-K COUNT

The Making Pre-K Count study tested the effects of an evidence-based pre-K math curriculum (Building Blocks), supported by two years of teacher training and in-classroom coaching, that was implemented in 69 pre-K sites in New York City.[2] In the study, whole pre-K sites were randomly assigned to receive the evidence-based math curriculum plus teacher training and coaching (n = 35), or to continue pre-K as usual (n = 34). The effects of Making Pre-K Count alone were estimated on one confirmatory sample (full implementation year sample) and four exploratory samples (full implementation year kindergarten analytic subsample, soft start year sample, soft start year consented subsample, and pooled sample). Each sample is described in greater detail below.

Full Implementation Year Sample

The Making Pre-K Count pre-Ks served children in two distinct years during the time of the study. The first was a soft start year to help teachers become familiar with the curriculum and receive training. Children in the second year of the program's implementation have been the main focus of the Making Pre-K Count study to date and are considered the confirmatory sample for this analysis. Administrative records are available for all children in the full implementation year sample still in the New York City school system. A total of 2,819 children were on the rosters in the 173 classrooms. Of those, 2,702 completed consent forms to participate in the study and were eligible for follow-up assessments.[3] Of those, 2,277 had third-grade data on any outcome, with each outcome varying in the amount of available data. Using the confirmatory outcome, math skills, a Wald test of joint significance indicated that the two groups of children were not systematically different along the available baseline demographic characteristics. (See Table A.1.) Wald tests using the samples for the other four exploratory outcomes (not shown) yielded the same result: that the two groups of children were not systematically different based on demographic characteristics. A test using the full implementation year sample (n = 2,702), including those who did not have third-grade data, yielded similar results.

Full Implementation Year Kindergarten Analytic Subsample

While the main focus of the study is the full implementation year sample, past analyses have only included a subsample of children who were randomly assessed at certain timepoints. To produce aligned estimates across time, the research team included an analysis of the same sample used in the

[1]Morris, Mattera, and Maier (2016); Mattera, Jacob, and Morris (2018).
[2]Clements and Sarama (2013).
[3]Students had to be born before 2010 or at least 4 years old to be eligible to participate. Students in public school pre-K programs were assumed to have met this age requirement in order to be enrolled in pre-K; birth dates for students enrolled in community-based pre-K programs were used to determine age eligibility for Making Pre-K Count. Of the 2,717 participants who consented to participate, 2,702 were age eligible.

Appendix Table A.1

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

Characteristic	Program Group	Control Group
Child demographics		
Race and ethnicity (%)		
Hispanic	52.4	55.6
Non-Hispanic White	6.8	1.5
Non-Hispanic Black	36.1	37.9
Other/multiracial[a]	4.8	5.0
Female (%)	52.3	52.1
Home language (%)[b]		
English	77.1	66.0
Age[c]	4.18	4.19
Parent demographics		
Highest level of education		
At least high school diploma/GED (%)	75.5	71.2
Joint test of difference between groups[d]	(F-value = 0.03)	
Sample size[e]	945	899

SOURCE: MDRC calculations from administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the full implementation year sample.
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.
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 represents the primary language spoken in the child's home.
[c]This is the age at the beginning of pre-K as of September 1, 2014.
[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.
[e]For the parent demographics, n=921 for the program group and n=862 for the control group.

Making Pre-K Count kindergarten impact report, (students who were assessed at the end of kindergarten).[4] Of the 1,325 students included in this analysis, 1,180 had third-grade data available on any outcome, with each outcome varying in the amount of available data. A sensitivity check was conducted to replicate the Making Pre-K Count kindergarten impact analysis with the full implementation year kindergarten analytic subsample from third grade—that is, those children that were still able to be tracked into third grade. Results of the kindergarten impact analysis were similar in magnitude, direction, and statistical significance using this sample. A Wald test of joint significance using the math sample indicated that the two groups of children were not systematically different based on demographics. (See Table A.2.)

[4]Mattera, Jacob, and Morris (2018).

Appendix Table A.2

Baseline Equivalence: Making Pre-K Count Versus Pre-K as Usual for Full Implementation Year Kindergarten Analytic Subsample

Characteristic	Program Group	Control Group
Child demographics		
Race and ethnicity (%)		
Hispanic	54.0	56.6
Non-Hispanic White	6.1	1.8
Non-Hispanic Black	35.0	36.4
Other/multiracial[a]	4.9	5.2
Female (%)	52.1	53.6
Home language (%)[b]		
English	76.2	64.2
Age[c]	4.18	4.17
Parent demographics		
Highest level of education		
At least high school diploma/GED (%)	74.5	68.9
Joint test of difference between groups[d]	(F-value = 0.04)	
Sample size[e]	474	500

SOURCE: MDRC calculations from administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

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 represents the primary language spoken in the child's home.
[c]This is the age at the beginning of pre-K as of September 1, 2014.
[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.
[e]For the parent demographics, n=462 for the program group and n=479 for the control group.

Soft Start Year Consented Subsample

Although the students who were in pre-K during the 2013-2014 soft start year were not the main focus of this study, an effort was made to collect consent from them. Of the estimated 3,120 students in Making Pre-K Count pre-Ks that year, 1,911 (61 percent) consented to participate in the study and have their data tracked over the years. By third grade, 1,520 of these students had third-grade data available on any outcome, with each outcome varying in the amount of available data. A Wald test of joint significance using the math sample indicated that the two groups of children were not systematically different based on demographics. (See Table A.3.)

Appendix Table A.3

Baseline Equivalence: Making Pre-K Count Versus Pre-K as Usual for Soft Start Year Consented Subsample

Characteristic	Program Group	Control Group
Child demographics		
Race and ethnicity (%)		
Hispanic	53.0	52.2
Non-Hispanic White	6.1	1.6
Non-Hispanic Black	37.0	42.0
Other/multiracial[a]	4.0	4.2
Female (%)	50.4	53.3
Home language (%)[b]		
English	69.7	65.4
Age[c]	4.22	4.20
Joint test of difference between groups[d]	(F-value = 0.03)	
Sample size	657	638

SOURCE: MDRC calculations from administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: 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 represents the primary language spoken in the child's home.

[c]This is the age at the beginning of pre-K as of September 1, 2013.

[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.

Soft Start Year Sample

Although the research team was unable to identify the full set of students from the soft start year without consent, deidentified administrative records were available for all students who attended pre-K in a public school implementing Making Pre-K Count that year. In addition to the 1,911 children in the soft start consented subsample, an additional 1,060 students were identified as attending pre-K in a Making Pre-K Count public school at some point in the first year of implementation. This sample therefore includes a larger number of students from the soft start year relative to the consented subsample; however, it does not include non-consenting students from community-based organizations and therefore disproportionately represents public school students. By third grade, 2,393 students in the soft start year sample had third-grade data available on any outcome, with each outcome varying in the amount of available data. A Wald test of joint significance using the math sample indicated that the two groups of children were not systematically different based on demographics. (See Table A.4.)

Appendix Table A.4

Baseline Equivalence: Making Pre-K Count Versus Pre-K as Usual for Soft Start Year Sample

Characteristic	Program Group	Control Group
Child demographics		
Race and ethnicity (%)		
Hispanic	50.1	51.6
Non-Hispanic White	6.9	1.4
Non-Hispanic Black	39.5	42.3
Other/multiracial[a]	3.6	4.7
Female (%)	50.7	52.8
Home language (%)[b]		
English	73.0	66.9
Age[c]	4.21	4.20
Joint test of difference between groups[d]	(F-value = 0.03)	
Sample size	1,007	995

SOURCE: MDRC calculations from administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: 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 represents the primary language spoken in the child's home.
[c]This is the age at the beginning of pre-K as of September 1, 2013.
[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.

Pooled Sample

A final exploratory sample includes all children from the soft start and full implementation years across the 69 Making Pre-K Count sites that were able to be tracked in administrative data (n=5,790). By third grade, 4,670 students had third-grade data available on any outcome, with each outcome varying in the amount of available data. A Wald test of joint significance using the math sample indicated that the two groups of children were not systematically different based on demographics. (See Table A.5.)

Appendix Table A.5

Baseline Equivalence: Making Pre-K Count Versus Pre-K as Usual for Pooled Sample

Characteristic	Program Group	Control Group
Child demographics		
Race and ethnicity (%)		
Hispanic	51.2	53.5
Non-Hispanic White	6.8	1.4
Non-Hispanic Black	37.9	40.2
Other/multiracial[a]	4.2	4.9
Female (%)	51.5	52.4
Home language (%)[b]		
English	75.0	66.5
Age[c]	4.19	4.19
Joint test of difference between groups[d]	(F-value = 0.03)	
Sample size	1,952	1,894

SOURCE: MDRC calculations from administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: 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 represents the primary language spoken in the child's home.

[c]This is the age at the beginning of pre-K as of September 1, 2013, for the soft start year sample and September 1, 2014, for the full implementation year sample.

[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.

HIGH 5s

High 5s was implemented in the year after children were in pre-K. Children who were in the 24 public schools that received Making Pre-K Count and stayed in the same public school were eligible for High 5s. In those Making Pre-K Count program public schools, children were individually randomly assigned within the school to either the High 5s program group in kindergarten (Making Pre-K Count plus High 5s group) or a kindergarten-as-usual group (Making Pre-K Count only group). Of the eligible students, 655 children were randomly assigned, 320 to the High 5s program group and 335 to the kindergarten-as-usual control group. These students make up the High 5s sample. Of the 655 children randomly assigned to the High 5s program group, 556 had available data by third grade, 456 of which had math data available. A Wald test of joint significance using the math sample indicated that the two groups of children were not systematically different along the available baseline characteristics. (See Table A.6.) Unlike the other tests, this test included child baseline skills as comparable characteristics in addition to demographics. A test only using demographic characteristics yielded

Appendix Table A.6

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

Characteristic	Program Group	Control Group
Child demographics		
Race and ethnicity (%)		
Hispanic	50.4	51.3
Non-Hispanic White	8.0	5.7
Non-Hispanic Black	37.2	37.8
Other/multiracial[a]	4.4	5.2
Female (%)	55.8	50.4
Home language (%)[b]		
English	77.9	80.0
Age[c]	4.19	4.18
Parent demographics		
Highest level of education		
At least high school diploma/GED (%)	77.6	73.9
Child skills at the end of pre-K (mean)		
Math		
ECLS-B math score (0-44)[d]	28.25	27.83
Woodcock-Johnson Applied Problems Standard Score[e]	103.99	103.00
Language		
ROWPVT Standard Score[f]	98.00	97.74
Executive function		
Pencil Tap: proportion correct (0-1)[g]	0.79	0.76
Arrows Mixed: proportion correct (0-1)[h]	0.85	0.80
Corsi Blocks forward: number correct[i]	3.06	3.10
PSRA Attention and Inhibition Score (0-3)[j]	2.74	2.63
Joint test of difference between groups[k]	(F-value = 1.16)	
Sample size[l]	226	230

SOURCES: MDRC calculations from administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools and the direct child assessments administered in spring 2015.

NOTES: The program group received Making Pre-K Count 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 represents the primary language spoken in the child's home.

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

(continued)

Table A.6 (continued)

[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 (ROWPVT-4; 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 in which 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. Because this test only includes students without missing data, it only uses a sample of 210 total students.

[l]The child skills variables have less data present than the child demographics. For the variable with the least data available, Arrows Mixed, 123 program students and 102 control students have data.

the same result: that the two groups of children were not systematically different based on demographic characteristics.

MAKING PRE-K COUNT PLUS HIGH 5S

Building from the Making Pre-K Count and High 5s random assignment designs, it is possible to estimate the effects of two years of early math enrichment compared with no enriched math (pre-K and kindergarten as usual). To do so, the following two-stage random assignment design was used:

- In the first stage of random assignment, as part of the Making Pre-K Count study, **public schools** (n = 47) were randomly assigned to either a control group or a group receiving the pre-K intervention, within blocks.

- In the second stage of random assignment, **children** in the pre-K program group (in public schools) in the full implementation year who stayed in the same school for pre-K and kindergarten were individually randomly assigned within schools to either a kindergarten providing math enrichment clubs or a control condition. In other words, children in the program public school sites were randomly assigned to receive High 5s or business-as-usual instruction in kindergarten.

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 for pre-K and kindergarten were randomly assigned early in the fall of the kindergarten school year to either the High 5s program or business-as-usual kindergarten instruction. Those children assigned to receive High 5s in kindergarten make up the program group for the sample (n = 320).

The control sample comprises children from the 23 public schools randomly assigned to the control group in the Making Pre-K Count study who stayed in the same school for pre-K and kindergarten and who were randomly selected for assessment in the kindergarten data collection.[5] Those children make up the pre-K-and-kindergarten-as-usual control group (n = 345). Of the 665 children in the kindergarten analysis, 587 had third-grade data available on any outcome, with each outcome varying in the amount of available data. A Wald test of joint significance using the math sample indicated that the two groups of children were not systematically different along the available baseline demographic characteristics. (See Table A.7.)

Appendix Table A.7

Baseline Equivalence: Making Pre-K Count Plus High 5s Versus Pre-K and Kindergarten as Usual

Characteristic	Program Group	Control Group
Child demographics		
Race and ethnicity (%)		
Hispanic	50.4	57.7
Non-Hispanic White	8.0	1.2
Non-Hispanic Black	37.2	36.1
Other/multiracial[a]	4.4	5.1
Female (%)	55.8	50.2
Home language (%)[b]		
English	77.9	69.0
Age[c]	4.19	4.18
Parent demographics		
Highest level of education		
At least high school diploma/GED (%)	77.6	66.9
Joint test of difference between groups[d]	(F-value = 1.29)	
Sample size	226	255

(continued)

[5]A small number of children in the 24 Making Pre-K Count program public schools who stayed in the same school from pre-K to kindergarten did not consent to participate in High 5s (n = 18). Children in the 23 Making Pre-K Count control public schools did not need to consent to High 5s, and therefore there is no way to match these ‘non-consenters’ in the control schools. To maintain external validity, the 18 ‘non-consenting’ children in the program group are randomly assigned post-hoc. In the kindergarten analysis, robustness checks without the 18 non-consenters included showed similar results to analyses with the 18 non-consenters included.

Appendix Table A.7 (continued)

SOURCE: MDRC calculations from administrative records from the New York City Department of Education, via the Research Alliance of New York City Schools.

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 represents the primary language spoken in the child's home.

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

[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 B

Third-Grade Analytic Models

Appendix B presents the analytic strategy for estimating the effects of Making Pre-K Count, High 5s, and two years of early math enrichment. The analytic strategy for estimating the impacts of Making Pre-K Count and High 5s on children's third-grade outcomes builds on prior analytic decisions made for kindergarten. The analyses for Making Pre-K Count, High 5s, and two years of early math enrichment were preregistered before starting impact analysis.[1]

MAKING PRE-K COUNT

To estimate the effect of one year of math enrichment in pre-K (Making Pre-K Count) on third-grade outcomes, this analysis compares the third-grade outcomes for children who attended the 35 pre-K programs that implemented Making Pre-K Count with outcomes for children who attended the 34 pre-K programs that delivered business-as-usual instruction.

Program impacts were estimated by comparing mean outcomes for the Making Pre-K Count group with corresponding means for the pre-K-as-usual control group, applying a regression adjustment for selected background characteristics and block dummy variables. For Making Pre-K Count, multilevel modeling was used to account for the nested structure of the data, with children nested within pre-K sites, which were nested within random assignment blocks.[2] By third grade, children had dispersed to new classrooms and schools. Although 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, those levels that are associated with random assignment (block and pre-K site) were carried forward.

The analysis across all confirmatory and exploratory samples included a standard set of covariates used 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 were due to random variation in the sample. Covariates for all samples included the following demographic information from administrative records: the student's race or ethnicity, gender, primary language at home, and age. For the full implementation year samples (including the confirmatory sample), additional covariates were available. As in the pre-K and kindergarten analyses, models also included the following covariates: parental education (a dummy variable for whether the parent had a high school diploma or equivalent or a higher degree) and a baseline measure of the child's level of English language proficiency (assessed by pre-LAS), executive function abilities (assessed by Corsi Blocks forward score and Spatial Conflict Arrows task), attention and impulsivity/self-regulation (assessed by the PSRA), and receptive language (ROWPVT). Because not all students in the full implementation year were assessed at baseline, missing baseline assessment data were imputed using multiple imputation.[3]

[1]The preregistered plans can be found at: https://osf.io/bm6va, https://osf.io/ujxnr, and https://osf.io/68yxg.

[2]Impacts on the full implementation year sample were also run including pre-K classrooms as an additional level in the model to replicate the original pre-K analytic model. Results of the third-grade impact analysis were similar in magnitude, direction, and statistical significance using this specification.

[3]As a robustness check, impacts on the full implementation year samples were also run dropping the baseline direct assessment covariates, which had higher levels of missingness than administrative demographic data. The analysis showed similar effects, and magnitude did not change direction substantially.

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

Level 1: Children in pre-K sites

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

Level 2: Sites

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

Where:

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

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

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

T_c = the treatment indicator, which equaled one if site *c* was randomized to treatment (an intervention) and zero if it was randomized to control status,

ε_{sc} = a random error for student *s* in site *c* that was independently and identically distributed across students in classrooms,

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

HIGH 5s

To estimate the effect of math enrichment in kindergarten (High 5s) on third-grade outcomes, this analysis compares the third-grade outcomes for children assigned to two years of math enrichment (Making Pre-K Count plus High 5s) with outcomes for children assigned to one year of math enrichment (Making Pre-K Count).

Program impacts were estimated by comparing mean outcomes for the children assigned to High 5s with corresponding means for the kindergarten-as-usual control group, applying a regression adjustment for selected background characteristics and a dummy variable for each public school. Covariates for this analysis included the following demographic information from administrative records: the student's race or ethnicity, gender, primary language at home, and age. Additional covariates from baseline included the following:[4] parental education (a dummy variable for whether the parent had a high school diploma or equivalent or a higher degree) and assessments from the spring of children's pre-K year, including executive function abilities (inhibition, cognitive flexibility, and working memory), attention and impulsivity/self-regulation, receptive language (ROWPVT), and math ability (ECLS-B and Woodcock Johnson Applied).

The following single-level model was used for third-grade child outcomes:

$$Y_s = \alpha_0 + \sum_{i>0} \alpha_i X_{is} + \Pi T_s + \sum_{c=1}^{24} \alpha_c Z_{cs} + \varepsilon_s$$

Where:

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 equaled 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 was independently and identically distributed across students.

TWO YEARS OF MATH

To estimate the effect of both years of math enrichment in pre-K and kindergarten (Making Pre-K Count and High 5s) on third-grade outcomes, this analysis compares the third-grade outcomes for children assigned to two years of math enrichment (Making Pre-K Count plus High 5s) with outcomes for children assigned to pre-K and kindergarten as usual (control condition).

Program impacts were estimated by comparing mean outcomes for the group of students assigned to Making Pre-K Count and High 5s with corresponding means for students in the pre-K-and-kindergarten-as-usual control group, applying a regression adjustment for selected background characteristics and dummy variables for the random assignment block. Multilevel modeling was used to account for the nested structure of the data, with children nested within pre-K sites, which were nested within random assignment blocks. By third grade, children had dispersed to new classrooms and schools. Although the pre-K site no longer accounted for a large portion of shared variance,

[4]Because randomization for High 5s occurred in kindergarten, the assessment covariates were measured at the end of pre-K, before the High 5s random assignment.

random assignment for this portion of the study occurred at the pre-K site level within random assignment blocks; therefore, those levels that are associated with random assignment (block and pre-K site) were carried forward.

The analysis included a standard set of covariates used 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 include the following demographic information from administrative records: the student's race or ethnicity, gender, primary language at home, and age. As in the kindergarten analyses, models also included the following covariates: parental education (a dummy variable for whether the parent had a high school diploma or equivalent or a higher degree) and a baseline measure of the child's level of English language proficiency (assessed by pre-LAS), executive function abilities (assessed by Corsi Blocks forwards score and Spatial Conflict Arrows task), attention and impulsivity (assessed by the PSRA), and receptive language (ROWPVT). Because not all students were assessed at baseline, missing baseline assessment data are imputed using multiple imputation.

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

Level 1: Children in pre-K sites

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

Level 2: Sites

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

Where:

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

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

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

T_c = the treatment indicator, which equaled one if site *c* was randomized to treatment (an intervention) and zero if it was randomized to control status,

ε_{sc} = a random error for student *s* in site *c* that was independently and identically distributed across students in classrooms,

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

Appendix C

Analyses for Making Pre-K Count for Exploratory Samples

As detailed in the report, the impacts of Making Pre-K Count on the confirmatory sample (full implementation year sample) were positive and not statistically significant for third-grade math and literacy outcomes. There was also a statistically significant reduction in chronic absenteeism and effects close to zero on retention in a grade or placement in special education in third grade for the confirmatory sample.

As described in Appendix A, the Making Pre-K Count study also included a number of exploratory samples—children who were in the schools in the first year of implementation (soft start year sample and soft start year consented sample), a subset of the full implementation year children who were randomly selected for assessment in pre-K (full implementation year kindergarten analytic sample), and the pooled sample of soft start year students and full implementation year students. Impacts were also estimated on the same outcomes for these exploratory samples. Results from the exploratory sample analyses showed the same pattern of effects as for the confirmatory sample. Tables C.1 and C.2 present the impacts of Making Pre-K Count on the confirmatory outcome (math skills) and exploratory outcomes for each of the subsamples.

Table C.1

Impacts of Making Pre-K Count on Third-Grade Math Outcomes, by Sample

Outcome Measure	Number of Children	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size[a]
Math[b]							
Full implementation year	1,844	-0.02	-0.12	0.10	0.19	0.08	0.10
Kindergarten analytic	974	-0.04	-0.20	0.16	0.03 **	0.07	0.16
Soft start year	2,002	0.00	-0.09	0.09	0.12	0.06	0.10
Consented	1,295	0.04	-0.10	0.14	0.03 **	0.06	0.15
Pooled	3,846	-0.01	-0.10	0.08	0.17	0.06	0.09
Sample size							
Blocks		16	16				
Sites		35	34				

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

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.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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.

Table C.2

Impacts of Making Pre-K Count on Third-Grade Literacy, Chronic Absenteeism, Retention, and Special Education Outcomes, by Sample

Outcome Measure	Number of Children	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size[a]
Literacy[b]							
Full implementation year	1,849	0.02	-0.09	0.11	0.12	0.07	0.11
Kindergarten analytic	976	0.04	-0.12	0.16	0.04 **	0.08	0.17
Soft start year	1,998	0.00	-0.09	0.09	0.12	0.06	0.09
Consented	1,292	0.03	-0.09	0.11	0.06 *	0.06	0.12
Pooled	3,847	0.01	-0.08	0.09	0.10 *	0.05	0.09
Chronic absenteeism (%)[c]							
Full implementation year	1,788	23.6	32.6	-9.0	0.00 ***	3.0	-0.19
Kindergarten analytic	927	23.5	30.9	-7.4	0.03 **	3.4	-0.16
Soft start year	1,902	22.1	26.8	-4.7	0.09 *	2.8	-0.11
Consented	1,184	17.1	23.5	-6.4	0.01 **	2.5	-0.15
Pooled	3,690	22.9	29.9	-7.0	0.01 **	2.6	-0.15
Retention (%)[d]							
Full implementation year	2,166	12.5	12.1	0.4	0.84	2.0	0.01
Kindergarten analytic	1,125	12.4	11.7	0.7	0.74	2.2	0.02
Soft start year	2,277	10.8	9.1	1.7	0.33	1.7	0.06
Consented	1,455	9.6	9.4	0.2	0.93	1.7	0.01
Pooled	4,443	11.7	10.6	1.0	0.52	1.6	0.03

(continued)

Appendix Table C.2 (continued)

Outcome Measure	Number of Children	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size[a]
Special education (%)[e]							
Full implementation year	2,276	18.6	20.1	-1.5	0.40	1.7	-0.04
Kindergarten analytic	1,180	17.2	21.0	-3.8	0.15	2.7	-0.09
Soft start year	2,385	16.3	16.6	-0.4	0.85	1.9	-0.01
Consented	1,512	14.4	17.6	-3.1	0.15	2.1	-0.08
Pooled	4,660	17.4	18.4	-1.0	0.48	1.4	-0.03
Sample size							
Blocks		16	16				
Sites		35	34				

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

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.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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]Citywide standardized z-score for state third-grade English language arts test.

[c]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.

[d]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[e]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

Appendix D

Subgroup Analyses for Making Pre-K Count

Chapter 3 of the report describes the impacts of Making Pre-K Count on the confirmatory outcome, third-grade math test scores, across a range of subgroups: for Hispanic and non-Hispanic students, for boys and girls, for students whose primary language at home is English and students whose language at home is another language, and for children entering pre-K with higher and lower relative skills. Appendix D presents the impacts of Making Pre-K Count across the same subgroups on the exploratory outcomes (third-grade literacy test scores, chronic absenteeism, retention in a grade, and placement in special education). As with math skills, Making Pre-K Count's effects on third-grade literacy skills and chronic absenteeism were similar in magnitude across racial or ethnic groups, gender, and primary language status. However, effects on literacy skills were substantively and statistically larger for children entering pre-K with weaker skills than children entering with stronger skills.

Table D.1

Impacts of Making Pre-K Count on Third-Grade Outcomes, by Race/Ethnicity (Hispanic Versus Non-Hispanic)

	Hispanic				Non-Hispanic					
Outcome Measure	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
Math[b]	-0.18	0.14	0.02 **	0.14	-0.03	0.09	0.31	0.09	0.05	0.65
Literacy[c]	-0.16	0.09	0.05 *	0.09	-0.04	0.12	0.11	0.13	-0.04	0.69
Chronic absenteeism (%)[d]	32.6	-8.4	0.00 ***	-0.18	27.2	-3.6	0.39	-0.08	-4.8	0.33
Retention (%)[e]	10.3	1.2	0.57	0.04	12.5	1.4	0.49	0.05	-0.2	0.95
Special education (%)[f]	20.2	-1.0	0.56	-0.02	16.4	-0.8	0.71	-0.02	-0.2	0.93
Sample size										
Blocks	16				16					
Sites	34				34					
Students[g]	1,013				881					

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the pooled sample (students from both the soft start year sample and the full implementation year sample).

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 in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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.

(continued)

Table D.1 (continued)

[b]Citywide standardized z-score for state third-grade math test.

[c]Citywide standardized z-score for state third-grade English language arts test.

[d]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.

[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

[g]The sample size refers to the number of students from the pooled sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the pooled analytic sample, 82 percent have data for math and at least 79 percent have data for all other outcomes in the table.

Table D.2

Impacts of Making Pre-K Count on Third-Grade Outcomes, by Gender (Male Versus Female)

	Male				Female					
Outcome Measure	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
Math[b]	-0.11	0.12	0.09 *	0.12	-0.11	0.07	0.30	0.08	0.05	0.62
Literacy[c]	-0.19	0.12	0.06 *	0.13	0.01	0.07	0.24	0.07	0.05	0.53
Chronic absenteeism (%)[d]	29.1	-4.0	0.24	-0.09	30.1	-9.2	0.00 ***	-0.20	5.1	0.23
Retention (%)[e]	11.9	0.7	0.76	0.02	9.3	1.3	0.45	0.04	-0.7	0.81
Special education (%)[f]	23.6	-0.5	0.80	-0.01	13.6	-1.4	0.38	-0.04	0.8	0.75
Sample size										
Blocks	16				16					
Sites	34				34					
Students[g]	901				993					

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the pooled sample (students from both the soft start year sample and the full implementation year sample).

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 in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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.

(continued)

Table D.2 (continued)

[b]Citywide standardized z-score for state third-grade math test.

[c]Citywide standardized z-score for state third-grade English language arts test.

[d]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.

[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

[g]The sample size refers to the number of students from the pooled sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the pooled analytic sample, 82 percent have data for math and at least 79 percent have data for all other outcomes in the table.

Table D.3

Impacts of Making Pre-K Count on Third-Grade Outcomes, by Home Language (English Versus Non-English)

Outcome Measure	English: Control Group Mean	English: Difference (Impact)	English: P-Value	English: Effect Size[a]	Non-English: Control Group Mean	Non-English: Difference (Impact)	Non-English: P-Value	Non-English: Effect Size[a]	Difference Between Subgroups	P-Value
Math[b]	-0.09	0.08	0.25	0.08	-0.16	0.13	0.04 **	0.14	-0.05	0.62
Literacy[c]	-0.05	0.08	0.18	0.08	-0.19	0.13	0.03 **	0.14	-0.05	0.55
Chronic absenteeism (%)[d]	32.3	-5.9	0.08 *	-0.12	20.5	-6.8	0.02 **	-0.17	0.9	0.84
Retention (%)[e]	11.0	0.9	0.61	0.03	10.2	1.5	0.47	0.05	-0.6	0.83
Special education (%)[f]	17.7	-0.4	0.83	-0.01	18.9	-1.8	0.45	-0.04	1.4	0.63
Sample size										
Blocks	16				16					
Sites	34				33					
Students[g]	1,259				635					

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the pooled sample (students from both the soft start year sample and the full implementation year sample).

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 in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies in sums and differences.

(continued)

Table D.3 (continued)

[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]Citywide standardized z-score for state third-grade math test.

[c]Citywide standardized z-score for state third-grade English language arts test.

[d]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.

[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

[g]The sample size refers to the number of students from the pooled sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the pooled analytic sample, 82 percent have data for math and at least 79 percent have data for all other outcomes in the table.

Table D.4

Impacts of Making Pre-K Count on Third-Grade Outcomes, by Entering Skill Level (Language and Self-Regulation)

	Low Skill Level				High Skill Level					
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
Math[b]										
Language (ROWPVT)[c]	-0.55	0.28	0.08 *	0.28	0.08	0.14	0.27	0.14	0.14	0.48
Self-regulation[d]	-0.40	0.37	0.01 ***	0.36	-0.06	0.14	0.28	0.14	0.22	0.24
Literacy[e]										
Language (ROWPVT)[c]	-0.52	0.25	0.06 *	0.27	0.12	0.15	0.23	0.15	0.10	0.58
Self-regulation[d]	-0.48	0.44	0.00 ***	0.43	0.07	0.08	0.55	0.08	0.37	0.04 ††
Chronic absenteeism (%)[f]										
Language (ROWPVT)[c]	27.4	-2.3	0.69	-0.05	31.6	-8.3	0.16	-0.18	6.0	0.48
Self-regulation[d]	32.3	-8.3	0.19	-0.18	25.0	2.7	0.65	0.06	-11.0	0.20
Retention (%)[g]										
Language (ROWPVT)[c]	17.8	0.0	0.99	0.00	10.4	-1.3	0.69	-0.04	1.3	0.82
Self-regulation[d]	19.0	-2.5	0.60	-0.06	8.4	1.5	0.65	0.06	-4.0	0.49
Special education (%)[h]										
Language (ROWPVT)[c]	24.7	-3.2	0.48	-0.07	15.2	-4.0	0.25	-0.12	0.8	0.89
Self-regulation[d]	25.1	-8.9	0.05 **	-0.21	15.7	-1.9	0.62	-0.05	-7.0	0.23

(continued)

Table D.4 (continued)

Outcome Score	Low Skill Level: Control Group Mean	Low Skill Level: Difference (Impact)	Low Skill Level: P-Value	Low Skill Level: Effect Size[a]	High Skill Level: Control Group Mean	High Skill Level: Difference (Impact)	High Skill Level: P-Value	High Skill Level: Effect Size[a]	Difference Between Subgroups	P-Value
Sample size										
Blocks										
Language (ROWPVT)[c]	16				16					
Self-regulation[d]	16				16					
Sites[i]										
Language (ROWPVT)[c]	33				34					
Self-regulation[d]	33				34					
Students[j]										
Language (ROWPVT)[c]	154				140					
Self-regulation[d]	136				159					

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

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 in pre-K. The control group did not receive math enrichment and participated in pre-K as usual.

Rounding may cause slight discrepencies 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]Citywide standardized z-score for state third-grade math test.

[c]Children's language skills were measured using the Receptive One-Word Picture Vocabulary Test (ROWPVT-4; Martin and Brownell, 2011), administered at pre-K entry in the fall of 2014.

[d]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.

[e]Citywide standardized z-score for state third-grade English language arts test.

[f]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.

[g]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[h]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

[i]The number of control sites ranges from 32 to 34 across outcomes.

[j]The sample size refers to the analytic sample for the study's confirmatory outcome, math skills.

Appendix E

Impacts on Chronic Absenteeism Across Grades

As described in Chapter 3, the impacts of Making Pre-K Count on chronic absenteeism in third grade were estimated for the confirmatory Making Pre-K Count sample. Making Pre-K Count had a favorable and consistent impact on children's chronic absenteeism in third grade across all Making Pre-K Count samples. This appendix presents the impacts of Making Pre-K Count on chronic absenteeism for children in the full implementation year sample from kindergarten through third grade. (See Table E.1.) The analytic sample presented in Table E.1 includes only children in the confirmatory third-grade sample, across each school year.[1]

The effect of Making Pre-K Count on chronic absenteeism began early and continued as children moved through elementary school. Making Pre-K Count had a favorable and statistically significant effect on chronic absenteeism for children in first, second, and third grade. Effect sizes ranged from -0.15 to -0.19. There was a favorable but not statistically significant effect in the kindergarten year (ES = -0.09).

[1]Sensitivity analyses including all available data for all full implementation year sample children at each timepoint show the same pattern of magnitude and statistical significance of the effects.

Table E.1

Impacts of Making Pre-K Count on Chronic Absenteeism in Kindergarten Through Third Grade

Outcome Measure	Number of Children	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size[a]
Chronic absenteeism (%)[b]							
Kindergarten	1,877	30.0	34.1	-4.1	0.27	3.7	-0.09
First grade	1,819	23.6	31.4	-7.8	0.02 **	3.3	-0.17
Second grade	1,790	23.0	29.8	-6.8	0.01 ***	2.6	-0.15
Third grade	1,788	23.6	32.6	-9.0	0.00 ***	3.0	-0.19
Sites		35	34				

SOURCE: MDRC calculations are based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

NOTES: Calculations are made using students from the full implementation year sample.

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.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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 outcome is defined as whether the student was chronically absent (attended <90 percent of school days) during the year.

Appendix F

Further Exploratory Analyses

The relatively large impact of two years of early math enrichment on both math and literacy scores at the end of third grade was surprising given Making Pre-K Count's relatively modest impacts for the confirmatory Making Pre-K Count sample and High 5's lack of a statistically significant impact in the third grade. Chapter 3 presents the analyses for the main hypothesis about why impacts were larger for this group—namely, that the sample eligible for two years of early math enrichment had low third-grade math scores in the absence of the intervention and therefore more room for their math skills to improve and the intervention to make a difference.

This appendix further describes exploratory analyses related to potential hypotheses about what could have contributed to the pattern of effects.

- **Were the children in the two groups (two years of early math enrichment group and no early math enrichment group) different at baseline?**

There is limited evidence of baseline differences. A Wald test of joint significance indicated that the two groups of children were not systematically different along the available baseline demographic characteristics, and the proportion of children who stayed in the same school for pre-K and kindergarten were roughly the same. (See Appendix Table A.7.)

- **Were the schools in these two groups different at baseline?**

There is limited evidence of baseline differences. The program public school sites (n = 24) and the control public school sites (n = 23) were similar at baseline in 2013.[1] (See left panel of Appendix Table F.1.) The schools had similar test scores for third-graders in 2013 and served demographically similar populations. Third-grade students in both the program and control schools scored approximately -0.30 standard deviations below the citywide mean in mathematics in the spring of 2013. Third-grade literacy scores for the program group schools were slightly higher than for the control schools in 2013, but not statistically significantly so.

The slight difference in baseline third-grade literacy scores was not enough to explain the large difference observed between the students in these two groups in third grade. Even after controlling for baseline third-grade test scores, the impacts of two years of early math enrichment remained similar. (See Appendix Table F.2.)

- **Were the schools in these two groups different by the time children who received Making Pre-K Count were in third grade?**

By 2019 (the year students who received Making Pre-K Count reached third grade), third-grade students in the Making Pre-K Count program sites were performing better than their counterparts in the Making Pre-K Count control sites. (See right panel of Appendix Table F.1.) In 2019, third-grade students in the Making Pre-K Count program public school sites scored 0.33 standard deviations below the citywide average in mathematics (and 0.31 standard deviations below in literacy), while third-grade students in the Making Pre-K Count control schools scored 0.50 standard deviations below the citywide average in mathematics (and 0.43 standard deviations below in literacy). Analyses

[1]Random assignment of sites occurred in spring 2013.

did not reveal differential demographic shifts in the population of students served by these schools over this time period that might account for these differences. (See Appendix Table F.1.)

Table F.1

Third-Grade Average Characteristics in 2013 and 2019 for Making Pre-K Count Schools, by Random Assignment Group

	2013		2019	
School Characteristic	Program Mean	Control Mean	Program Mean	Control Mean
Average third-grade test score				
Math[a]	-0.31	-0.32	-0.33	-0.50
Literacy[b]	-0.23	-0.28	-0.31	-0.43
Demographics (%)				
Race and ethnicity				
Hispanic	49.5	51.7	50.4	53.5
Non-Hispanic white	4.1	1.3	5.1	2.1
Non-Hispanic black	44.2	43.5	41.2	39.9
Asian	1.5	2.5	1.9	3.0
Other/multiracial[c]	0.8	0.9	1.4	1.5
Female	48.9	49.4	48.5	49.6
English language learners	15.6	15.4	14.0	16.3
Students with disabilities[d]	16.5	18.0	23.1	23.7
Students living in poverty[e]	94.2	96.8	88.0	90.0
Sample size				
Blocks	11	11	11	11
Sites	24	23	24	23
Students	2,490	2,093	2,027	1,701

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education.

NOTES: The program group comprises 24 schools randomly assigned to the Making Pre-K Count program. The control group comprises 23 schools randomly assigned to pre-K-as-usual.

2013 corresponds to the 2012-2013 academic school year, the year before any Making Pre-K Count implementation.

2019 corresponds to the 2018-2019 academic school year, the year that the full implementation year sample was in third grade.

[a]Citywide standardized z-score for state third-grade math test.

[b]Citywide standardized z-score for state third-grade English language arts test.

[c]Other includes students who did not report their race or who reported as Native American.

[d]Students in poverty includes students with families who qualified for free or reduced price lunch or were eligible for Human Resources Administration benefits.

[e]Students with disabilities includes students who had an Individualized Education Program (IEP).

Table F.2

Impacts of Making Pre-K Count and High 5s on Third-Grade Outcomes (Controlling for School's Baseline Performance)

Outcome Measure	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Standard Error	Effect Size[a]
Math[b]	-0.07	-0.44	0.37	0.01	***	0.13	0.37
Literacy[c]	-0.06	-0.37	0.31	0.00	***	0.11	0.32
Chronic absenteeism (%)[d]	22.6	33.9	-11.3	0.01	***	4.2	-0.24
Retention (%)[e]	14.0	11.1	2.8	0.38		3.2	0.09
Special education (%)[f]	14.2	19.9	-5.6	0.20		4.4	-0.14
Sample size							
Blocks	11	10					
Sites	24	22					
Students[g]	226	255					

SOURCE: MDRC calculations based on administrative records from the New York City Department of Education, via the Research Alliance for New York City Schools.

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.

Impacts were estimated by comparing third-grade outcomes for the group assigned to Making Pre-K Count 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.

Rounding may cause slight discrepencies 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]Citywide standardized z-score for state third-grade math test.

[c]Citywide standardized z-score for state third-grade English language arts test.

[d]The outcome is defined as whether the student was chronically absent (attended <90 percent of school days) in third grade.

[e]The outcome is defined as whether the student was below grade level in third grade. It excludes students who do not have a valid grade due to enrollment in self-contained special education classrooms.

[f]The outcome is defined as whether the student had an Individualized Education Program (IEP) in third grade.

[g]The sample size refers to the number of students from the two years of math sample for which test score data were available for math, the study's confirmatory outcome. The analytic sample refers to students with any outcome data. For the two years of math analytic sample, 82 percent have data for math and at least 83 percent have data for all other outcomes in the table.

References

Allensworth, Elaine M., and John Q. Easton. 2007. *What Matters for Staying On-Track and Graduating in Chicago Public High Schools: A Close Look at Course Grades, Failures, and Attendance in the Freshman Year*. Chicago: Consortium on Chicago School Research at the University of Chicago.

Ansari, Arya, and Kelly M. Purtell. 2018. "Absenteeism in Head Start and Children's Academic Learning." *Child Development* 89, 4: 1,088–1,098.

Bailey, Drew H., Jade M. Jenkins, and Daniela Alvarez-Vargas. 2020. "Complementarities Between Early Educational Intervention and Later Educational Quality? A Systematic Review of the Sustaining Environments Hypothesis." *Developmental Review* 56: Article 100910.

Baltimore Education Research Consortium. 2011. *Destination Graduation: Sixth Grade Early Warning Indicators for Baltimore City Schools. Their Prevalence and Impact*. Baltimore: Baltimore Education Research Consortium.

Blair, Clancy, Hilary Knipe, and David Gamson. 2008. "Is There a Role for Executive Functions in the Development of Mathematics Ability?" *Mind, Brain, and Education* 2, 2: 80–89.

Civil Rights Data Collection. 2016. *2015-16 Civil Rights Data Collection*. Washington, DC: U.S. Department of Education, Office of Civil Rights.

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

Clements, Douglas H., Julie Sarama, and Carrie Germeroth. 2016. "Learning Executive Function and Early Mathematics: Directions of Casual Relations." *Early Childhood Research Quarterly* 36: 79–90.

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.

Connolly, Faith, and Linda S. Olson. 2012. *Early Elementary Performance and Attendance in Baltimore City Schools' Pre-Kindergarten and Kindergarten*. Baltimore: Baltimore Education Research Consortium.

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

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *The American Economic Review* 85, 3: 341–364.

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1, 3: 111–134.

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

Dumas, Denis, Daniel McNeish, Julie Sarama, and Douglas Clements. 2019. "Preschool Mathematics Intervention Can Significantly Improve Student Learning Trajectories Through Elementary School." *AERA Open* 5, 4: 1–15.

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: 1,428–1,446.

Duncan, Greg J., and Katherine Magnuson. 2011. "The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems." Pages 47–69 in Greg J. Duncan and Richard J. Murnane (eds.), *Whither Opportunity: Rising Inequality, Schools, and Children's Life Chances*. New York: Russel Sage Foundation.

Duncan, Greg J., and Katherine Magnuson. 2013. "Investing in Preschool Programs." *Journal of Economic Perspectives* 27, 2: 109–132.

Ehrlich, Stacy B., Julia A. Gwynne, and Elaine M. Allensworth. 2018. "Pre-Kindergarten Attendance Matters: Early Chronic Absence Patterns and Relationships to Learning Outcomes." *Early Childhood Research Quarterly* 44, 4: 136–151.

Engel, Mimi, Amy Claessens, and Maida A. Finch. 2013. "Teaching Students What They Already Know? The (Mis)Alignment Between Mathematics Instructional Content and Student Knowledge in Kindergarten." *Educational Evaluation and Policy Analysis* 35, 2: 157–178.

Engel, Mimi, Amy Claessens, Tyler Watts, and George Farkas. 2015. "The Misalignment of Kindergarten Mathematics Content." Paper presented at the annual meeting of the American Education Research Association, Chicago.

Faria, Anne-Marie, Nicholas Sorensen, Jessica Heppen, Jill Bowdon, Suzanne Taylor, Ryan Eisner, and Shandu Foster. 2017. *Getting Students on Track for Graduation: Impacts of the Early Warning Intervention and Monitoring System After One Year*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *The American Economic Review* 92, 4: 999–1,012.

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: 1–23.

Gottfried, Michael A. 2014. "Chronic Absenteeism and Its Effects on Students' Academic and Socioemotional Outcomes." *Journal of Education for Students Placed at Risk* 19, 2: 53–75.

Gray-Lobe, Guthrie, Parag A. Pathak, and Christopher R. Walters. 2021. "The Long-Term Effects of Universal Preschool in Boston." Working Paper 28756. Cambridge, MA: National Bureau of Economic Research.

Hamre, Bridget K., and Robert C. Pianta. 2001. "Early Teacher-Child Relationships and the Trajectory of Children's School Outcomes through Eighth Grade" *Child Development* 72, 2: 625–638

Hamre, Bridget K., and Robert C. Pianta. 2005. "Can Instructional and Emotional Support in the First-Grade Classroom Make a Difference for Children at Risk of School Failure?" *Child Development* 76, 5: 949–967.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24, 3: 411–482.

Hernandez, Donald J. 2011. *Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation*. Baltimore: The Annie E. Casey Foundation.

Hill, Carolyn J., Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey. 2008. "Empirical Benchmarks for Interpreting Effect Sizes in Research." *Child Development Perspectives* 2, 3: 172–177.

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.

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

Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *The Quarterly Journal of Economics* 122, 1: 159–208.

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 Edition*. Novato, CA: Academic Therapy Publications.

Mattera, Shira K., Robin Jacob, and Pamela A. Morris. 2018. *Strengthening Children's Math Skills with Enhanced Instruction: The Impacts of Making Pre-K Count and High 5s on Kindergarten Outcomes*. New York: MDRC.

McCormick, Meghan P., Robin Neuhaus, Erin E. O'Connor, Hope I. White, E. Parham Horn, Samantha Harding, Elise Cappella, and Sandee McClowry. 2021. "Long-Term Effects of Social-Emotional Learning on Academic Skills: Evidence from a Randomized Trial of *INSIGHTS*." *Journal of Research on Educational Effectiveness* 14, 1: 1–27.

Morris, Pamela, Shira K. Mattera, Nina Castells, Michael Bangser, Karen Bierman, and Cybele Raver. 2014. *Impact Findings from the Head Start CARES Demonstration: National Evaluation of Three Approaches to Improving Preschoolers' Social and Emotional Competence. OPRE Report 2014-44*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

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.

New York State Education Department. 2020. "NYC Public Schools at a Glance." Website: data.nysed.gov/profile.php?instid=7889678368.

Reid, Jeanne L., Samantha A. Melvin, Sharon Lynn Kagan, and Jeanne Brooks-Gunn. 2019. "Building a Unified System for Universal Pre-K: The Case of New York City." *Children and Youth Services Review* 100, C: 191–205.

Research Alliance for New York City Schools. 2019. “How Has Attendance in NYC Schools Changed Over Time?” Website: www.steinhardt.nyu.edu/research-alliance/research/spotlight-nyc-schools/how-has-attendance-nyc-schools-changed-over-time

Rittle-Johnson, Bethany, Emily R. Fyfe, Kerry G. Hofer, and Dale C. Farran. 2016. “Early Math Trajectories: Low-Income Children’s Mathematics Knowledge from Ages 4 to 11.” *Child Development* 88, 5: 1,727–1,742.

Romero, Mariajosé, and Young-Sun Lee. 2007. *A National Portrait of Chronic Absenteeism in the Early Grades*. New York: National Center for Children in Poverty.

Rose, Stephanie, and Karen Schimke. 2012. *Third Grade Literacy Policies: Identification, Intervention, Retention.* Denver: Education Commission of the States (NJ3).

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, Alissa A. Lange, Douglas H. Clements, and Christopher B. Wolfe. 2012. “The Impacts of an Early Mathematics Curriculum on Oral Language and Literacy.” *Early Childhood Research Quarterly* 27, 3: 489–502.

Schweinhart, Lawrence J. 2013. “Long-Term Follow-Up of a Preschool Experiment.” *Journal of Experimental Criminology* 9, 4: 389–409.

Simon, Odelia, Karen Nylund-Gibson, Michael Gottfried, and Rebeca Mireles-Rios. 2020. “Elementary Absenteeism Over Time: A Latent Class Growth Analysis Predicting Fifth and Eighth Grade Outcomes.” *Learning and Individual Differences* 78, 2: Article 101822.

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.

Snow, Catherine E., Susan M. Burns, and Peg Griffin. 1998. *Preventing Reading Difficulties in Young Children.* Washington, DC: National Academy Press.

U.S. Department of Education. 2019. *Chronic Absenteeism in the Nation’s Schools: A Hidden Educational Crisis*. Washington, DC: U.S. Department of Education.

Watts, Tyler W., Greg J. Duncan, Douglas H. Clements, and Julie Sarama. 2018. “What Is the Long-Run Impact of Learning Mathematics During Preschool?” *Child Development* 89, 2: 539–555.

What Works Clearinghouse. 2020. *What Works Clearinghouse Standards Handbook, Version 4.1*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

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: 418–431.

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

About MDRC

MDRC, a nonprofit, nonpartisan social and education policy research organization, is committed to finding solutions to some of the most difficult problems facing the nation. We aim to reduce poverty and bolster economic mobility; improve early child development, public education, and pathways from high school to college completion and careers; and reduce inequities in the criminal justice system. Our partners include public agencies and school systems, nonprofit and community-based organizations, private philanthropies, and others who are creating opportunity for individuals, families, and communities.

Founded in 1974, MDRC builds and applies evidence about changes in policy and practice that can improve the well-being of people who are economically disadvantaged. In service of this goal, we work alongside our programmatic partners and the people they serve to identify and design more effective and equitable approaches. We work with them to strengthen the impact of those approaches. And we work with them to evaluate policies or practices using the highest research standards. Our staff members have an unusual combination of research and organizational experience, with expertise in the latest qualitative and quantitative research methods, data science, behavioral science, culturally responsive practices, and collaborative design and program improvement processes. To disseminate what we learn, we actively engage with policymakers, practitioners, public and private funders, and others to apply the best evidence available to the decisions they are making.

MDRC works in almost every state and all the nation's largest cities, with offices in New York City; Oakland, California; Washington, DC; and Los Angeles.