Making Preschool More Productive

HOW CLASSROOM MANAGEMENT TRAINING CAN HELP TEACHERS

Pamela Morris
C. Cybele Raver
Megan Millenky
Stephanie Jones
Chrishana M. Lloyd

NOVEMBER 2010

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY

The Foundations of Learning Demonstration

Making Preschool More Productive

How Classroom Management Training Can Help Teachers

Pamela Morris (MDRC/New York University)
C. Cybele Raver (New York University)
Megan Millenky (MDRC)
Stephanie Jones (Harvard University)
Chrishana M. Lloyd (MDRC)

November 2010

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY

Funders of the Foundations of Learning Demonstration

The George Gund Foundation
The Grable Foundation
The John D. and Catherine T. MacArthur Foundation
The Joyce Foundation
The Kresge Foundation
McCormick Foundation
The Nicholson Foundation
The Pew Charitable Trusts
Robert Wood Johnson Foundation

Dissemination of MDRC publications is supported by the following funders that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Ambrose Monell Foundation, The Annie E. Casey Foundation, Carnegie Corporation of New York, The Kresge Foundation, Sandler Foundation, and The Starr Foundation.

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

The findings and conclusions in this report do not necessarily represent the official positions or policies of the funders.

Cover photo courtesy of the John D. and Catherine T. MacArthur Foundation. Photographer: Alex Fledderjohn.

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

Overview

Educators and researchers increasingly recognize that high-quality early childhood programs are an important way to prepare disadvantaged children for later school success. One critical (and often overlooked) aspect of quality is addressing children's ability to engage positively with peers and teachers and to focus their attention and behavior during classroom activities. Evidence suggests that improving young children's emotional and behavioral adjustment is both an important outcome in its own right and can be a pathway to improved academic achievement for low- and high-risk children alike. In addition, challenging behavior may divert teachers' attention from instructional time for all children in the classroom. This is not a peripheral problem in preschool classrooms, where a sizable minority of children have behavioral challenges. Preschool teachers often discuss the need for additional training in how to address these issues.

This report presents results from the Newark, New Jersey, site of the Foundations of Learning (FOL) Demonstration — an intervention and random assignment evaluation of a program aimed at equipping teachers with the skills and strategies they need to help guide children's behavior and emotional development. The FOL intervention was tested in two cities — Newark and Chicago — and it combined teacher training in effective classroom management with weekly classroom consultation. In the Newark site, 51 preschool classrooms (one per center) serving primarily 4-year-old children were selected to participate in the study; 26 classrooms were randomly assigned to implement the FOL intervention, and 25 were assigned to conduct preschool as usual. Differences between the two groups were analyzed at the end of the intervention year and the following year to assess the added value of FOL over and above standard practice in preschool classrooms.

Key Findings

The evidence emerging from the Newark site shows that investments in teachers' professional development improve children's preschool experiences, although the long-term effects on children remain uncertain.

- FOL improved teachers' ability to address children's behavior and to provide a positive emotional climate in the classroom. It also improved teachers' management of classroom time, their use of engaging teaching methods, and the amount of instructional time that children experienced in their classrooms.

- Based on ratings by independent trained observers, FOL reduced children's conflicts with teachers and peers and increased their levels of engagement in the learning tasks of preschool, but it did not otherwise change the quality of teacher-child or peer interactions.

- Based on limited data for the year following the intervention, very few of the intervention's effects on children were sustained as they entered kindergarten classrooms. However, teachers who were trained in the intervention appear to continue to engage in the positive practices they learned.

Additional publications on this intervention will be released in the coming years. This will complement information emerging on other promising social-emotional interventions currently being tested in preschool classrooms, providing extensive information to policymakers and practitioners about where to put their attention in efforts to improve preschool quality.

Contents

List of Tables, Figures, and Boxes

Table

Table

Figure

Box

Preface

The Foundations of Learning (FOL) project comes at a particularly important time in the field of early education and child care. Public support for preschool programming is high and, in its wake, has brought a marked increase in the number of children served by early childhood programs. With these programs comes the promise of solving one of the nation's most persistent social problems: the growing achievement gap between poor children and their more affluent peers. Yet, without good information about how to boost the quality of preschool programs, delivering on that promise is a challenge.

FOL is also important because of its explicit focus on children's social-emotional development as a primary target of the intervention model. Driven in part by the passage of the No Child Left Behind Act of 2001, program administrators' attention has been increasingly focused on building children's academic readiness before their entry into formal education. At the same time, teachers repeatedly express a need for effective strategies to address children's emotional and behavioral problems, which they feel ill equipped to address. Such findings suggest that strategies to improve children's behavioral adjustment are a necessary complement to the ongoing work on strategies to improve children's literacy.

This report shares impact results of MDRC's Foundations of Learning Demonstration in Newark, New Jersey. The program model that was tested in this site provides intensive training in classroom management skills for lead and assistant teachers, supported by weekly in-class support from a master's-level clinician to reinforce the lessons from the training and to provide direct services to children. The goal of the FOL evaluation is to test the effectiveness of this model of professional development and clinical consultation in order to provide the underpinning for high-quality preschool education and child care.

The positive but still short-lived results presented here are part of a larger body of work to help understand ways to maintain and improve a quality preschool experience for children. In coming years, MDRC will publish further work on the Chicago site of the FOL demonstration, including a benefit-cost analysis. In addition, MDRC is conducting a large-scale trial of three different social-emotional enhancements in the context of Head Start programs nationally, as part of the Head Start CARES project conducted by the Administration for Children and Families. These studies, along with similar evaluations by other researchers, will continue to build the body of evidence on the most effective ways that preschool programs can support children's social-emotional development as part of an effort to improve their school readiness.

Gordon Berlin
President

Acknowledgments

This report on the Newark, New Jersey, site of the Foundations of Learning (FOL) demonstration benefited from a collaboration among many organizations and individuals. From the outset, the support of the Newark Public Schools has been instrumental in planning and conducting the project. We particularly thank Dr. Gayle Griffin, Assistant Superintendent; Nancy Rivera, Director of Early Childhood Education; Shirley Grundy, Director of Guidance at the time the study was conducted (now retired); Dr. Marbella Barerra, Institutional Review Board Director; and Kathleen Tague and Patricia DeMarco, supervisors in the Office of Early Childhood; as well as the Preschool Intervention and Referral Team members and resource teachers. Babu Yalamanchili, Marisol Peña, Laurie Newell, John Duggan, and Carol Little were all helpful in providing data.

Carrying out a research demonstration of this type is possible only with the commitment and cooperation of the participating programs. We thank all the teachers and administrators at the schools, Head Start centers, and community-based programs that were part of the demonstration. Clark Thompson and Ernestine Simpson of the Newark Preschool Council were especially helpful during both the pilot phase and the full demonstration period.

At Family Connections — the subcontractor responsible for implementing the classroom consultation component of the FOL program — we thank Paula Sabreen, Executive Director; Sheila Berard, Associate Director; Dorothy Jordan, Clinical Classroom Consultant Coordinator; and each of the Clinical Classroom Consultants who worked so diligently to implement a high-quality program.

Greg O'Donnell, Darlene Jones-Lewis, Kimya Barden, and Dorothy Jordan conducted the teacher training sessions. The University of Virginia's Center for Advanced Study of Teaching and Learning and Megan Siebert trained coders for classroom observations; a team of dedicated coders visited Newark classrooms throughout the demonstration.

Survey Research Management, led by Linda Kuhn, fielded surveys throughout the demonstration and located students for follow-up data collection.

A number of foundation funders provided indispensable support for the evaluation. They are gratefully acknowledged at the front of the report.

The FOL research effort has been a true partnership. At MDRC, we thank the following key members of the team: Mike Bangser, Francesca Longo, Ximena Portilla, Vivian Mateo, and Farrah Parkes. Shirley James and her team were responsible for keying and verifying the data. Glee Holton and Shelley Rappaport helped recruit preschool programs for the demonstration.

Helen Gorden provided administrative support. Karen Paget answered our Institutional Review Board questions regarding data collection. Gordon Berlin, Howard Bloom, Robert Ivry, Ginger Knox, and John Hutchins provided valuable comments during the drafting process. Robert Weber edited the report, and Vivian Mateo and David Sobel prepared it for publication.

Outside MDRC, Christine Li-Grining of Loyola University and Fuhua Zhai of New York University provided valuable guidance and input by drawing on their experiences with CSRP (formerly the Chicago School Readiness Project).[1] Karen McFadden, a doctoral fellow from New York University, helped to compile and analyze data.

Most of all, we thank the children in the FOL program and their parents. We hope that the lessons from this and future reports will help strengthen the quality of preschool programs in Newark and elsewhere.

The Authors

[1]CSRP is not associated with The Chicago School®, which is a trademark of The Chicago School of Professional Psychology.

Executive Summary

Even as policymakers embrace greater investments in early childhood programs, questions remain about how to ensure that preschools are of high enough quality to promote children's development. One critical (and often overlooked) aspect of quality is the capacity to address children's emotional and behavioral adjustment — that is, their ability to engage positively with peers and teachers and to focus their attention and behavior during classroom activities.

Evidence suggests that improving young children's emotional and behavioral adjustment is both an important outcome in its own right and can be a pathway to improved academic achievement for low- and high-risk children alike. Not only are preschoolers who have behavioral challenges more likely to face long-term difficulties throughout their school careers, but their behavior may divert teachers' attention from instructional time for all children in the classroom. Studies have shown that as many as 3 to 4 children in every preschool class of 15 to 20 present behavioral challenges. Unfortunately, preschool teachers generally receive very little training about how to address these issues.

This report presents results from the Newark, New Jersey, site of the Foundations of Learning (FOL) Demonstration, an intervention and random assignment evaluation of a curriculum designed to target children's behavior and emotional adjustment through the training of preschool teachers. The demonstration adopted the model used by CSRP (formerly the Chicago School Readiness Project) and adapted it slightly to fit a new policy context.[1] The remainder of this report refers to the model as "FOL." The FOL intervention was tested in two cities — Newark and Chicago — and it combined teacher training in effective classroom management with weekly classroom consultation. Consultants coached and mentored the teachers in the new strategies learned in the training workshops, and they provided individualized support to the highest-risk children in each FOL classroom. MDRC evaluated the results of this intervention. The findings for the preschool year indicate that FOL had promising effects on classroom quality, teachers' productive use of classroom time, and some outcomes for children. Yet these effects were not sustained, particularly for high-risk children, as they transitioned to kindergarten.

[1]CSRP is not associated with The Chicago School®, which is a trademark of The Chicago School of Professional Psychology.

The Foundations of Learning Model

The FOL intervention addresses children's challenging behavior by training preschool teachers to proactively support children's positive behavior while more effectively limiting and redirecting their aggressive and disruptive behavior.[2] The model was initially developed in the context of an earlier trial led by Dr. Cybele Raver,[3] who is now a member of the FOL research team.

The intervention includes four components delivered across the school year:

- **Teacher training.** Lead and assistant teachers are invited to attend five Saturday training sessions. The workshops are an adapted version of *The Incredible Years* curriculum developed by Dr. Carolyn Webster-Stratton.[4] The workshops provide instruction on how to develop positive relationships with children; present classroom strategies that teachers can use, such as setting clear rules; and provide teachers with techniques to develop children's social skills, anger management, and problem-solving ability.

- **Classroom-level consultation.** To complement the training, teachers are assigned a master's-level Clinical Classroom Consultant (CCC) to work with them in the classroom one day per week throughout the school year. The CCCs model and reinforce the content of the training sessions.

- **Stress management.** In winter, teachers participate in a 90-minute stress management workshop at their program site. CCCs also help support the teachers' use of stress management skills and techniques.

- **Individualized child-centered consultations.** Beginning in the spring, the CCCs provide one-on-one clinical services for a small number of children who have not responded sufficiently to teachers' improved classroom management. By design, the individualized clinical consultation is delivered only *after* children have had ample time to react to the new teaching strategies.

[2]Barrera et al. (2002); Brotman et al. (2005); Dumas, Prinz, Smith, and Laughlin (1999); Gorman-Smith et al. (2006); Webster-Stratton, Reid, and Hammond (2001).

[3]Raver et al. (2008); Raver et al. (2009a); Raver et al. (2009b).

[4]For more information, see the Web site: http://www.incredibleyears.com.

The Design of Foundations of Learning and the Implementation Context

The findings discussed in this report focus on the Newark FOL site, which included 51 preschools serving primarily 4-year-old children. FOL operated in each of the three primary preschool venues in Newark — Head Start centers, community-based child care centers, and public schools — and was conducted in collaboration with the Newark Public Schools, Newark Preschool Council, and Family Connections (a community-based counseling and family services agency). In each preschool, one classroom was selected for participation in the study. Of the 51 preschools, 26 were randomly assigned to receive the FOL intervention (the "program" group), and 25 were assigned to the control group, where they experienced their school year as any other preschool classroom in Newark. This experimental design represents the gold standard of evaluation research. In short, with this design, the study reliably assesses the added value of FOL over and above standard practice in preschool classrooms.

FOL was implemented in the context of preschool classrooms that were subject to the requirements of a series of New Jersey Supreme Court decisions in the *Abbott v. Burke* class action case, which required the state to increase education funding for disadvantaged districts. Abbott mandates include smaller class sizes (limited to 15 students), lower teacher-student ratios (two teachers per classroom), higher teacher salaries, and stricter teacher credentialing, among other features. In this context, it is important to note that the "bar" in Newark was set relatively high for improvements in center quality, compared with more typical urban districts.

As described in detail in a 2009 implementation report, the FOL intervention was implemented with fidelity and quality in the FOL preschools.[5] Most teachers received the training in the new strategies, reported that this training was of high quality, and received consultation to support implementation, generally as the model intended — suggesting that the demonstration is a fair test of FOL.

Impacts of Foundations of Learning on Classroom Context

Did the components of the FOL intervention lead to measurable improvements in teachers' behavior in their classrooms? To address this question, observers were sent to observe classrooms (blind to whether they were in the program group or control group classroom), where they used a standardized observational tool. Findings are summarized in Table ES.1, which compares these observer ratings across the two groups of classrooms. Each measure in the table is reported on a scale from 1 to 7, with 1 and 2 indicating low levels, 3 to 5 indicating moderate levels, and 6 and 7 indicating high levels. Stars (asterisks) indicate the differences that

[5]Lloyd and Bangser (2009).

Table ES.1

Program Impacts on Observed Ratings of Teacher Behavior in the Classroom, Preschool Year

Variable	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Positive classroom management					
Composite[a]	5.8	5.2	0.6 **	0.3	0.75
Positive climate	5.6	5.0	0.6	0.4	0.60
Negative climate	1.1	1.8	-0.6 ***	0.2	-0.90
Teacher sensitivity	5.2	4.8	0.4	0.3	0.46
Behavior management	5.4	4.7	0.8 **	0.4	0.72
Use of classroom time					
Management of classroom time	5.4	4.9	0.5 *	0.3	0.63
Amount of instructional time (minutes)	35.6	25.1	10.6 **	4.4	0.96
Quality of language instruction					
Composite	4.3	3.8	0.5	0.3	0.56
Regard for student perspectives	5.1	4.9	0.2	0.3	0.28
Use of engaging teaching methods	4.2	3.5	0.6 *	0.3	0.61
Promoting understanding through conversation	3.5	3.0	0.4	0.4	0.44
Encouragement of students' language use	4.3	3.6	0.7	0.5	0.54
Sample size	26	25			

SOURCES: MDRC calculations using classroom observations in September-October 2007 and April-May 2008.

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

The table presents adjusted means that control for random assignment blocks and baseline (fall) CLASS dimension scores. For each dimension, observers rated classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

[a]"Negative climate" is reverse-coded for the composite score.

are statistically significant and, therefore, are unlikely to be due to chance. The effect size (shown as a percentage of a change in standard deviation) allows a comparison of impacts for measures that are assessed on different scales.

- **The FOL intervention improved teachers' ability to address children's behavior and to provide a positive emotional climate in the classroom.** Program group teachers used more positive affect (positive climate), displayed less sarcasm and anger (negative climate), and showed a greater ability to comfort children (teacher sensitivity) and prevent misbehavior (behavior management) by setting clear expectations and using effective praise

than did their counterparts in the control group. Therefore, the first hurdle for the intervention was cleared — showing benefits in those aspects of classroom management that teachers were trained in during the FOL training sessions.

- **The FOL intervention also improved the management of classroom time, the use of engaging teaching methods, and the amount of instructional time.** In particular, instructional time was significantly higher in the FOL classrooms than in the control classrooms, by an average of 10 minutes out of a 120-minute observation period. This would translate to 50 minutes more instruction a week, or an entire week's more instruction over a school year. These improvements are consistent with the governing hypothesis behind the demonstration: that addressing teachers' classroom management skills may reduce nonproductive time in preschool classrooms.

- **While FOL improved the management of classroom time, it did not affect the quality of language instruction that children received.** Although not a primary outcome of the intervention model, the study sought to determine whether FOL enabled teachers to engage in higher-quality language interactions with children during instructional activities or whether, on the other hand, focusing on emotional and behavioral adjustment interfered with instructional support for children (which would result in reductions in the quality of instruction in FOL classrooms). Neither appears to be true, as there are no consistent statistically significant differences between FOL and control classrooms on measures of the quality of language used in the classroom.

Impacts of Foundations of Learning on Children

A second observation team rated a subset of children on their conflict and positive interactions with teachers and peers as well as on the extent to which the children were engaged in classroom activities, using the same 1-to-7 scale that was used for the classroom observations. In addition, teachers completed surveys on all children to rate their perceptions of children's problem behavior and positive social behavior.

The findings on both these sets of data are presented in Table ES.2.

- **FOL led to reductions in conflicts with teachers and peers, but the intervention did not otherwise change the quality of teacher-child or peer interactions.** Children in FOL classrooms were observed to have statistically lower levels of conflict, on average, than were children in control classrooms.

The Foundations of Learning Demonstration

Table ES.2

Program Impacts on Observed and Teacher Ratings of Child Outcomes, Preschool Year

Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Observations					
Problem behavior					
Teacher conflict	1.2	1.5	-0.2 ***	0.1	-0.40
Peer conflict	1.4	1.6	-0.2 *	0.1	-0.27
Positive social behavior					
Teacher communication	2.2	2.4	-0.2	0.1	-0.20
Teacher positive engagement	3.2	3.4	-0.2	0.2	-0.27
Peer communication	2.5	2.6	-0.1	0.2	-0.14
Peer sociability	3.4	3.5	-0.1	0.1	-0.11
Peer assertiveness	2.1	2.3	-0.2	0.2	-0.21
Approach to learning					
Task engagement	4.9	4.6	0.2 *	0.1	0.31
Task self-reliance	3.1	3.1	-0.1	0.2	-0.07
Task behavior control	5.4	5.1	0.3 *	0.2	0.34
Overall classroom student engagement	5.7	5.2	0.6 *	0.3	0.60
Teacher reports					
Problem behavior					
Internalizing problems	2.7	2.3	0.4	0.6	0.11
Externalizing problems	4.1	3.7	0.4	0.7	0.08
Teacher-student conflict	12.4	12.3	0.1	0.9	0.02
Positive social behavior					
Social competence	4.0	4.0	0.0	0.1	0.06
Teacher-student closeness	34.5	35.8	-1.3	0.9	-0.24
Approach to learning					
Work-related skills	4.8	4.8	0.1	0.1	0.08
Preacademic skills					
Language and literacy skills	35.1	32.6	2.5	1.7	0.27
Math knowledge	25.8	25.4	0.4	1.7	0.05
Sample size - observations of students	130	121			
Sample size - teacher reports on students	283	248			
Sample size - classrooms	26	25			

(continued)

Table ES.2 (continued)

SOURCES: Based on MDRC calculations of classroom observations and a teacher survey.

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

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

The observed outcomes "Problem behavior," "Positive social behavior," and "Approach to learning" come from the inCLASS observations. "Overall classroom student engagement" comes from the CLASS. For each dimension, observers rated children and classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

Teacher-reported outcomes control for the child's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

These effects were most pronounced for children who entered preschool with the highest levels of behavior problems. However, there were no effects on positive aspects of teacher-child and peer interactions — a disappointing finding, given that an aim of the intervention was not only to reduce conflict but also to improve children's relationships.

- **Children in the program group demonstrated greater levels of engagement in the classroom than did children in the control group.** Children in FOL classrooms were also rated higher on their ability to regulate behavior during tasks than were children in control classrooms. If children are able to spend more time on task, they may be able to take greater advantage of the formal and informal learning opportunities in the preschool classroom.

- **Surprisingly, *teachers* did not report differences in children's behavior between the program and control groups.** As shown in the bottom panel of Table ES.2, even though the independent research team saw FOL children as having fewer behavior problems and being more engaged, there are no statistically significant differences between the two groups of children in teacher ratings of children's problem behavior, positive social behavior, and approach to learning. Researchers found no differences among those at low and high levels of behavior problems at preschool entry. The findings for the teacher-reported outcomes are somewhat surprising; one hypothesis is that the training that teachers received primed them to see challenging behaviors, even as it increased their capacity to effectively manage these behaviors when they occurred.

The Year *Following* Foundations of Learning

Limited information is available about the year following the delivery of the FOL intervention in preschool classrooms. The children were dispersed to a large number of schools (about 100) for their kindergarten year, and FOL and control children were together in a large number of the kindergarten classrooms. Conclusions are based solely on kindergarten teacher reports, which presents some advantages and disadvantages. On the one hand, kindergarten teachers were largely blind to children's FOL program-group status, which makes them less biased reporters. On the other hand, teachers' reports are somewhat less reliable than information that would be collected from trained observers.

The effects of FOL on kindergarten teachers' ratings of child outcomes are presented in Table ES.3. Unlike the preceding tables, this one presents only the control group level and the impact of FOL for each of the measures collected. These are shown for the full sample and for two groups of children defined by their level of behavior problems (low or high) when they entered preschool.

- **Based on kindergarten teacher reports, overall FOL had very few sustained effects on children the year after they received the intervention.** Surprisingly, the only significant effects were observed in the negative direction: Teachers reported higher levels of behavior problems among children from the FOL classrooms than among children from the control classrooms. It is possible that sustained effects might require kindergarten teachers who employ similar skills and strategies to redirect behavior.

- **Impacts in kindergarten differed somewhat, depending on children's initial level of behavior problems.** For children with the lowest levels of behavior problems, FOL had no ongoing effects on problem behavior or positive social behavior. However, significant *positive* effects were found on these children's learning-related skills (a measure of student engagement) and on their language and literacy skills — showing sustained effects for these lowest-risk children. For children with elevated behavior problem scores in the fall of preschool, the kindergarten follow-up found evidence of *increased* withdrawn and sad behavior. No other statistically significant differences were found among children who received FOL in preschool and those who did not.

Did the intensive investment in preschool teachers' professional development — substantial training plus one day per week of in-person consulting — result in changes in their *ongoing* practice in the following school year, when they were no longer receiving that support? Findings are presented in Figure ES.1.

The Foundations of Learning Demonstration

Table ES.3

Program Impacts on Teacher Ratings of Child Outcomes, by Full Sample and Level of Behavior Problems, Kindergarten

	Full Sample			Child's Level of Behavior Problems						
				Low			High			
Outcome	Control Group Mean	Difference (Impact)	Effect Size	Control Group Mean	Difference (Impact)	Effect Size	Control Group Mean	Difference (Impact)	Effect Size	H-Stars
Problem behavior										
Internalizing problems	2.4	0.8 *	0.24	1.8	0.0	0.00	3.5	1.2 *	0.34	
Externalizing problems	3.8	1.0 *	0.16	3.0	-0.3	-0.05	5.3	1.5	0.22	
Teacher-student conflict	11.5	0.8	0.11	10.5	-0.1	-0.02	13.6	0.2	0.02	
Positive social behavior										
Social competence	4.0	0.0	0.03	4.1	0.1	0.10	3.7	0.0	0.04	
Teacher-student closeness	33.9	-0.8	-0.14	33.6	0.8	0.13	33.7	-1.2	-0.21	
Approach to learning										
Work-related skills	4.9	0.0	-0.04	5.0	0.4 *	0.33	4.6	-0.3	-0.22	**
Academic skills										
Language and literacy skills	33.5	0.3	0.03	33.9	2.5 *	0.26	32.6	-1.7	-0.19	**
Math knowledge	26.5	0.0	0.00	26.9	1.2	0.16	26.1	-1.0	-0.15	
Sample size	230			105			99			

SOURCE: Based on MDRC calculations from responses to teacher survey.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent. The final column lists the H-stars, which show the statistical significance of the difference between the subgroup impacts.

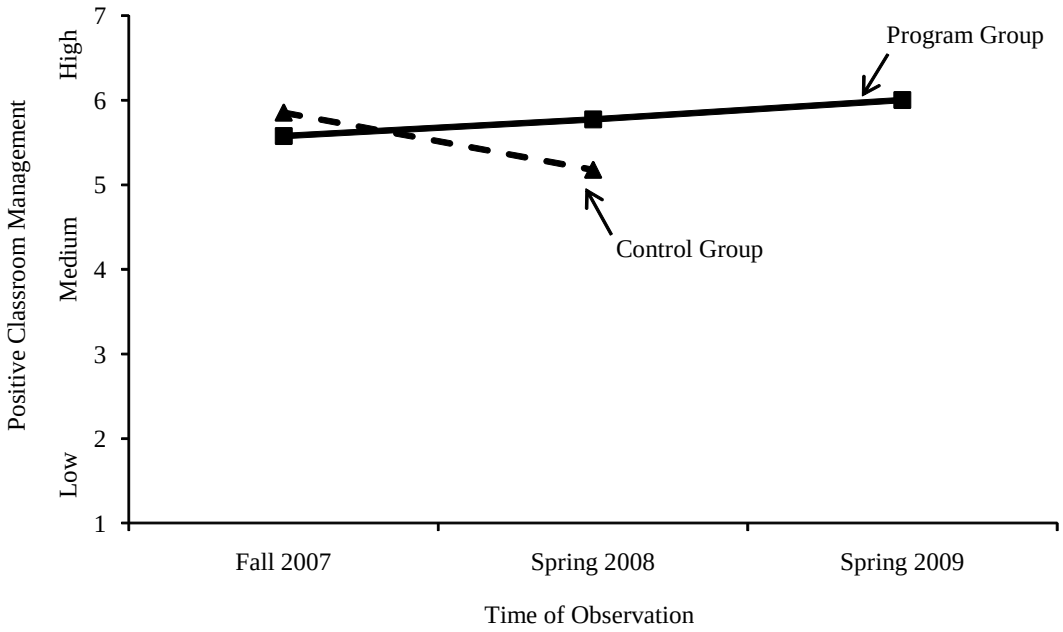
Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics. Outcome controls for the child's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

The Foundations of Learning Demonstration

Figure ES.1

Observed Ratings of Teachers' Positive Classroom Management During Intervention and Following Year



SOURCES: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007, April-May 2008, and April-May 2009.

NOTES: Scores represent adjusted means that control for random assignment block. Spring 2008 and spring 2009 control for baseline fall CLASS dimensions as well.

Spring 2008 scores for the program group and the control group are statistically different (p-value = less than 5 percent).

- **When observed in the year following the FOL intervention, teachers who were assigned to the intervention appeared to have continued to engage in the positive practices that they had learned.** When comparing the teachers' scores on positive classroom management from the spring of the intervention year and from the following spring, scores were largely maintained. Concern that teachers might not continue to use the skills and strategies they had learned in the prior year without the direct support of the intervention was unfounded.

Conclusion

The evidence emerging from the Newark site of the FOL demonstration is encouraging on some measures, showing that investments in teachers' professional development can make a difference in children's experience in preschool. In short, the intervention changed not only the primary outcome that was targeted — teachers' positive classroom management — but also the productive use of classroom time. Benefits to children were observed during preschool on some but not all aspects of their behavior in the classroom. Yet those benefits were not sustained as the children moved to their kindergarten classrooms. Understanding how to sustain effects beyond preschool is critical to addressing the long-term needs of low-income children, especially those at highest risk of emotional and behavioral challenges.

If training in classroom management can have benefits for the provision of preschool instruction, the next question is how are preschool teachers using that increased productive time — that is, do teachers use the time to teach the kinds of skills that preschool children need to transition successfully into kindergarten? The fact that these early investments in teachers' professional development are sustained beyond the one year of intensive intervention efforts suggests that pairing this intervention with a cognitively focused curriculum in a second year might be a promising approach for enhancing the school readiness of preschool children.

Additional findings on this intervention are forthcoming, including from the Chicago site of the FOL demonstration. The FOL study will add to the research emerging on other promising social-emotional interventions in preschool classrooms, including the Head Start Classroom-based Approaches and Resources for Emotion and Social skill promotion (CARES) project, sponsored by the Administration for Children and Families in the U.S. Department of Health and Human Services, which is testing several different social-emotional curriculum enhancements across Head Start centers as part of a national demonstration.

References for the Executive Summary

Barrera, Manuel, Anthony Biglan, Ted K. Taylor, Barbara K. Gunn, Keith Smolkowski, Carol Black, Dennis V. Ary, and Rollen C. Fowler. 2002. "Early Elementary School Intervention to Reduce Conduct Problems: A Randomized Trial with Hispanic and Non-Hispanic Children." *Prevention Science* 3: 83-94.

Brotman, Laurie Miller, Kathleen Kiely Gouley, Daniel Chesir-Teran, Tracy Dennis, Rachel G. Klein, and Patrick Shrout. 2005. "Prevention for Preschoolers at High Risk for Conduct Problems: Immediate Outcomes on Parenting Practices and Child Social Competence." *Journal of Clinical Child and Adolescent Psychology* 34, 4: 724-734.

Dumas, Jean E., Ronald J. Prinz, Emilie Phillips Smith, and James Laughlin. 1999. "The EARLY ALLIANCE Prevention Trial: An Integrated Set of Interventions to Promote Competence and Reduce Risk for Conduct Disorder, Substance Abuse, and School Failure." *Clinical Child and Family Psychology Review* 2: 37-53.

Gorman-Smith, Deborah, Patrick H. Tolan, David B. Henry, Elena Quintana, Kelly Lutovsky, and Amy Leventhal. 2006. "Schools and Families Educating Children: A Preventive Intervention for Early Elementary School Children." Pages 113-135 in Patrick Tolan, José Szapocznik, and Soledad Sambrano (eds.), *Preventing Youth Substance Abuse: Science-Based Programs for Children and Adolescents.* Washington, DC: American Psychological Association.

Lloyd, Chrishana M., and Michael Bangser. 2009. *Promoting Preschool Quality Through Effective Classroom Management: Implementation Lessons from the Foundations of Learning Demonstration.* New York: MDRC.

Raver, C. Cybele, Stephanie M. Jones, Christine P. Li-Grining, Molly Metzger, Kina Smallwood, and Latriese Sardin. 2008. "Improving Preschool Classroom Processes: Preliminary Findings from a Randomized Trial Implemented in Head Start Settings." *Early Childhood Research Quarterly* 63, 3: 253-255.

Raver, C. Cybele, Stephanie M. Jones, Christine P. Li-Grining, Fuhua Zhai, Kristen Bub, and Emily Pressler. 2009a. "CSRP's Impact on Low-Income Preschoolers' Pre-Academic Skills: Self-Regulation as a Mediating Mechanism." Unpublished paper.

Raver, C. Cybele, Stephanie M. Jones, Christine P. Li-Grining, Fuhua Zhai, Molly Metzger, and Bonnie Solomon. 2009b. "Targeting Children's Behavior Problems in Preschool Classrooms: A Cluster-Randomized Controlled Trial." *Journal of Consulting and Clinical Psychology* 77: 302-316.

Webster-Stratton, Carolyn, M. Jamila Reid, and Mary Hammond. 2001. "Preventing Conduct Problems, Promoting Social Competence: A Parent and Teacher Training Partnership in Head Start." *Journal of Clinical Child Psychology* 30, 3: 238-302.

Chapter 1

Introduction

Policymakers increasingly recognize that early childhood education is a promising strategy for improving the school readiness of disadvantaged young children. Yet even as policymakers and administrators make great strides in investments in early childhood programs, they confront a difficult challenge: How can quality preschools be created and maintained? This question is critical, considering that preschools of mediocre or lower quality may undermine children's development.[1]

One key aspect of preschool quality is addressing children's emotional and behavioral adjustment — including their ability to engage positively with peers and teachers and to control their attention and behavior during classroom activities. Evidence suggests that improving young children's healthy emotional and behavioral development is both an important outcome in its own right and can be a pathway to improved academic achievement. Recent research documents high levels of behavior problems for children in preschool classrooms. Unsurprisingly, then, teachers consistently emphasize their need for professional development and other supports to help them address children's behavioral issues.

This report presents findings from the Newark, New Jersey, site of the Foundations of Learning (FOL) demonstration, a random assignment evaluation conducted by MDRC of an intervention that trains preschool teachers to better support children's behavior and emotional development. The demonstration adopted the model used by CSRP (formerly the Chicago School Readiness Project) and adapted it slightly to fit a new policy context.[2] The remainder of this report refers to the model as "FOL." The FOL intervention was tested in two cities — Newark and Chicago — and it combined teacher training in effective classroom management with weekly classroom consultation. Consultants coached and mentored the teachers in the new strategies learned in the training workshops and provided individualized support to the highest-risk children. As discussed below, the rationale behind FOL's focus on children's emotional and behavioral adjustment is threefold: (1) children's behavioral and emotional problems represent a highly salient issue facing teachers of low-income children; (2) these problems may be impeding children's ability to take advantage of learning opportunities in the classroom; and (3) teachers have little training in how to address these issues in their classrooms.

[1]Magnuson, Ruhm, and Waldfogel (2007).

[2]CSRP is not associated with The Chicago School®, which is a trademark of The Chicago School of Professional Psychology.

This report focuses on the effects of the Newark FOL intervention on teachers' ability to support children's emotional and behavioral adjustment and to manage their difficult behaviors, on teachers' management of classroom time, on children's ability to engage positively with peers and teachers, and on children's engagement in preschool learning. Findings on the effects of this model as these children progress into kindergarten are also presented. A 2009 report documents the implementation of FOL in this site, drawing operational lessons from this experience for those interested in mounting similar models.[3]

The Policy Context

FOL's focus on quality improvements comes at a particularly important juncture. Educators and researchers increasingly recognize that high-quality early childhood programs are an important way to prepare children for later school success, and, correspondingly, they have advocated for a marked expansion in preschool education services. For example, growth in state-funded prekindergartens accelerated in the 1990s and 2000s, fueled by public support for early education along with the increased labor market participation of mothers. State prekindergarten education has expanded from 28 states in the early 1990s to 38 states in 2007.[4] Georgia led efforts to expand to "universal" prekindergartens serving children regardless of income level, with several states (Oklahoma, New York, and West Virginia) following close behind. Moreover, state-funded prekindergartens are only one of a broader set of programs providing care and education to preschool-age children that have seen expansions over this period. In fact, early childhood programs benefit from many government expenditures that are not specifically funding preschools; federal expenditures for programs affecting young children were approximately $20 billion in 2006.[5] Furthermore, federal expenditures on young children are likely to increase in future years due to President Obama's strong focus on newborn care and early childhood education. The 2010 budget request alone includes increases of over $9 billion for home visiting programs and preschool initiatives.[6]

At the same time, there is a growing need for preschool intervention services to help close the pernicious gap between the school achievement of low-income children and their more affluent peers. Studies find that low-income children fare more poorly on indices of

[3]Lloyd and Bangser (2009).

[4]Barnett et al. (2008a).

[5]This number is based on MDRC calculations of data gathered on funding for programs for young children, including the Child Care and Development Block Grant, the Child Care and Development Fund, Temporary Assistance for Needy Families (TANF), Head Start, the Social Services Block Grant, the Department of Education, and the Department of Housing and Urban Development.

[6]U.S. Office of Management and Budget (2010).

academic achievement than their higher-income peers,[7] with scores for low-income children being as large as 1 standard deviation below national norms.[8] In a nationally representative survey of kindergarten teachers, 30 percent reported that at least half the children in their class lacked academic skills and had difficulty following directions and working as part of a group — critical skills for engaging in the learning tasks of school.[9] Even more concerning is that the gaps in school achievement that are observed at the start of children's formal schooling remain or increase over the course of the elementary years.[10]

Against this backdrop of concern for children's achievement and the growth of preschool services came the passage of the No Child Left Behind (NCLB) Act of 2001. With accountability being linked to children's performance on "high-stakes" academic tests, NCLB focused preschool administrators' attention on building children's academic readiness for school prior to their entry into formal education. At the same time, developmental researchers raised the concern that meeting NCLB benchmarks might lead to an exclusive focus on cognitively oriented outcomes. They called for a complementary focus on children's social, emotional, and behavioral adjustment as a key part of children's success in school.[11] Correspondingly, surveys of teachers repeatedly show that their foremost concern is whether children are emotionally and behaviorally ready for the demands of formal learning in classroom settings.[12]

Why Focus on Children's Emotional and Behavioral Adjustment?

Children in preschool who have behavioral challenges are more likely to face social, behavioral, and academic difficulties throughout their school careers than are more behaviorally well-adjusted children. Conversely, children who learn to focus their attention and regulate their impulsivity in the face of classroom distractions are likely to have greater opportunities for learning than their behaviorally disruptive peers. In addition, a child who has difficulty negotiating the classroom environment may disrupt other children's academic and social progress in the classroom. When children act out aggressively or become sad and withdrawn, teachers may be diverted from instructional time to manage these behaviors.

Behavior problems are not a minor or peripheral problem in preschool classrooms. Studies have documented rates of emotional and behavior problems among preschool children to be as high as 15 percent to 20 percent.[13] In a classroom of 15 to 20 children, that means as

[7]Lee and Burkham (2002).
[8]Kaiser, Xinsheng, Hancock, and Foster (2002).
[9]Rimm-Kaufman, Pianta, and Cox (2000).
[10]Alexander, Entwisle, and Kabbani (2001); Entwisle and Hayduk (1988).
[11]Raver (2002).
[12]U.S. Department of Education, National Center for Education Statistics (2001).
[13]Campbell (1995); Lavigne et al. (1996).

many as 3 or 4 children may be acting out — and taking a substantial portion of the teacher's attention away from more productive activities. Not only does this contribute to teachers' stress and burnout, but children with low levels of social competence and high rates of behavior problems represent an especially large percentage of school districts' high-expenditure pupils.[14] These children are more likely to repeat a grade early in elementary school[15] and are more likely to receive special education services.[16]

What does research reveal about teaching practices and children's behavior problems? Observational studies find that teachers and children can easily become caught up in cycles of negative interactions[17] in which adults inadvertently exacerbate children's acting out and aggressive behavior through harsh and ineffective techniques (for example, by failing to set limits on children's misbehavior or being overly negative with a child who is not cooperating in the classroom). As might be expected, children respond poorly to these ineffective strategies, becoming more emotionally negative and disruptive. Adults, for their part, can become exasperated and may give up on their own attempts to control children's negative behavior, thereby inadvertently reinforcing it.[18]

Unfortunately, preschool teachers generally receive very little training about how to address these issues in their preschool classrooms, which can lead to ineffective classroom management, increased stress and burnout, and high rates of turnover. One approach to addressing this problem is to train preschool teachers to proactively support children's positive behavior while more effectively limiting their aggressive and disruptive behavior.[19] A number of studies show that efforts to train teachers and parents in these skills can be effective.[20] The question is whether these strategies can make a difference when implemented on a larger scale and in preschools serving large numbers of low-income children. In addition, it is unclear whether improving these skills among teachers will result in changes in children's approach to learning, in their social engagement with others, and in their emotion regulation skills.

[14]Chambers, Kidron, and Spain (2004).

[15]Beebe-Frankenberger, Bocian, MacMillan, and Gresham (2004).

[16]Centers for Disease Control and Prevention, National Center for Health Statistics (2005); Wagner and Blackorby (2002).

[17]Patterson (1982); Patterson, Reid, and Dishion (1992); Patterson (1996).

[18]Dishion, French, and Patterson (1995).

[19]Barrera et al. (2002); Brotman et al. (2005); Dumas, Prinz, Smith, and Laughlin (1999); Gorman-Smith et al. (2006); Webster-Stratton, Reid, and Hammond (2001).

[20]Brotman et al. (2005); Webster-Stratton, Reid, and Hammond (2001).

The Foundations of Learning Intervention

The FOL intervention sought to provide teachers with specific tools that they could use in handling the daily classroom challenges presented by children's emotional and behavioral difficulty. In so doing, FOL aimed to improve the quality of daily interactions among teachers and their students, to increase teachers' ability to manage the classroom environment, and to increase children's emotional and behavioral adjustment. As shown in Figure 1.1, the intervention includes four components delivered throughout the preschool year:[21]

- **Teacher training.** Lead and assistant teachers are invited to attend five Saturday training sessions, once a month for six hours each, from late September to January. The workshops are an adapted version of *The Incredible Years* curriculum developed by Dr. Carolyn Webster-Stratton.[22] The workshops instruct teachers in how to develop positive relationships with children and their families; present classroom strategies that teachers can use, such as setting clear rules, outlining predictable limits, and instituting a discipline structure that minimizes classroom disruptions and avoids confrontation; and provide teachers with techniques to develop children's social skills, anger management, and problem-solving ability.

- **Classroom-level consultation.** To complement the training, teachers are assigned a master's-level Clinical Classroom Consultant (CCC) to work with them in the classroom one day per week throughout the school year. The CCCs play an important role in modeling and reinforcing the content of the training sessions and in acting as a sounding board for teachers.

- **Stress management.** In January or February, lead and assistant teachers participate in a customized, 90-minute stress management workshop at their program sites. In the months leading up to and following the workshop, the CCCs help support the teachers' use of stress management skills and techniques.

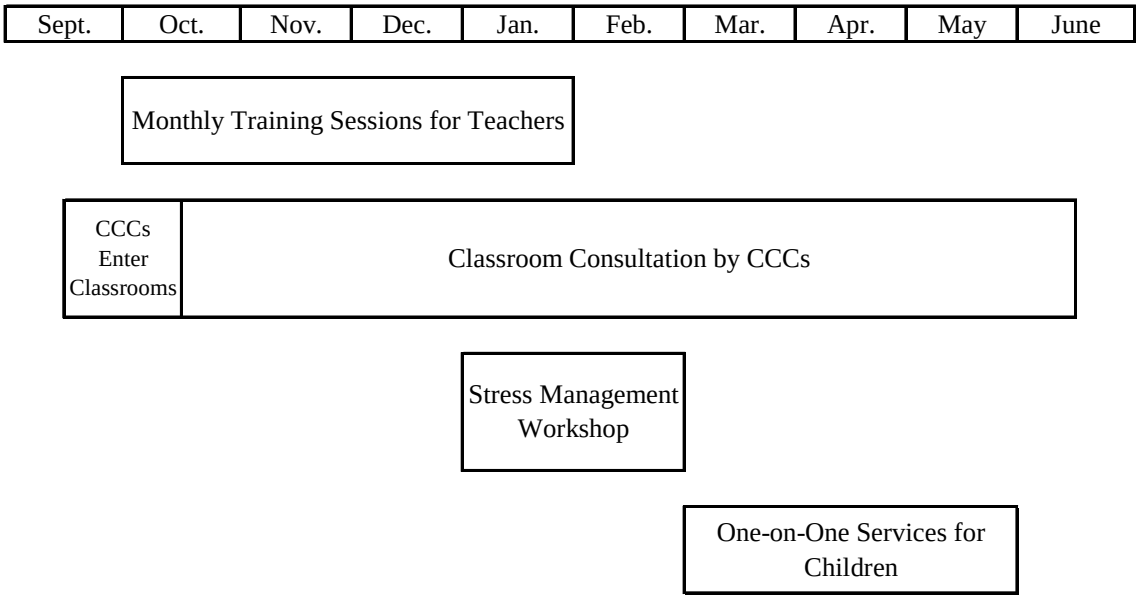
- **Individualized child-centered consultation.** Beginning in March, the CCCs provide one-on-one clinical services for a small number of children who have not responded sufficiently to the teachers' improved classroom management.

[21]Appendix E gives information about the costs associated with implementing FOL in Newark.
[22]For more information, see the Web site: http://www.incredibleyears.com.

The Foundations of Learning Demonstration

Figure 1.1

2007-2008 Intervention Timeline



NOTE: CCC = Clinical Classroom Consultant.

The timing for the delivery of these components is deliberate. First, the teacher trainings are spaced one month apart throughout the fall and winter. This allows teachers to learn components of the model over time and to work over the next weeks to "carry back" the new strategies that they have learned from the training into their classrooms, gradually integrating the new strategies into their daily practice. Second, although the CCC enters the classroom in September, she or he does not start active consultation with teachers until October, after the first training session. This allows the CCC to understand the institutional context in which the teacher is working, to get to know the children, and to build a productive working relationship with the teacher. This relationship provides a critical foundation for successful consultations. The stress management workshop builds on these relationships and is timed to coincide with heightened winter stress. Finally, the individualized child services are deferred until the final third of the school year (March to May). In this way, the teacher training can be fully implemented and allows children to have sufficient time to respond to the changes in teacher practice, before deciding which children should be targeted for these one-on-one services.

The Evolution of the Foundations of Learning Demonstration

The FOL intervention model, and the demonstration as a whole, drew heavily on the lessons from CSRP. Designed and developed by Dr. Cybele Raver, who is now a member of the FOL research team, CSRP operated in 18 Head Start sites in high-poverty Chicago neighborhoods from 2004 to 2006. The evaluation of CSRP used a rigorous research design, in which the 18 sites were randomly assigned to one of two groups: Half received the multicomponent CSRP intervention, and the other half served as a control group.[23]

The results from the CSRP evaluation indicate that the intervention improved the quality of the classroom environment as well as outcomes for children. Encouraged by these findings, MDRC decided to test the model on a larger scale. Following a feasibility study and a yearlong pilot phase of FOL in Newark in the 2006-2007 school year, MDRC proceeded with full-scale demonstrations in Newark during the 2007-2008 school year and in Chicago in the 2008-2009 school year. Table 1.1 presents the timelines of CSRP and FOL.

As in CSRP, the FOL research design includes the random assignment of preschool sites either to receive the enhanced services (the program sites) or to continue their standard operations (the control sites). Yet the original CSRP evaluation and the FOL demonstration in Newark differ in some respects, two of which are most important:

- **Scale.** With 51 FOL sites (26 program sites and 25 control sites) in one year, the full FOL demonstration phase in Newark operated on a larger scale than the CSRP evaluation, which operated in 18 sites (9 program sites and 9 control sites) over two cohorts.[24]

- **Delivery mechanism.** CSRP was a university-based intervention, while FOL was embedded into a typical service delivery framework established in Newark. For example, the clinical consultants in CSRP were retained by the University of Chicago, where the research team was based; in Newark, however, the CCCs and their coordinator were employees of Family Connections, a community-based counseling and family services agency. By imple-

[23]CSRP included a teacher's aide in control classrooms; the presence of the teacher's aide in control classrooms ensured that any positive impacts of the intervention were not attributable to an improved adult-child ratio in the CSRP classrooms.

[24]CSRP Cohort 1 operated from 2004 to 2005 and consisted of 10 sites (5 program sites and 5 control sites). Cohort 2 operated from 2005 to 2006 and consisted of 8 sites (4 program sites and 4 control sites).

Table 1.1

Evolution of the Foundations of Learning Demonstration

Characteristic	CSRP	FOL Pilot in Newark	FOL Full Scale in Newark	FOL Full Scale in Chicago
School years	2004-2006	2006-2007	2007-2008	2008-2009
Total number of sites	18	17	51	20
Program sites	9	9	26	10
Control sites	9	8	25	10
Total number of classrooms	35	17	51	40

SOURCES: Raver et al. (2008) and MDRC calculations from random assignment.

menting FOL in a larger number of classrooms and by subcontracting the services of a local agency, the FOL demonstration can evaluate whether the intervention could be implemented with fidelity in a larger, less tightly controlled study that is more similar to the "real world" demands faced by state and local policy professionals and administrators.

The findings presented in this report focus on the full-scale effort in the Newark site, which involved 51 preschools during the 2007-2008 school year. Outcomes are presented for classrooms and children during the year of the intervention (the preschool year) and for one year following (the kindergarten year). This allows an examination not only of whether the intervention made a difference in the preschool classroom but also of whether preschool children sustain these benefits across a key educational transition (that is, the transition to elementary school), "taking with them" into their *next* school environment the skills gained in the context of the FOL intervention.

The Context for Implementing Foundations of Learning in Newark

The Newark FOL demonstration was conducted in close collaboration with the Newark Public Schools, Newark Preschool Council, and Family Connections. In addition, MDRC staff played an active role in ensuring the fidelity of the model throughout its implementation.

Newark is ahead of most of the country in its implementation of structural changes to promote quality in preschools. The New Jersey Supreme Court's landmark decisions in the *Abbott v. Burke* class action case required the state to increase education funding to disadvantaged districts. In short, the court ruled that the state must provide additional resources, oversight, and regulations for 31 low-income school districts in the state, including Newark.

More specifically, at the preschool level, the Abbott rulings mandate that the state provide free preschool education to all children whose parents wish to enroll them. Additionally, the court ordered that:[25]

- All lead teachers are required to have a minimum of a bachelor's degree. This requirement, coupled with a substantial increase in financial resources, make Abbott-funded Newark preschool teachers among the highest paid and most credentialed in the nation.[26]
- All classrooms must have a lead teacher and an assistant teacher.
- Each preschool classroom must have no more than 15 students.
- The size of each preschool classroom must be at least 950 square feet.[27]
- Each preschool classroom must have a standard set of materials and distinct activity areas for children, including a sand and water table, a dramatic play area, a library, and an area specifically for playing with blocks.
- All facilities must provide special education, bilingual education, transportation, and health services.
- Teachers must use a standard curriculum assigned by the school district. At the time of this study, Newark utilized the Creative Curriculum.[28]

It is important to keep in mind that the impact findings presented in this report represent the incremental effect of adding the FOL approach on top of these requirements of the Abbott mandates.

As described in detail in the 2009 implementation report, there was initially some concern regarding whether the FOL intervention was even needed in Newark, given the substantial level of resources already mandated and provided in its publicly funded preschool system.[29] However, during the 2006 feasibility study, MDRC team members interviewed Newark educational administrators and teachers, who responded that there was a dearth of resources to specifically address the emotional and behavioral needs of preschool children. The pilot findings from the 2006-2007 school year further confirmed this need.

---

[25]New Jersey Department of Education (2008).

[26]Barnett et al. (2008a); U.S. Department of Labor, Bureau of Labor Statistics (2009).

[27]However, since this could require construction or renovation of facilities, preschools with classrooms below this size were initially allowed to remain within the Abbott system.

[28]For more information, see the Web site: http://www.creativecurriculum.net.

[29]Lloyd and Bangser (2009).

How Well Was Foundations of Learning Implemented in Newark?

Before turning to a discussion of the impact results — comparing centers that were offered the intervention with those that were not — it is important to understand whether the intervention was implemented with fidelity to the model. That is, it is important to understand whether the comparisons between the two groups of centers represent a "fair test" of the FOL intervention.

As discussed in more detail in the 2009 implementation report, the intervention was, in fact, implemented with fidelity and quality, as evidenced by the following findings:[30]

- **Trainings for teachers were well attended, and the quality of trainings was high.** For example, teachers gave high ratings (between 4.6 and 4.7 on a scale of 5.0) when asked whether the training content was clear, the training environment was conducive to learning, the trainers themselves were effective and clear, and the training enhanced the teachers' professional development.

- **The number of hours (or "dosage") of classroom consultation that CCCs provided to teachers was less than the full amount scheduled, but it exceeded what is typically offered in early childhood consultation models.**[31] CCCs provided an average of just over 162 hours (or about 23 days) of in-classroom consultation over the course of the academic year. Teachers gave high ratings to the quality of the consultation.

- **Stress management workshops turned out to be more important to teachers than intervention developers had expected.** All the treatment classrooms in the FOL demonstration received an on-site stress management workshop. Out of 52 teachers, 49 attended their on-site workshop and rated it highly.

- **As planned, children who needed additional services were provided individualized consultation from CCCs.** A small number of children in each classroom were selected by the consultant and the teacher to receive additional one-on-one support. The advantage of having these services provided by the CCCs is that they were delivered without delay, in the familiar setting of the classroom, and on a consistent basis (weekly or biweekly). This manner of delivering services was well received by children and teachers.

[30]Lloyd and Bangser (2009).
[31]Brennan, Bradley, Allen, and Perry (2008).

Theory of Change: How Might the Components of Foundations of Learning Affect Teachers and Children?

This section describes the theory of change underlying the FOL intervention. As depicted in the conceptual model presented in Figure 1.2, the primary targets of FOL are teachers' positive classroom management skills. These skills were the core components of the Incredible Years training that the teachers received as well as the content of CCCs' modeling and coaching sessions with teachers.

FOL was developed on the premise that managing children's problem behavior was diverting teachers' attention from providing instruction to children in preschool classrooms. Therefore, by changing the way in which teachers managed children's behavior, it was thought that lesson time would be more productive and that downtime in classrooms would be reduced.

While not a formal pathway of influence, FOL also *might* have changed the quality of instructional time, but the direction of that influence was far less clear. On the one hand, it was possible that an intervention that focused on behavior management and the affective climate of the classroom could translate into better teaching. On the other hand, teachers' attention to classroom management might have diverted them from providing cognitively oriented instruction. Both hypotheses are tested in the analyses presented in Chapter 3.

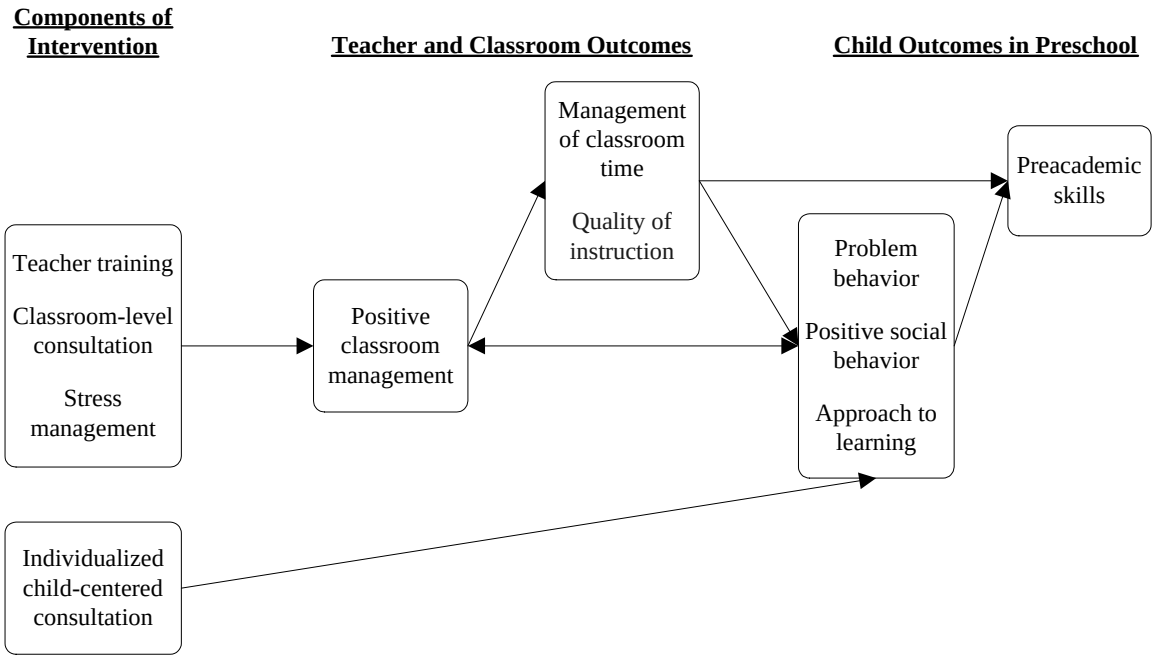
These changes in classroom-level interactions between teachers and children were expected to affect children's problem behavior (acting out and withdrawn behaviors), their positive social behavior (social interactions with teachers and peers), and their approach to learning (their engagement in the learning tasks of preschool). The provision of individualized child-centered consultation provided by the CCC was also thought to benefit the highest-risk children directly.

Notably, FOL focused on children's emotional and behavioral skills, not their preacademic skills directly. Any effects on academic skills would need to occur *through* changes in the classroom or changes in children's emotional and behavioral adjustment. Finally, while not shown in Figure 1.2, there was the hope that, by developing their social and emotional skills in preschool, children would be able to elicit more positive interactions from their teachers in the new kindergarten setting, thus potentially reinforcing or amplifying the initial benefits of the intervention, across time. However, whether or not children do so might depend on the quality of the classrooms that they enter in the early elementary years.

The Foundations of Learning Demonstration

Figure 1.2

Conceptual Model of Intervention Effects



The Organization of This Report

The remainder of the report is organized as follows:

Chapter 2 describes the design of the study, the research sample, and the key outcome measures.

Chapter 3 presents the key findings on the effects of FOL during the intervention year. First, the chapter presents the effects of FOL on those aspects of teacher behavior and classroom processes that were directly targeted by the intervention (positive classroom management), followed by the effects on nontargeted classroom outcomes, such as management of classroom time. Second, the chapter presents the effects of FOL on children's outcomes in preschool.

Chapter 4 presents two sets of findings from the follow-up year: effects on children as they transition to new school environments and effects on teachers as they move into their second year of implementing FOL.

Lastly, **Chapter 5** summarizes the key findings, places them in context, and discusses the future research agenda in this area.

Chapter 2

The Evaluation Sample, Measures, and Analytic Strategy

Chapter 2 discusses the key features of the research design and measurement strategies used in the Newark, New Jersey, site of the Foundations of Learning (FOL) demonstration. The chapter first describes how sites were recruited and were randomly assigned to program and control groups. Then it provides information on the context in which FOL was implemented, with a description of the classrooms, students, and teachers in the research sample. Finally, the chapter presents the analysis strategy used in estimating the impact of FOL.

Recruitment, Random Assignment, and Data Collection

As noted in Chapter 1, a total of 51 preschool sites were recruited for the full-scale phase of the demonstration in Newark in the summer of 2007. All were Abbott-funded preschools being overseen by the Newark Public School system, but they were located in three distinct venues: Approximately half the preschool sites were located in community-based centers, and the remaining sites were split between Head Start and public school sites. This parallels the distribution of locations of Newark preschools in the 2007-2008 school year: Of approximately 125 preschool sites, 59 (about half) operated in community-based child care centers; 33 (about one-fourth) operated in Head Start centers; and 33 (about one-fourth) operated in public schools.[1] One classroom within each preschool, serving primarily 4-year-old children, was selected as the "study" classroom. See Appendix A for further details on the selection of the research sample.

The 51 participating preschools were grouped first by venue (community-based sites, school-based sites, Head Start centers) and then, within venue, by child racial/ethnic composition and city ward — for a total of 12 groups of sites.[2] These groups, or blocks, varied in size from two to nine sites, or centers. Random assignment was conducted *within* each block to ensure representation of sites across blocking characteristics in both the program group and the control group: 26 centers with 26 classrooms were randomly assigned to the FOL intervention,

[1]These calculations of Abbott-funded preschool classrooms/sites were made by MDRC, based on information provided by the Newark Public Schools. The New Jersey Supreme Court's landmark decisions in *Abbott v. Burke* required the state to provide additional resources, oversight, and regulations for 31 low-income school districts in the state, including Newark.

[2]The city of Newark is divided into five political wards: North, South, Central, West and the Ironbound (East). The Newark population consists mainly of African-Americans and Hispanics, largely divided by race/ethnicity in specific wards or neighborhoods. In addition, the city also boasts a large Portuguese or Portuguese-speaking population that is located in the Ironbound district.

and 25 centers with 25 classrooms were alternately assigned to the control group, where they experienced their school year like any other preschool classroom in Newark. Consistent with accepted practice in the field, block assignment is included as a covariate in all impact analyses to account for the restrictions on random assignment across blocks. (The analytic approach is discussed further below.)

All children in the 51 participating classrooms were eligible to be part of the study sample, and both the lead and the assistant teachers in the FOL-assigned classrooms were invited to participate in the teacher trainings. Table 2.1 presents a timeline of data collection: In short, the data include information on classrooms, teachers, and children and were collected at multiple times during the preschool year and the follow-up year. Information on children was collected from three sources: parents (at baseline) and teachers and trained observers (at both baseline and the preschool follow-up year). All participating sites were included in the analysis for observed ratings of classrooms and children, resulting in 100 percent participation in these key components of the study. For teacher-reported outcomes, response rates depended on parental consent and teacher participation in the data collection effort, both of which achieved benchmarks of acceptability. See Appendix A for more information on parental consent for research participation and the survey response rates.

The Research Sample: Baseline Characteristics of Teachers, Classrooms, and Students

Table 2.2 presents baseline characteristics of lead teachers and classrooms in the research sample. Not surprisingly, given the random assignment nature of the study, no statistically significant differences appear between classrooms and teachers that were assigned to the program group and those that were assigned to the control group. Teachers were, on average, about 37 years of age, and about half had taught preschool for six or more years at baseline. Consistent with Abbott requirements, all lead teachers had a bachelor's degree or higher.[3] The sample is predominantly female.

As shown in the second panel of Table 2.2, more than half of classrooms consisted of predominantly African-American children, and one-fourth consisted of predominantly Hispanic children. Participating classrooms averaged an enrollment of 14 children — well within Abbott class-size mandates. While classrooms could be characterized as higher quality based on these

[3]Due to a chronic illness of the lead teacher in one intervention site, the assistant teacher completed the self-survey and reports on children.

The Foundations of Learning Demonstration

Table 2.1

Data Collection Timeline

Type of Data	Preschool Fall Sept.-Oct. 2007	Preschool Winter Feb. 2008	Preschool Spring Apr.-May 2008	Postintervention Follow-Up Mar.-Apr. 2009
Parent survey	X			
Teacher report on children	X		X	X
Teacher self-survey	X		X	
CLASS[a]	X	X	X	X
inCLASS[b]			X	

SOURCE: MDRC.

NOTES: [a]CLASS = Classroom Assessment Scoring System. March-April 2009 classroom observations were done for the program group teachers only.

[b]inCLASS = Individualized Classroom Assessment Scoring System.

dimensions of teacher credentials and teacher-student ratio, they scored surprisingly low, with an average score of 4, on the Space & Furnishings subscale of the ECERS-R, a standardized observational measure that assesses the size, furnishings, and arrangement of the classroom and available outdoor equipment and space for play.[4]

Consistent with these classroom findings, Table 2.3 illustrates that there also were few differences between *children* in the program and control group classrooms at baseline. The sample is roughly split between boys and girls who averaged just over age 4 at the start of the preschool year. Their racial/ethnic composition, as reported by parents, depicts the diversity of Newark's population: over 40 percent black, nearly 10 percent white, and approximately 35 percent identifying as Hispanic. About half the children lived in a single-parent household, and approximately one-third lived in households with two or more other children. On average, the children's households were receiving more than one government benefit, such as housing assistance, Temporary Assistance for Needy Families (TANF), food stamps, Medicaid, the State Children's Health Insurance Program (SCHIP), or Social Security benefits.

[4]Cryer, Harms, and Riley (2003). A review of studies on comparable preschool populations shows a range of mean scores for the ECERS-R Space & Furnishings subscale, from 4.6 to 5.4. The studies include the Early Head Start Evaluation, Head Start Impact Study, and Abbott Preschool Longitudinal Effects Study.

The Foundations of Learning Demonstration

Table 2.2

Baseline Characteristics of Teachers and Classrooms

Characteristic	Program Group Mean	Control Group Mean	Difference	Standard Error
Characteristics of lead teachers[a]				
Female (%)	88.5	88.0	0.5	9.2
Age (years)	37.0	38.2	-1.3	2.8
Race/ethnicity (%)				
Black/African-American[b]	52.0	66.7	-14.7	14.7
Hispanic	23.1	17.4	5.7	11.8
White[b]	24.0	14.3	9.7	11.9
Taught preschool for 6 or more years (%)	53.8	56.5	-2.7	14.5
Holds bachelor's degree or higher (%)	96.2	100.0	-3.8	4.1
Characteristics of classrooms				
Racial/ethnic composition (%)				
Predominantly black/African-American	53.8	60.0	-6.2	14.1
Predominantly Hispanic	26.9	20.0	6.9	12.1
Predominantly Portuguese	15.4	4.0	11.4	8.3
Mixed	3.8	16.0	-12.2	8.3
Average score for space and furnishings[c]	4.1	4.1	0.1	0.2
Number of children enrolled in the class	13.8	13.6	0.2	0.4
Number of children present on an average day	12.4	11.7	0.6	0.5
Sample size - teachers	26	23		
Sample size - classrooms	26	25		

SOURCES: MDRC calculations from responses to teacher self-survey and classroom observations.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.
Because of missing data, the sample sizes for age, race/ethnicity, "Taught preschool for 6 or more years," and education, do not total 51.

[a]In one classroom, the assistant teacher acted as the lead teacher because of an illness of the lead teacher.
[b]This group includes only teachers not also reporting Hispanic.
[c]Based on the Early Childhood Environment Rating Scale average score for space and furnishings.

Outcome Measures

This section describes the measures of classroom quality and child behavior that were used to assess the impact of FOL on classrooms and individual children. Results of the impact analysis using these measures are presented in Chapters 3 and 4.

The Foundations of Learning Demonstration

Table 2.3

Baseline Characteristics of Students

Characteristic	Program Group Mean	Control Group Mean	Difference	Standard Error
Female (%)	48.6	48.0	0.6	4.0
Age (years)[a]	4.4	4.3	0.1 *	0.0
Race/ethnicity (%)[b]				
Black, not Hispanic	42.2	43.7	-1.5	3.9
White, not Hispanic	9.5	9.1	0.4	2.3
Hispanic	35.8	34.4	1.5	4.1
Other	0.4	1.2	-0.8	0.8
Number of benefits received in household	1.4	1.4	-0.1	0.1
3 or more children in the household (%)	33.1	30.4	2.8	3.7
Parent is 22 years old or younger (%)	7.1	6.4	0.7	2.2
Single-parent household (%)	47.8	50.0	-2.2	4.4
Primary language spoken at home is Spanish (%)	18.2	17.5	0.7	3.2
Sample size	319	304		

SOURCES: MDRC calculations from responses to teacher and parent surveys.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.
Means are adjusted for random assignment block but not for nesting of students within classrooms.
[a]Age at start of school year, September 2007, calculated from date of birth.
[b]Means do not equal 100 percent because the values represent adjusted means and race/ethnicity is not available for all students.

Classroom Quality

This study used a highly regarded measure of classroom quality — the *Classroom Assessment Scoring System* (CLASS) — to assess the impact of FOL on classroom outcomes. CLASS ratings are based on independent observers' detailed ratings of the quality of interactions between teachers and children in the classrooms.[5] Importantly, each rating is collected by independent observers who are blind to the random assignment status of the preschool classrooms. These observations were made in each of the 51 classrooms as a baseline measure in September 2007 (Appendix Table A.1) and at two follow-up points (winter and spring of the

[5]La Paro, Pianta, and Stuhlman (2004); Pianta, La Paro, and Hamre (2006).

preschool year). As discussed in Chapter 4, the program group teachers were observed again in the year following the intervention as well. Commensurate with standard practice in the field, 20 percent of the observations for each measure were double-coded to reduce the risk of rater drift and to ensure coder reliability. The interrater reliability for CLASS averaged 0.88 across the three time points in the intervention year. For each period of observation, observers watched the classroom for 20 minutes, followed by 10 minutes for compiling their notes into a 1-to-7 score on each of the CLASS dimensions. All ratings were calculated as average scores of these independent observers across four periods of observations on a single day, beginning first thing in the morning (typically with breakfast). Box 2.1 describes the individual CLASS dimensions.

In addition to looking at the individual dimensions as outcomes, two composite scores were created, based on the theoretically expected relationships among the dimensions and factor analytic work of the baseline CLASS data. Notably, these composites differ slightly from the domains offered by Pianta et al. (See Appendix Table B.1 for factor loadings.)[6] Specifically, dimensions included teachers' *positive classroom management*, which includes ratings of positive climate, negative climate, teacher sensitivity, and behavior management. A second composite score, the *quality of language instruction*, includes regard for student perspectives, use of engaging teaching methods, promoting understanding through conversation, and encouragement of students' language use. Composite scores in this report are averaged across dimensions that make up the total score. One remaining dimension — *management of classroom time* — was considered separately.[7]

Outcomes for Children

To understand how children were faring in the classroom, teachers assessed children's problem behavior and positive social interactions at the start of the school year and again in the spring of the preschool and kindergarten years. In addition, in spring of the preschool year and

[6]Pianta, La Paro, and Hamre (2006) CLASS domains are (1) Emotional Support: Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives; (2) Classroom Organization: Behavior Management, Productivity (referred to in this report as "management of classroom time"), and Instructional Learning Formats (referred to in this report as "use of engaging teaching methods"); (3) Instructional Support: Concept Development, Quality of Feedback (referred to in this report as "promoting understanding through conversation"), Language Modeling (referred to in this report as "encouragement of students' language use"); and (4) Student Outcomes: Student Engagement.

[7]Scores for the CLASS dimension Concept Development were not used in this evaluation because of lower reliability scores and reported coding difficulties from trained observers. Also, a final dimension, *overall classroom student engagement*, is discussed with the measures presented on children, as it represents an outcome of children's behavior rather than classroom quality per se.

Box 2.1

Observed Ratings of Classroom Climate

Average scores across observers across four cycles of observations on a single day were computed for each of these CLASS dimensions. In each case, scores ranged from 1 to 7.

Positive Classroom Management

Positive climate reflects the overall emotional tone of the classroom. Observers evaluate the relationships between teachers and students, positive affect or joint experiences of positive emotions (group laughter), the degree of respect the teacher shows toward the students, and interactions among the students.

Negative climate gauges the severity and intensity (duration) of both teacher- and peer-expressed negativity in the classroom. "Negativity" refers to sarcasm, anger, and disrespect by the teacher and to arguing, aggression, and bullying by students.

Teacher sensitivity measures the extent to which teachers comfort students, address students' needs (academic and emotional), and create age-appropriate activities. Observers evaluate whether a teacher effectively addresses students' concerns and notices when students need assistance and whether they feel comfortable enough to seek the teacher's support.

Behavior management reflects a teacher's ability to prevent student misbehavior by setting clear expectations, using effective praise to reinforce positive behavior, and proactively monitoring and redirecting possible problems. Time taken to deal with behavioral issues and occurrences of misbehavior are also factors.

Quality of Language Instruction

Regard for student perspectives gauges the degree to which the teacher's interactions with students and the activities prepared for the classroom support students' autonomy, expression of their own ideas, taking on responsibilities, and socialization with their peers. Teachers should show flexibility during lessons and should respond to student's ideas but should not attempt to overcontrol the students' movements.

Use of engaging teaching methods focuses on the teacher's preparation of lessons and activities and the mode with which these are delivered. Teachers should use a variety of different materials that promote exploration and inquiry and should actively question students; as a result, students should demonstrate active engagement in the tasks.

(continued)

Box 2.1 (continued)

Quality of Language Instruction (continued)

Promoting understanding through conversation assesses how a teacher phrases and uses feedback to promote understanding of a topic and not simply to express correctness. Teachers should prolong student-initiated conversations in an effort to get the student to garner a greater understanding of a specific topic or in order to give hints when a student provides an incorrect response.

Encouragement of students' language use measures the quality and amount of a teacher's use of language to stimulate and encourage student conversation and responses. Teachers should frequently converse with children using open-ended questions, repetition and extension, self- and parallel talk, and advanced language.

Management of Classroom Time

Management of classroom time measures the degree to which instructional time is effectively managed, including minimal downtime, through lesson planning, use of classroom routines, and the teacher's ability to minimize disruptions. Management of classroom time also takes into account the amount of time spent transitioning from one activity to another.

in spring of the kindergarten year, information was also collected from teachers about children's approach to learning. Box 2.2 describes these measures. The outcomes are evaluated in Chapter 3 (preschool outcomes) and Chapter 4 (kindergarten-year outcomes).

In addition to these teacher-reported outcomes for children, the *Individualized Classroom Assessment Scoring System* (inCLASS) observational tool was collected, assessing individual children using a rubric similar to CLASS.[8] These observations were collected during the preschool spring follow-up period for five preselected children in each classroom. Children were stratified by gender and by baseline behavior problems scores as reported by teachers. The goal was to ensure representation of both boys and girls at low, moderate, and high initial levels of behavior problems. As done with CLASS observations, 20 percent of the observations for this measure were double-coded to ensure coder reliability, resulting in an interrater reliability score of 0.88. Box 2.3 describes each dimension of this measure.

[8]Previously, inCLASS was titled "Classroom Assessment Scoring System — Child Version (CLASS-C)" (Downer et al., 2008).

Box 2.2

Teacher-Reported Outcomes for Children

Problem Behavior

The **Behavior Problem Index** assessed children's externalizing and internalizing problem behavior. The measure determines the frequency, range, and type of childhood problem behavior for children age 4 and older.[*] Teachers were asked to rate each of 30 items according to how characteristic it was of the child, using a 3-point scale ("not true," "sometimes true," "very true"). A total score was created by summing answers to the individual items. A 14-item *Externalizing Problems* subscale was created to assess the extent to which the child engaged in acting out and aggressive behaviors. A 14-item *Internalizing Problems* subscale assessed the extent to which the child was anxious or depressed. See Appendix B.2 for factor loadings and a list of all items.

The *Attention Problems* subscale of the **Caregiver-Teacher Report Form** (CTRF)[†] asked teachers to answer a series of questions about the child's behavior "now or within the past two months." The 3-point scale allowed teachers to report whether behaviors (for example, "fidgets," "wanders away," or "fails to carry out assigned tasks") occur often ("very true"), sometimes ("somewhat true"), or never ("not true"). The subscale score is an average of nonmissing items, and scores for this sample range from 0 to 18.

Teachers' perceptions of conflict and closeness with children were assessed with the **Student-Teacher Relationship Scale** (STRS).[‡] The subscales *Teacher-Student Conflict* and *Teacher-Student Closeness* (included in "Positive Social Behavior," below) include a total of 15 items that use a 5-point Likert-type rating scale to assess a teacher's perceptions of his or her relationship with a child, a child's interactive behavior with the teacher, and a teacher's beliefs about the child's feelings toward the teacher.

Positive Social Behavior

Teachers reported on a child's positive behavior in the classroom using the *Compliance with Teachers' Directives* (Compliance) and *Social Competence* (Social Competence and Sensitivity) subscales of the **Positive Behavior Scale** (PBS).[§] The teacher responds on a 5-point scale, ranging from 1 ("never") to 5 ("all of the time") for the 11-item Social Competence subscale (for example, "gets along well with other children," "shows concern for other people's feelings") and the 8-item Compliance with Teachers' Directives subscale (for example, "thinks before he/she acts," "usually does what I tell him/her"). Reported scores are an average of nonmissing items.

(continued)

Box 2.2 (continued)

Approach to Learning

Teachers assessed children's task engagement using the 16-item *Work-Related Skills* (Work-Related) subscale of the **Cooper-Farran Behavioral Rating Scales** (CFBRS).[‖] The full measure is designed for use by teachers in assessing classroom behavior, with teachers being asked to report on children's behavior during such classroom activities as "designated work time," and it has been used extensively with preschool and kindergarten children. Showing good predictive validity for children's later academic outcomes, the 7-point scale has descriptive phrases, which differ by item, to anchor responses to points. The scores shown for the Work-Related Skills (Work-Related) subscale used in this report are an average, not standardized, score of nonmissing items.

Preacademic Skills

Academic skills were assessed using the **Academic Rating Scale** (ARS).[#] The scale was designed to indirectly assess the process and products of children's learning in school. The 21 items are divided into three subscales: *General Knowledge* (5 items), *Language and Literacy* (9 items), *Mathematical Knowledge* (7 items). Teachers compare the target child with peers on a 5-point scale reflecting the degree ("not yet" to "shows proficiency") to which a child demonstrates skills, knowledge, and behaviors. Total and subscale scores were calculated as an average of nonmissing items.

SOURCES:
[*]Zill and Peterson (1986).
[†]Achenbach (1997).
[‡]Pianta (2001).
[§]Quint, Bos, and Polit (1997).
[‖]Cooper and Farran (1991).
[#]National Center for Education Statistics (n.d.).

Unlike the observations of the entire classrooms using the CLASS dimensions, the observations using the inCLASS measure assess individual preschool child behavior. Behaviors are scored across 10 dimensions that are grouped into three categories of classroom interactions: problem behavior, positive social behavior, and approach to learning.[9] For an inCLASS observation cycle, observers watch a specified child for 10 minutes, followed by 5 minutes for

[9]Downer et al. (2008) domains are classified into three categories, including (1) Teacher Interactions: Positive Engagement (referred to in this report as "teacher positive engagement"), Teacher Conflict, and Teacher Communication; (2) Peer Interactions: Peer Sociability, Peer Conflict, Peer Assertiveness, and Peer Communication; and (3) Task Orientation: Task Engagement, Self-Reliance, and Task Behavior Control.

Box 2.3

Observed Ratings of Child Behavior

Average scores across observers across four cycles of observations on a single day were computed for each of these CLASS dimensions. In each case, scores ranged from 1 to 7.

Problem Behavior

Teacher conflict measures the magnitude, consistency, and buildup of the child's interactions with the teacher when they are characterized by tension, resistance, and negativity. The degrees of verbal or physical aggression, noncompliance, negative affect, whining, and complaining directed toward the teacher are assessed.

Peer conflict gauges the magnitude, consistency, and buildup of the child's interactions with peers when they are characterized by tension, resistance, and negativity. The degrees of verbal or physical aggression, uncooperative behavior, negative affect, and whining and complaining — as well as whether the child sets children against one another, bullies, or is intrusive — are also evaluated.

Positive Social Behavior

Teacher communication reflects the child's communication with all adults. This includes the child's ability to initiate and sustain conversations and the overall use of language to convey needs and emotions as well as to communicate socially and to share opinions with teachers.

Teacher positive engagement assesses the degree to which the child is emotionally connected to teachers. This can be expressed by seeking out and enjoying interactions with teachers, the child's proximity to adults during tasks, and the child's display of shared positive emotions (for example, laughing, smiling) with teachers.

Peer communication reflects the child's communication with peers. Similar to teacher communication, this includes the child's ability to initiate, join, and sustain conversations and the overall use of language to convey needs and emotions as well as to communicate socially and to share opinions with peers.

Peer sociability measures the child's experiences of positive emotions and behaviors with peers. The child's desire to seek out and the willingness to respond to interactions with peers and the child's popularity are evaluated. Observers also assess the child's proximity to and conversation and eye contact with other children as well as his or her social awareness and ability to share.

Peer assertiveness assesses the child's positive strategies used to initiate peer interactions and the child's display of leadership and self-confidence during interactions. This includes the child's ability to join groups and communicate needs to peers as well as whether the child is imitated by peers.

(continued)

Box 2.3 (continued)

Approach to Learning

Task engagement measures the amount of time that the child remained actively engaged, focused, and on-task. Also evaluated are the child's ability to follow directions, enthusiasm for classroom tasks and activities, and willingness to participate voluntarily in these tasks.

Task self-reliance reflects the child's ability to seek out learning opportunities, work independently, and make the best use of classroom resources (including the teacher). The child's persistence during frustrating tasks and the ability to link concepts to personal experience are also assessed.

Task behavior control measures the child's ability to regulate movement, physical activity, and verbalizations so that they match the expectations of the classroom activity. This dimension also focuses on patience and the ability to keep one's hands to oneself and to respect others' personal space.

Overall classroom student engagement is the only student-focused dimension of the CLASS observations and captures students' focus and participation during activities, across all children in the classroom. A distinction is made between active and passive engagement and whether the engagement is sustained throughout the day.

compiling their notes into a 1-to-7 score for each of the 10 dimensions; 1 represents a low score, and 7 is the highest score. All outcomes in Box 2.3 were calculated as average scores across multiple cycles of observations in a single day. The first 10 dimensions in Box 2.3 are part of inCLASS, and the final dimension — overall classroom student engagement — was observed as part of the CLASS and represents an aggregate classroom measure.

The Analysis Strategy

The cluster-randomized experimental design utilized in this study represents the "gold standard" of evaluation research, ruling out the possibility that omitted variables — such as classroom quality or teacher skills — might bias the results.

The analyses presented in Chapters 3 and 4 compare regression-adjusted means of outcomes for FOL and control classrooms and children. Controls for block assignment are included in all regression models. In addition, a grand mean imputation strategy was used for missing teacher and student baseline characteristics in this analysis.

Conducting random assignment at the center level and including only one classroom per participating center allowed the use of a one-level model for the classroom- and teacher-level outcomes. The classroom-level regression includes controls for baseline characteristics, including random assignment block and baseline scores on CLASS dimensions.

Teacher Outcomes: One-Level Model

The model for the teacher outcomes adds covariates for selected baseline characteristics of teachers.

$$Y_j = \alpha + \beta_0 T_j + \sum_{k>0} \beta_k X_j + e_j$$

where:

Y_j = the outcome for classroom j at a given time

T_j = one for classrooms in the program group and zero otherwise

X_{kj} = the kth baseline characteristic for classroom j

e_j = a random error term for classroom j

α = the regression-adjusted mean outcome for classrooms in the control group

β_0 = the impact of the intervention on the outcome

β_k = a regression coefficient for the kth baseline characteristic

Child Outcomes: Two-Level Mixed Model

For the child outcomes, a two-level model is utilized to account for the nesting of children within classrooms (and centers). At the classroom level, the model controls for baseline characteristics, including random assignment block and baseline CLASS composite scores. At the child level, the model controls for age, gender, race/ethnicity, and baseline teacher-reported behavior problems score. To account for possible bias in teacher-reported outcomes, the following baseline characteristics were added to the model: years teaching preschool and teacher perceptions of job demand, classroom management skills, and stress.

$$Y_{ij} = \alpha + \beta_0 T_j + \sum_{k>0} \beta_k X_{kij} + e_j + \varepsilon_{ij}$$

where:

Y_{ij} = the outcome for student i from classroom j at a given time

T_j = one for all students from centers in the program group and zero otherwise

X_{kij} = the kth baseline characteristic for student i from classroom j

e_j = a random error term for classroom j

(continued)

ε_{ij} = a random error term for student i from classroom j

α = the regression-adjusted mean outcome for classroom in the control group

β_0 = the impact of the intervention on the outcome

β_k = a regression coefficient for the kth baseline characteristic

Chapter 3

The Impact of Foundations of Learning on Classrooms and Children in the Preschool Year

Chapter 3 presents impact findings for the Newark, New Jersey, site of the Foundations of Learning (FOL) demonstration. Specifically, these findings include an analysis of the effect of the intervention model (Chapter 1) on both the quality of preschool classrooms and the outcomes for individual children, during the year that FOL was implemented. These findings answer the primary questions of the demonstration: (1) whether teachers who were randomly assigned to receive FOL would be able to better structure emotionally positive, behaviorally supportive classroom environments than similar teachers who were randomly assigned to the control group; (2) whether teachers who were assigned to receive FOL showed improvements in their management of classroom time; and (3) whether preschool children were positively affected by the implementation of FOL in their classrooms.

The implementation findings discussed briefly in Chapter 1 indicate that the FOL intervention in Newark was implemented with fidelity and quality. That is, teachers were found to receive the training in the new strategies; they reported that this training was high quality; and they received consultation in the new strategies well above the control group levels. Therefore, comparisons of the data on classrooms and children assigned to FOL and those assigned to the control group provide a fair test of the model. Given this, it is reasonable to expect that the FOL classrooms might differ from control classrooms in teachers' ability to address children's behavioral adjustment. By extension, these changes might lead to differences in children's skills in managing their behavior and emotions in the classroom.

Findings in Brief

- **The FOL intervention improved teachers' ability both to address children's behavior and to provide a positive emotional climate in the classroom.** Based on information collected by independent observers, teachers in the FOL group showed significantly better skills in managing children's behavior problems and in providing an emotionally positive and supportive classroom climate than did their counterparts who were randomly assigned to the control group.

- **The FOL intervention also improved management of classroom time, use of engaging teaching methods, and the amount of instructional time, but it did not otherwise increase or decrease the quality of language in-**

struction that children received. These improvements in teachers' management of classroom time are consistent with the governing hypothesis behind the demonstration: namely, that children with challenging behaviors may divert teachers' attention away from providing instruction to managing behavior in the classroom. With regard to language instruction, teachers' engagement in behavioral management strategies does not appear to have either supported or undermined other aspects of their teaching.

- **FOL led to reductions in conflicts but did not otherwise change the quality of teacher-child or peer interactions; surprisingly, observers recorded these differences, but teachers did not.** Based on observations conducted by trained observers, children in FOL classrooms exhibited lower levels of behavior problems than did children in control group classrooms, but, notably, were not rated differently on other aspects of positive social interactions (communication or sociability, for example). However, somewhat surprisingly, *teachers* did not report differences in children's behavior between the program and control groups, even though the independent observation team did.

- **Children in the FOL intervention group demonstrated greater levels of engagement in the classroom than did children in the control group.** Children in FOL classrooms were rated by observers as higher on engagement in classroom activities and on their ability to regulate behavior during tasks than were children in control classrooms.

Classroom Quality

As discussed in Chapter 2, FOL-assigned and control-assigned classrooms in the demonstration were observed by trained observers using a standardized assessment tool, the Classroom Assessment Scoring System (CLASS), which rates the emotional and the instructional climate of educational settings. (See Chapter 2 and Box 2.1.) Observers — who were blind to classroom random assignment status and who were not part of the research team — rated classrooms using a 1-to-7 score for each of the CLASS dimensions, with 1 and 2 indicating low scores, 3 to 5 indicating moderate scores, and 6 and 7 indicating high scores. The impact of FOL was examined on each of the nine CLASS dimensions as well as on two composite scores.

Positive Classroom Management

The first goal and research question was to examine whether teachers who were randomly assigned to receive the FOL intervention would be better able to structure emotionally

positive, behaviorally supportive classroom environments than were similar teachers who were randomly assigned to the control group. The results are shown in Table 3.1, where average levels of classroom quality are reported for the FOL-assigned classrooms and for the control group classrooms. Stars (asterisks) in the third column indicate differences that are statistically significant, indicating that the differences between the two groups are not likely due to chance. The effect size allows one to compare the impacts in this study with those from other studies, yielding a measure of the magnitude of the impact for different measures. The effect size indicates the proportion of a change in standard deviation that is due to FOL.

As shown in Table 3.1, the FOL program group classrooms were rated significantly better than the control group classrooms on a composite measure of positive classroom management — a measure that reflects teachers' positive affect with children, lack of sarcasm and anger, and ability to comfort children and prevent misbehavior by setting clear expectations and using effective praise. It is important to note that the control classrooms scored at the high end of the moderate range of this measure (with scores just above 5), meaning that the FOL program group had a relatively high bar to exceed. These high scores are not surprising, given the Abbott resources and requirements for preschool in Newark that are discussed in Chapter 1. Nevertheless, the classrooms in the program group scored about half a point higher, on average (on a 7-point scale), than the control classrooms. Looking beyond the group averages, approximately two-thirds of program group classrooms have high scores on this composite measure (a score of 6 or 7), compared with about half the classrooms in the control group (not shown).

To better understand this finding, effects were examined on the separate dimensions of the CLASS that make up the composite measure. The two dimensions of positive classroom management that showed the strongest impacts were observers' ratings of (1) classroom negative climate (which is assessed by teachers' use of emotionally negative classroom management practices, such as yelling at children), with program group classrooms scoring at 1.1, on average, and their control counterparts with scores about a half point higher (at 1.7, on average); and (2) behavior management, with program classrooms' average scores at nearly 5.5 and control group classrooms below 5. Notably, both of these dimensions were targeted by FOL directly as part of the training that the teachers' received.

To put these effects in perspective, it is helpful to consider new findings on the points on the 1-to-7 continuum at which classroom quality is associated with improved outcomes for young children. The question is whether an improvement in the 5-to-6 range on the scale might matter in terms of classroom quality and outcomes for children. Fortunately, research studies find that it does, indeed, make a difference (in terms of associations with both social-emotional

The Foundations of Learning Demonstration

Table 3.1

Program Impacts on Observed Ratings of Classroom Climate and Instructional Time

Variable	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Positive classroom management					
Composite[a]	5.8	5.2	0.6 **	0.3	0.75
Positive climate	5.6	5.0	0.6	0.4	0.60
Negative climate	1.1	1.8	-0.6 ***	0.2	-0.90
Teacher sensitivity	5.2	4.8	0.4	0.3	0.46
Behavior management	5.4	4.7	0.8 **	0.4	0.72
Use of classroom time					
Management of classroom time	5.4	4.9	0.5 *	0.3	0.63
Amount of instructional time (minutes)	35.6	25.1	10.6 **	4.4	0.96
Sample size	26	25			

SOURCES: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007 and April-May 2008.

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

The table presents adjusted means that control for random assignment blocks and baseline (fall) CLASS dimension scores. For each dimension, observers rated classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

"Management of classroom time" refers to the "Productivity" dimension of the CLASS.

[a]"Negative climate" is reverse-coded for the composite score.

and academic outcomes for children) whether classrooms score a 5, a 6, or a 7.[1] Thus, raising the level of classroom quality at this range of the scale may, indeed, be important for outcomes for children. These findings suggest that the FOL intervention's ability to shift classroom quality substantially above a score of 5 on positive classroom management of the CLASS is substantively important.

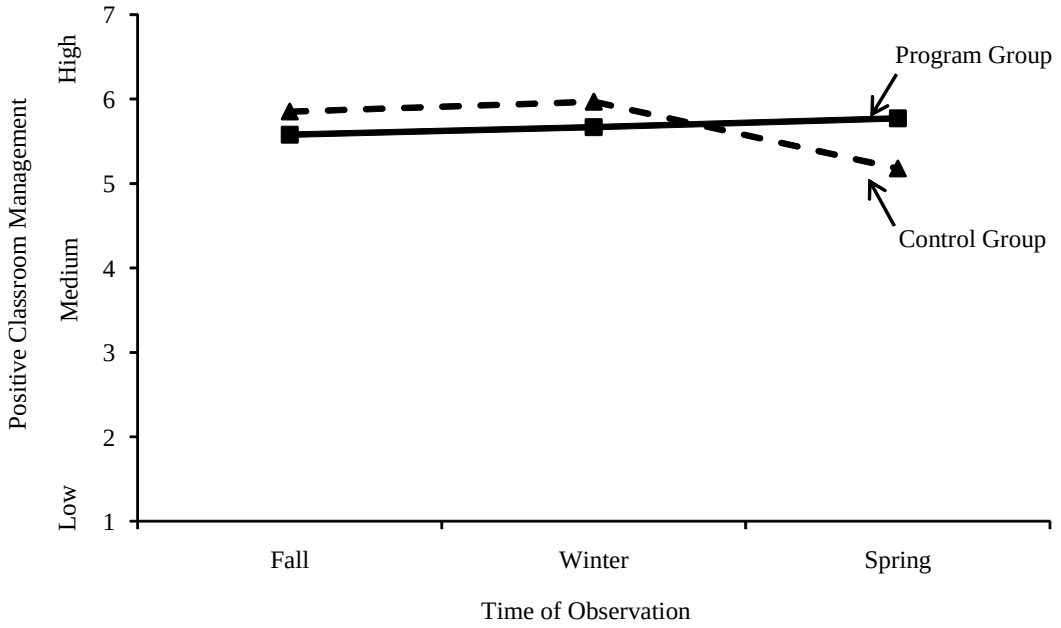
Figure 3.1 shows how classrooms scored on this composite measure of positive classroom management in the fall, prior to the implementation of the intervention; in the winter of the intervention year; and in the spring. Given the small sample, it is not surprising, even with random assignment, that program and control groups show small differences at baseline in the fall — with the control groups being slightly (but not statistically significantly) higher than the program groups. Note that the program group line stays relatively flat, indicating little change in

[1]Burchinal, Vandergrift, Pianta, and Mashburn (2010).

The Foundations of Learning Demonstration

Figure 3.1

Scores for Positive Classroom Management, Program and Control Groups During Preschool Year



SOURCES: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007, January-February 2008, and April-May 2008.

NOTES: Scores represent adjusted means that control for random assignment block. Winter and spring control for baseline fall CLASS dimensions as well.

Spring scores for the program group and the control group are statistically different (p-value = less than 5 percent).

positive classroom management. By contrast, the control group classrooms decline in this dimension of quality from winter to spring. Thus, FOL was primarily effective at helping teachers *maintain* a positive classroom climate over the course of the year. These findings are quite consistent with other studies that show that classroom quality, indeed, declines over the course of a year, as teachers and children develop more negative styles of interacting over time.

Use of Classroom Time

The second research question asked whether teachers who received the FOL services would show improvements in their productive use of time in the classroom, even though this

area of instruction was not directly targeted by the FOL intervention. Evidence here suggests that targeting emotionally and behaviorally supportive classroom practices through FOL led to statistically significant benefits for teachers' use of productive time in the classroom. As shown in the second half of Table 3.1, program group teachers were rated higher on the extent to which they *managed* classroom time, including better planning, use of classroom routines, and ability to minimize disruptions; FOL classrooms scored about a half a point higher on the 7-point scale than control classrooms.

These findings are illustrated in Figure 3.2, which shows that management of classroom time increased in the program group classrooms from winter to spring, while it decreased in the control classrooms. Again, with such a small sample, small differences emerge at baseline that are controlled for in analyses conducted on outcomes in spring.

In addition to these standardized ratings of quality, observers also assessed the amount of time that teachers actually spent in leading small- and large-group instruction. Consistent with the higher ratings for teachers' use of productive time in the classroom, instructional time was significantly higher in the FOL classrooms, by an average of 10 minutes. In a 120-minute observation period, an average of 35 minutes was spent in teacher-led instruction in FOL classrooms, compared with 25 minutes in control group classrooms. This significant gain — a 40 percent change — may be a result of fewer disruptions by children during large-group activities and a reduction in the amount of transition downtime between activities, although this was not explicitly tested in these analyses. While 10 minutes may not seem like a great deal of time, it is important to remember that this represented a difference when the FOL and control classrooms were compared on a single day of observation, for only 120 minutes total. If such gains were representative of gains achieved every weekday of a 40-week year, this would translate to 50 minutes more of instruction a week for children in FOL classrooms, or the equivalent of a week's more instruction over the course of a school year than children in control classrooms received.

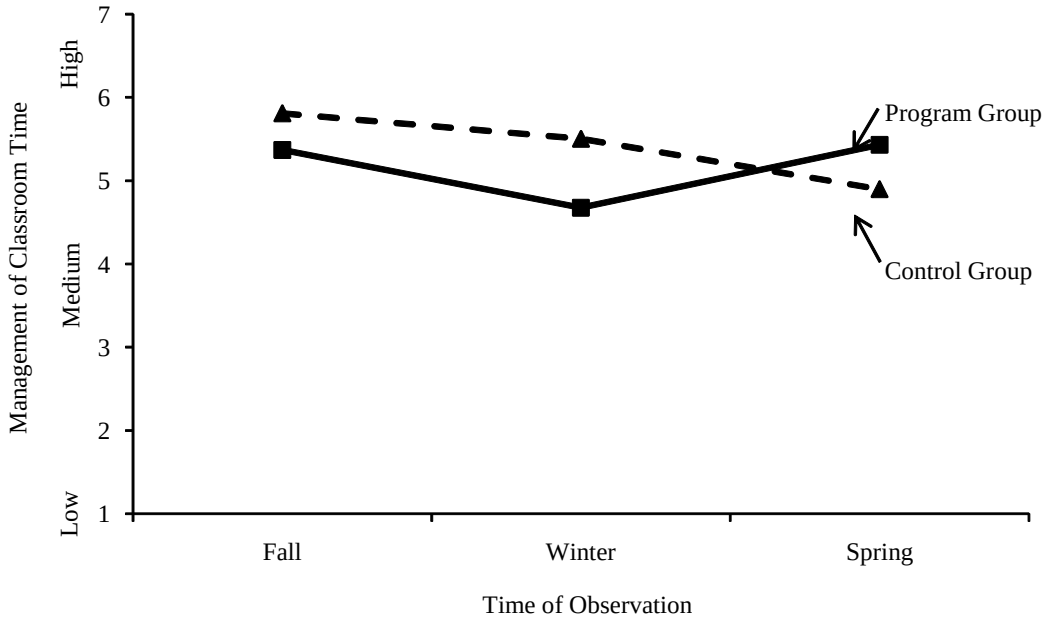
To place these findings in perspective, a major issue in the quality debate has been the recognition that there is a great deal of time lost in preschool classrooms to tedious, off-task activities, such as getting children organized to complete small- and large-group projects, getting children to line up, and so on.[2] These findings of an increase in the number of minutes spent in teacher-led instruction and an improvement in the ways that teachers use time more productively suggest that there are clear, concrete steps that centers can take to address this facet of classroom quality.

[2]Early et al. (2010); Pianta et al. (2005).

The Foundations of Learning Demonstration

Figure 3.2

Scores for Management of Classroom Time, Preschool Year



SOURCES: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007, January-February 2008, and April-May 2008.

NOTES: "Management of classroom time" refers to the "Productivity" dimension of the CLASS.

Scores represent adjusted means that control for random assignment block. Winter and spring control for baseline fall CLASS dimensions as well.

Scores for the program group and the control group are statistically different (p-value = less than 5 percent for the fall and winter; p-value = less than 10 percent for the spring).

In addition to these key aspects of teachers' behavior, the CLASS also includes measures of the instructional quality that teachers provide to children. While not a primary outcome of the intervention model, it was important to test whether, on the one hand, greater behavioral management led to different kinds of language used with children during instructional activities or whether, on the other hand, focusing on emotional and behavioral adjustment interfered with instructional support for children (which would result in reductions in quality of instruction in FOL-assigned classrooms).

Table 3.2 presents these results. Teachers who were randomly assigned to receive the FOL intervention scored higher on an instructional measure assessing their use of engaging

The Foundations of Learning Demonstration

Table 3.2

Program Impacts on Observed Ratings of Language Instruction

Variable	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Quality of language instruction					
Composite	4.3	3.8	0.5	0.3	0.56
Regard for student perspectives	5.1	4.9	0.2	0.3	0.28
Use of engaging teaching methods	4.2	3.5	0.6 *	0.3	0.61
Promoting understanding through conversation	3.5	3.0	0.4	0.4	0.44
Encouragement of students' language use	4.3	3.6	0.7	0.5	0.54
Sample size	26	25			

SOURCES: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007 and April-May 2008.

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

The table presents adjusted means that control for random assignment blocks and baseline (fall) CLASS dimension scores. For each dimension, observers rated classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

"Use of engaging teaching methods" refers to the "Instructional learning formats" dimension of the CLASS. "Promoting understanding through conversation" refers to the "Quality of feedback" dimension of the CLASS. "Encouragement of students' language use" refers to the "Language modeling" dimension of the CLASS.

teaching methods but otherwise did not differ from control group teachers in other aspects of the quality of language instruction they were providing. That is, on the composite measure and most of the individual measures of instructional quality, there were no significant differences between classrooms assigned to the FOL intervention and those assigned to the control group.

These findings help rule out the possibility that FOL-like interventions represent a potential threat to the quality of cognitively oriented instruction. The findings suggest that teachers' implementation of improved classroom management practices did not represent a zero-sum game whereby time might have potentially been lost to instruction as teachers focused more time and attention on behavioral management. On the contrary, given the findings on management of classroom time discussed above, these findings suggest that there are crossover effects of emotionally and behaviorally oriented intervention efforts on teachers' provision of longer, more organized, and more productive classroom instruction.

Outcomes for Children

The third research question was whether FOL would also affect the behavior of the pre-school children in the intervention-assigned classrooms. To examine this, a separate observation team was sent to rate individual children's behavior in the classroom. In each of the 51 study classrooms, these raters observed five children, ultimately including an equal number of boys and girls and children rated in the fall by teachers at low, moderate, and high levels of behavior problems. Observers rated children on conflictual interactions with teachers and peers; on positive social interactions with peers and teachers; and, finally, on their engagement in the learning tasks of school (their approach to learning). As with the classroom observations, ratings ranged from a low of 1 to a high of 7, with scores of 3 to 5 being considered in the moderate range. (Chapter 2 more fully describes these outcome measures.)

Findings are shown in Table 3.3. Ratings of problem behavior are in the low range in all classrooms — with scores, on average, just above 1 (the lowest rating on a scale). Such low scores are not surprising. Conflict is a rare but highly salient event in preschool classrooms, and observing such behavior in the course of one 80-minute observation period is unusual. Yet, even with these overall low ratings, FOL intervention classrooms were observed to have statistically lower levels of conflict, on average. For example, ratings on teacher conflict were nearly 1.5 in control group classrooms, while FOL-group classrooms had ratings about a quarter point lower. Effects on peer conflict were similar. It was thought that these findings might have implications for reducing teachers' feelings of stress in the classroom, yet there is no evidence that FOL reduced work-related stress or burnout, on a limited set of measures collected from teachers used to address these issues. (See Appendix D.)

In contrast, no differences were found between program and control classrooms on ratings of children's sociability with peers and positive engagement with teachers. This lack of effects on the positive aspects of teacher-child and peer interactions is disappointing, given that an aim of this intervention was not only to reduce conflict but also to improve children's relationships.

As shown in the bottom panel of Table 3.3, children in FOL classrooms were rated higher on measures of their approach to learning than children in control classrooms. Children in FOL classrooms were scored about a quarter of a point higher both on their engagement in activities and on their ability to control their behavior during those activities. In this case, statistically significant differences are observed in the overall classroom level of engagement as well as across all children in the classroom, indicating a higher level of focus and participation

The Foundations of Learning Demonstration

Table 3.3

Program Impacts on Observed Ratings of Child Behavior

Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Problem behavior					
Teacher conflict	1.2	1.5	-0.2 ***	0.1	-0.40
Peer conflict	1.4	1.6	-0.2 *	0.1	-0.27
Positive social behavior					
Teacher communication	2.2	2.4	-0.2	0.1	-0.20
Teacher positive engagement	3.2	3.4	-0.2	0.2	-0.27
Peer communication	2.5	2.6	-0.1	0.2	-0.14
Peer sociability	3.4	3.5	-0.1	0.1	-0.11
Peer assertiveness	2.1	2.3	-0.2	0.2	-0.21
Approach to learning					
Task engagement	4.9	4.6	0.2 *	0.1	0.31
Task self-reliance	3.1	3.1	-0.1	0.2	-0.07
Task behavior control	5.4	5.1	0.3 *	0.2	0.34
Overall classroom student engagement	5.7	5.2	0.6 *	0.3	0.60
Sample size - students	130	121			
Sample size - classrooms	26	25			

SOURCE: Based on MDRC calculations of classroom observations in April-May 2008.

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

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

The outcomes "Problem behavior," "Positive social behavior," and "Approach to learning" come from the inCLASS observations. "Overall classroom student engagement" comes from the CLASS. For each dimension, observers rated children and classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

among all children during classroom activities.[3] To give a sense of the magnitude of these effects, 75 percent of program group classrooms had high scores on student engagement, while only half of the control group classrooms received such high scores. If children are able to spend more time on task, they are likely to be able to take greater advantage of the formal and informal learning opportunities in the preschool classroom.

In addition to having unbiased, independent observers rate children's behavior, the study also asked teachers to report on children's positive and problematic behaviors, their

[3]Because different teams of observers were deployed to focus on the overall classroom environment and on individual children, the consistency across these findings is noteworthy.

approach to learning, and their preacademic skills. These findings are presented in Table 3.4. Surprisingly, teachers did not report differences in children's behavior between the program and control groups, even though the independent research team did (as presented above). Whether examining behavioral outcomes or approach to learning, no statistically significant differences are seen between the two groups of classrooms.

However, when teachers were asked to rate the prevalence of challenging behaviors in their classrooms — how often, on a 1-to-3 scale they observed a set of 28 behavior problems in the classroom as a whole — there is a statistically significant difference between FOL and control classrooms (not shown in table). Average scores on this 4-point scale were 2.3 in the control group classrooms, compared with 1.9 in the program group classrooms (with an effect size of nearly 1 standard deviation). It is not clear why FOL teachers report a difference in the level of problem behavior in their *classrooms* relative to control teachers while no differences are found between FOL and control teachers when teachers are reporting on *individual* children.

The lack of findings on the teacher-reported outcomes is somewhat surprising, in that other studies of social-emotional enhancements in preschool have shown that successful interventions typically change teachers' perceptions as well as observed aspects of behavior. One hypothesis is that the training that teachers in FOL received primed them more to see challenging behaviors, even as it increased their capacity to effectively manage these behaviors when they occurred. The fact that teachers do report that their classrooms as a whole were less behaviorally problematic but that individual children were not lends support to this hypothesis, in that, when thinking about individual children, FOL-assigned teachers did not report lower levels of problem behaviors.

In comparing impacts reported by the two data sources, one concern was that the divergent findings may be due to the fact that the independent observations were collected on a subset of children in each classroom whereas teacher-reported data were collected on all children whose parents consented to their participation in the demonstration. Therefore, additional analyses were conducted to confirm whether these differences in impacts are due to the different samples or whether they reflect differences in findings across sources. Appendix Tables C.1 and C.2 show the impacts on observed child behavior and on teacher-reported outcomes for those children who are included in both data sources. Results parallel those in Tables 3.3 and 3.4, indicating that the findings are not due to differences in the samples of children that were assessed via the two types of reporters.

Teachers were also asked about children's need for, and referral to, special services. Although no statistically significant differences were found between children in FOL and control

The Foundations of Learning Demonstration

Table 3.4

Program Impacts on Teacher-Reported Child Outcomes

Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Problem behavior					
Internalizing problems	2.7	2.3	0.4	0.6	0.11
Externalizing problems	4.1	3.7	0.4	0.7	0.08
Teacher-student conflict	12.4	12.3	0.1	0.9	0.02
Attention problems	3.6	3.5	0.1	0.6	0.02
Positive social behavior					
Compliance with teachers' directives	4.0	4.0	0.1	0.1	0.08
Social competence	4.0	4.0	0.0	0.1	0.06
Teacher-student closeness	34.5	35.8	-1.3	0.9	-0.24
Approach to learning					
Work-related skills	4.8	4.8	0.1	0.1	0.08
Preacademic skills					
Language and literacy skills	35.1	32.6	2.5	1.7	0.27
Math knowledge	25.8	25.4	0.4	1.7	0.05
General knowledge	19.7	18.5	1.3	1.0	0.28
Sample size - students	283	248			
Sample size - classrooms	26	23			

SOURCE: Based on MDRC calculations from responses to teacher survey.

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

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

Outcome controls for the child's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

"Internalizing problems" and "Externalizing problems" refer to the internalizing and externalizing subscales of the BPI. BPI internalizing and externalizing scales were created based on factor analysis work. "Teacher-student conflict" and "Teacher-student closeness" refer to the Student-Teacher Relationship Scale conflict and closeness subscales. "Attention problems" refers to the Caregiver-Teacher Report Form attention problems subscale. "Compliance with teachers' directives" and "Social competence" refer to the Positive Behavior Scale compliance and social competence subscales. "Work-related skills" refers to the Cooper-Farran Behavioral Rating Scales work-related skills subscale. "Language and literacy skills," "Math knowledge," and "General knowledge" refer to the language and literacy, math knowledge, and general knowledge subscales of the Academic Rating Scale.

group classrooms, the proportion of children about whom teachers had concerns is noteworthy (not shown). Across both groups, while only about 4 percent of children were *referred* for services, nearly 30 percent were reported to be of *concern* because of developmental issues. About half that percentage were of concern regarding social-emotional development, suggesting that this population remains at risk.

Finally, analyses were conducted to detect whether the impacts presented above differed for children at different levels of behavioral risk at the start of the school year (see Table 3.5). The expectation was that FOL may have a different pattern of effects for children with higher versus lower levels of behavior problems at the beginning of the year, as reported by teachers. More specifically, FOL was expected to have stronger reductions in problematic behavior for children with higher levels of behavior problems and stronger improvements in their approach to learning (that is, task engagement) for children with lower levels of problems. Indeed, as shown in Table 3.5, the impacts on independent observations of peer conflict are statistically significantly different for the higher-risk children. The FOL intervention helped decrease the observed conflict for these higher-risk children, while there are no statistically significant differences between program and control group children for the lower-risk subgroup. No differences emerged for this subgroup for other observed or teacher-reported outcomes, however, including observations of children's approach to learning, where there was an effect of FOL for the full sample (higher levels of task engagement among children in FOL classrooms than among children in control classrooms).

The analysis also examined two other exploratory subgroups — defined by child's gender (for boys and for girls) and by child's race/ethnicity (for Hispanic and for black, non-Hispanic). See Appendix Tables C.3 and C.4.

The Foundations of Learning Demonstration

Table 3.5

Program Impacts on Observed and Teacher Ratings of Child Outcomes, by Level of Behavior Problems

	Low					High					
Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	H-Stars
Observations											
Problem behavior											
Teacher conflict	1.3	1.3	0.0	0.1	0.04	1.4	1.6	-0.2	0.1	-0.40	
Peer conflict	1.5	1.5	0.0	0.2	-0.06	1.5	1.8	-0.3 **	0.1	-0.51	
Positive social behavior											
Teacher communication	2.1	2.5	-0.4 *	0.2	-0.47	2.3	2.3	0.0	0.2	-0.01	
Teacher positive engagement	3.1	3.5	-0.4	0.2	-0.46	3.2	3.3	-0.1	0.3	-0.13	
Peer communication	2.6	2.6	0.0	0.3	-0.05	2.5	2.6	-0.1	0.2	-0.18	
Peer sociability	3.5	3.7	-0.2	0.2	-0.24	3.4	3.6	-0.2	0.2	-0.23	
Peer assertiveness	2.0	2.3	-0.3	0.4	-0.29	2.2	2.2	-0.1	0.2	-0.09	
Approach to learning											
Task engagement	4.9	4.8	0.2	0.2	0.19	4.8	4.6	0.2	0.2	0.27	
Task self-reliance	2.9	3.1	-0.1	0.4	-0.09	3.1	3.2	-0.1	0.3	-0.08	
Task behavior control	5.6	5.2	0.3	0.3	0.36	5.1	4.8	0.3	0.2	0.35	
Teacher reports											
Problem behavior											
Internalizing problems	1.6	1.2	0.4	0.5	0.18	4.0	2.6	1.4	1.1	0.39	
Externalizing problems	1.7	1.7	0.0	0.6	-0.01	6.9	5.2	1.7	1.3	0.28	
Teacher-student conflict	9.6	10.0	-0.4	0.9	-0.10	15.3	13.5	1.9	1.4	0.26	
Positive social behavior											
Social competence	4.3	4.2	0.1	0.2	0.09	3.7	3.8	0.0	0.1	-0.06	
Teacher-student closeness	35.8	35.8	0.0	1.1	0.01	33.5	35.9	-2.4 *	1.2	-0.39	

(continued)

Table 3.5 (continued)

Outcome	Low: Program Group Mean	Low: Control Group Mean	Low: Difference (Impact)	Low: Standard Error	Low: Effect Size	High: Program Group Mean	High: Control Group Mean	High: Difference (Impact)	High: Standard Error	High: Effect Size	H-Stars
Teacher reports (continued)											
Approach to learning											
Work-related skills	5.1	5.2	-0.1	0.2	-0.09	4.5	4.6	-0.1	0.2	-0.08	
Preacademic skills											
Language and literacy skills	36.1	34.9	1.2	1.7	0.13	32.8	32.5	0.3	2.0	0.03	
Math knowledge	26.7	27.3	-0.6	1.8	-0.08	23.8	25.3	-1.5	1.9	-0.21	
Sample size - observations of students	50	40				58	38				
Sample size - teacher reports on students	125	94				128	100				
Sample size - classrooms[a]	26	25				26	25				

SOURCES: Based on MDRC calculations of classroom observations in April-May 2008 and from responses to teacher survey.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent. The final column lists the H-stars, which show the statistical significance of the difference between the subgroup impacts. The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

The observed outcomes "Problem behavior," "Positive social behavior," and "Approach to learning" come from the inCLASS observations. For each dimension, observers rated children and classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

Teacher-reported outcomes control for the student's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

Subgroup was created by calculating a median split of baseline teacher-reported BPI scores across the entire sample. BPI scores higher than a 4 fall into the "high" category. Baseline BPI scores were not available for all children. Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

"Internalizing problems" and "Externalizing problems" refer to the internalizing and externalizing subscales of the BPI. BPI internalizing and externalizing scales were created based on factor analysis work. "Teacher-student conflict" and "Teacher-student closeness" refer to the Student-Teacher Relationship Scale conflict and closeness subscales. "Attention problems" refers to the Caregiver-Teacher Report Form attention problems subscale. "Compliance with teachers' directives" and "Social competence" refer to the Positive Behavior Scale compliance and social competence subscales. "Work-related skills" refers to the Cooper-Farran Behavioral Rating Scales work-related skills subscale. "Language and literacy skills," "Math knowledge," and "General knowledge" refer to the language and literacy, math knowledge, and general knowledge subscales of the Academic Rating Scale.

[a]For teacher-reported outcomes, the sample size is 23 for the control group classrooms.

Chapter 4

One Year Later: Outcomes for Children and Teachers After the Implementation of Foundations of Learning

The impact findings that are reported in Chapter 3 show the short-term benefits of training preschool teachers in the Newark, New Jersey, site of the Foundations of Learning (FOL) intervention to better support children's behavior and emotional development. Specifically, at the end of the intervention year, the FOL model (described in Chapter 1) had improved the quality of the program classrooms relative to the control classrooms, and children in the FOL classrooms demonstrated reduced conflict and greater task engagement than children in the control classrooms, although they were not more socially competent.

Chapter 4 explores whether the effects of FOL extended to the following school year, in two ways. First, the chapter presents information about whether FOL led to differences in *children's* behavior and academic skills as they transitioned into their next school environment, typically kindergarten. Second, the chapter examines whether, for *teachers*, the investments that were made in improving the quality of classrooms might be sustained over successive cohorts of students, across multiple school years. Notably, the FOL demonstration was not funded to answer either of these questions comprehensively. However, the study is able to provide initial answers to these questions by capitalizing on teacher-reported data that were collected from children's kindergarten teachers[1] and on observations of FOL classrooms one year later.

Findings in Brief

- **Based on reports of kindergarten teachers, overall FOL had very few sustained effects on children during the year after they received the intervention.** On measures of children's positive social interactions and their approach to learning, there are no statistically significant differences between children who were in the FOL classrooms in preschool and those who were in the control classrooms. Surprisingly, the only significant effects were observed in the reverse direction: Higher levels of problem behavior were reported by teachers of children from the FOL classrooms than by teachers of children from the control classrooms. It is possible that sustained effects on children of a preschool intervention depend on the quality of the

[1]In a few cases, these were preschool teachers (n = 24) and first-grade teachers (n = 5), since some children transitioned to these classrooms instead of to kindergarten classrooms.

classrooms that they enter in these early elementary years, which was not assessed in this study.

- **Impacts in kindergarten differed somewhat depending on children's initial level of behavior problems.** For children with the lowest levels of behavior problems, there were no effects on problem or positive social behavior, but significant *positiv*e effects were found on children's work-related skills (a measure of student engagement) and their language and literacy skills. For the children with more elevated behavior problem scores in the fall of preschool, there is evidence of *increased* withdrawn and sad behavior — but no other statistically significant differences among children who received FOL in preschool and those who did not.

- **When observed in the year following the intervention, teachers who were assigned to the FOL classrooms appear to have continued to engage in the positive practices that they had learned as a result of the intervention.** When comparing teachers' scores on positive classroom management from the spring of the preschool year and from the following spring, scores were largely maintained for the FOL-assigned teachers. This is very good news, as there was concern that teachers might not continue to utilize what they had learned in the prior year without the direct support of the intervention.

Were Child-Level Improvements Sustained? A Preliminary Look, One Year Later

As described in Chapter 2, information on children's behavioral and academic skills was collected from the children's kindergarten teachers during the year following the implementation of FOL. Before turning to the findings, what was the expectation regarding the long-term effects of the intervention? In short, would preschoolers' gains be sustained as they made the transition to kindergarten?

From an optimistic perspective, one hypothesis might be that children with initial behavioral difficulty benefited so substantially from improved classroom climates in their preschool year that those children would carry improved behavioral adjustment with them, into new classroom contexts. That is, children may be able to elicit more positive interactions from teachers having experienced the FOL intervention during preschool. Moreover, even more behaviorally well-adjusted children may have benefited from greater opportunities for learning in more well-managed FOL classrooms, compared with similar children in control group

classrooms. It was, indeed, the expectation that children's gains in preschool might be sustained into the next year.

From a more skeptical perspective, there might be many reasons to expect that intervention impacts might not be sustained. First, children's behavioral gains in the preschool year may be context dependent. That is, improvements in behavior may reflect those children's greater ability to work with specific teachers who have also improved their classroom management skills. Behavioral skill may not be as much like luggage or a toolkit as one might hope: It may be the case that children are not able to "carry" those new skills to new relationships with teachers who may range widely in their effectiveness in classroom management.

Moreover, kindergarten may represent a particularly challenging transition: Compared with preschool classrooms, kindergarten classrooms comprise a larger number of children and a smaller number of teachers, and they are housed in larger buildings. Elementary schools generally focus less attention on supporting a positive emotional and social climate and have greater expectations of independence and discipline than preschool classrooms. In short, like any group of preschoolers, children who were enrolled in FOL likely entered a range of school settings on their first weeks in kindergarten, with some children entering emotionally supportive and well-managed classroom settings and other children entering more chaotic or emotionally and socially negative classroom settings. Against such a backdrop of variation in classroom quality, children's FOL-induced gains might not be sustained.

Children in the study were tracked to their next school environment, and their teachers were surveyed with regard to the children's problem behavior (acting out and withdrawn behavior), their positive social behavior with teachers and peers, their approach to learning, and their early academic skills. As discussed in Appendix A, children were dispersed to a large number of schools (about 100) for their kindergarten years, and FOL and control children were together in a large number of the kindergarten classrooms.

Notably, at this time point in the study, conclusions are based solely on kindergarten teacher reports, representing some advantages and disadvantages. On the one hand, these teachers were largely blind to children's random assignment status, which makes them less biased reporters of behavior than the reports from teachers the year earlier. On the other hand, teachers' reports are still influenced by their own experiences and beliefs, which makes them somewhat less reliable than information that would be collected from trained observers; that is, there is greater measurement error. It is not clear whether information from trained observers would corroborate the results reported here by teachers.

Table 4.1 shows the overall findings for all children in their kindergarten year. On most measures, there are no statistically significant differences on outcomes for children assigned to

The Foundations of Learning Demonstration

Table 4.1

Program Impacts on Teacher-Reported Child Outcomes, Kindergarten Year

Outcome	Program Group Mean	Control Group Mean	Difference (Impact)		Standard Error	Effect Size
Problem behavior						
Internalizing problems	3.24	2.45	0.79	*	0.39	0.24
Externalizing problems	4.73	3.77	0.96	*	0.56	0.16
Teacher-student conflict	12.30	11.55	0.75		0.66	0.11
Attention problems	4.45	4.34	0.11		0.51	0.02
Positive social behavior						
Compliance with teachers' directives	3.95	4.04	-0.08		0.08	-0.10
Social competence	3.98	3.96	0.02		0.09	0.03
Teacher-student closeness	33.13	33.90	-0.78		0.61	-0.14
Approach to learning						
Work-related skills	4.82	4.87	-0.04		0.13	-0.04
Academic skills						
Language and literacy skills	33.79	33.50	0.29		1.09	0.03
Math knowledge	26.56	26.55	0.02		0.72	0.00
General knowledge	19.09	19.26	-0.17		0.55	-0.03
Sample size	259	230				

SOURCE: Based on MDRC calculations from responses to teacher survey.

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

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

Outcome controls for the child's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

"Internalizing problems" and "Externalizing problems" refer to the internalizing and externalizing subscales of the BPI. BPI internalizing and externalizing scales were created based on factor analysis work. "Teacher-student conflict" and "Teacher-student closeness" refer to the Student-Teacher Relationship Scale conflict and closeness subscales. "Attention problems" refers to the Caregiver-Teacher Report Form attention problems subscale. "Compliance with teachers' directives" and "Social competence" refer to the Positive Behavior Scale compliance and social competence subscales. "Work-related skills" refers to the Cooper-Farran Behavioral Rating Scales work-related skills subscale. "Language and literacy skills," "Math knowledge," and "General knowledge" refer to the language and literacy, math knowledge, and general knowledge subscales of the Academic Rating Scale.

the program group classrooms when they were in preschool and those assigned to the control group classrooms. The only significant effects are in the reverse direction: Teachers reported higher levels of externalizing and internalizing behavior problems among children from the FOL classrooms than among children from the control classrooms.

As with the analyses presented in Chapter 3, it is important to examine whether FOL offered different benefits for groups of children who had higher and lower levels of behavioral problems at the beginning of the preschool year, based on teacher reports. That is, FOL may confer particular advantages in reducing the risk of later school difficulty for children who might initially show the highest levels of behavioral difficulty, while it might confer the greatest opportunities for learning for those children who are more well adjusted emotionally or behaviorally at entry into preschool. During the preschool year, the impacts of FOL on conflict were found primarily for children with the highest levels of behavioral difficulty, while impacts of FOL on children's approach to learning (engagement) did not differ across these two groups of children (see Chapter 3).

Findings are shown in Table 4.2. When splitting the sample by children's initial level of behavior problems, slightly different patterns of effects are seen for the two groups of children, but there are still very few effects overall. For the children with low levels of behavior problems, FOL had no effects on their problem behavior or positive social behavior. However, significant positive effects can be seen on children's work-related skills (a measure of their approach to learning) and on their language and literacy skills. These effects are statistically significantly different than the effects for the children with high levels of behavior problems at the start of preschool. For these latter children, there is some evidence of *increased* problem behavior — more withdrawn and sad behavior among children who came from FOL classrooms — but no other statistically significant differences among children who received FOL in preschool and those who did not.

Were Classroom-Level Improvements Sustained? A Preliminary and Descriptive Look at Foundations of Learning Teachers, One Year Later

Information from Observers

Based on independent ratings of teachers' classroom practices, were the FOL improvements in classroom practices sustained? It was unclear at the outset of the intervention what the effects might be in this area. On the one hand, it might be reasonable to expect that teachers would retain the skills that they had learned from FOL training, thus sustaining improved

The Foundations of Learning Demonstration

Table 4.2

Program Impacts on Teacher-Reported Child Outcomes, by Level of Behavior Problems, Kindergarten Year

	Low					High					
Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	H-Stars
Problem behavior											
Internalizing problems	1.8	1.8	0.0	0.5	0.00	4.7	3.5	1.2 *	0.7	0.34	
Externalizing problems	2.7	3.0	-0.3	0.8	-0.05	6.8	5.3	1.5	1.0	0.22	
Teacher-student conflict	10.4	10.5	-0.1	0.9	-0.02	13.7	13.6	0.2	1.2	0.02	
Attention problems	2.7	3.6	-0.9	0.7	-0.20	6.1	5.5	0.6	0.8	0.10	
Positive social behavior											
Compliance with teachers' directives	4.2	4.2	0.0	0.1	-0.03	3.7	3.8	-0.1	0.1	-0.13	
Social competence	4.2	4.1	0.1	0.1	0.10	3.8	3.7	0.0	0.1	0.04	
Teacher-student closeness	34.4	33.6	0.8	0.9	0.13	32.5	33.7	-1.2	1.0	-0.21	
Approach to learning											
Work-related skills	5.4	5.0	0.4 *	0.2	0.33	4.3	4.6	-0.3	0.2	-0.22	**
Academic skills											
Language and literacy skills	36.4	33.9	2.5 *	1.3	0.26	31.0	32.6	-1.7	1.5	-0.19	**
Math knowledge	28.1	26.9	1.2	1.0	0.16	25.1	26.1	-1.0	1.1	-0.15	
General knowledge	20.0	19.6	0.4	0.8	0.07	18.6	18.9	-0.4	0.8	-0.08	
Sample size	121	105				120	99				

(continued)

Table 4.2 (continued)

SOURCE: Based on MDRC calculations from responses to teacher survey.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent. The final column lists the H-stars, which show the statistical significance of the difference between the subgroup impacts.

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

"Internalizing problems" and "Externalizing problems" refer to the internalizing and externalizing subscales of the BPI. BPI internalizing and externalizing scales were created based on factor analysis work. "Teacher-student conflict" and "Teacher-student closeness" refer to the Student-Teacher Relationship Scale conflict and closeness subscales. "Attention problems" refers to the Caregiver-Teacher Report Form attention problems subscale. "Compliance with teachers' directives" and "Social competence" refer to the Positive Behavior Scale compliance and social competence subscales. "Work-related skills" refers to the Cooper-Farran Behavioral Rating Scales work-related skills subscale. "Language and literacy skills," "Math knowledge," and "General knowledge" refer to the language and literacy, math knowledge, and general knowledge subscales of the Academic Rating Scale.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

Outcome controls for the child's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

classroom practices. The hope was that teachers would show *no significant decreases* in the skills that they had mastered, even after the withdrawal of professional supports, such as coaching and repeated trainings. A more skeptical hypothesis might be that teachers' skills might not be sustained and instead might fall back to baseline levels. For example, the new school year might bring a range of challenges that might tax teachers' newly learned skills; these could include a new cohort of behaviorally challenging preschoolers, new staff to work with, and programmatic transitions such as changes in center leadership and administration.

All teachers were provided with a short refresher training (either through an abbreviated, one-day training and a set of materials to remind them of the highlights of the training or through printed materials only).[2] The thinking behind providing this booster was that it would be similar to what teachers would be receiving if the intervention were implemented on a larger scale.

Observations were conducted of FOL-assigned lead teachers who were still teaching preschool the following year. Unlike the analyses reported in Chapter 3, these analyses compare FOL teachers' classroom practices in spring of the follow-up year with their *own* performance one year earlier, near the completion of their participation in the intervention.[3] In short, these observational findings do not reflect impact analyses (or comparisons of the program and control groups), and they should not be interpreted as such.

The results of the independent observations are presented in Table 4.3 and are summarized in Figure 4.1. These results are relatively positive. For example, most of the effects in Table 4.3 are neutral: No differences arose between how teachers were observed at the two time points on the composite measure of positive classroom management, on the management of classroom time, on most measures of instruction, and on overall classroom student engagement. This is encouraging, given that the first assessment was conducted when the intervention was in full swing, as compared with one year later — with a new classroom of children and nearly a year after the Clinical Classroom Consultants (CCCs) had stopped working with teachers. In addition, teachers who were initially enrolled in the program group continued to show *further* improvements in their ability to maintain positive classroom climate, their sensitivity to children, and their use of engaging teaching methods, relative to their own performance in the spring of the prior academic year, which was not expected. On the less positive side, however, teachers were also observed to maintain a more emotionally negative climate, on average, than they had

[2]Notably, results are similar whether teachers received the one-day training or only the printed materials, although the power to differentiate the effects is very limited.

[3]This analysis includes 21 teachers for whom data are available from both time points. See Appendix A for further information.

The Foundations of Learning Demonstration

Table 4.3

Means for Observed Ratings of Classroom Climate and Instructional Time, Program Group in Spring 2008 Compared with Program Group in Spring 2009

Variable	Program Group 2008 Mean	Program Group 2009 Mean	Difference	
Positive classroom management				
Composite[a]	5.7	5.9	0.2	
Positive climate	5.5	5.9	0.4	*
Negative climate	1.2	1.6	0.5	**
Teacher sensitivity	5.2	5.7	0.5	**
Behavior management	5.3	5.6	0.3	
Use of classroom time				
Management of classroom time	5.3	5.4	0.1	
Quality of language instruction				
Composite	4.2	4.5	0.3	
Regard for student perspectives	5.1	5.4	0.3	
Use of engaging teaching methods	4.0	4.9	0.9	***
Promoting understanding through conversation	3.4	3.2	-0.2	
Encouragement of students' language use	4.2	4.5	0.3	
Overall classroom student engagement	5.6	5.8	0.2	
Sample size	21	21		

SOURCES: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in April-May 2008 and April-May 2009.

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

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

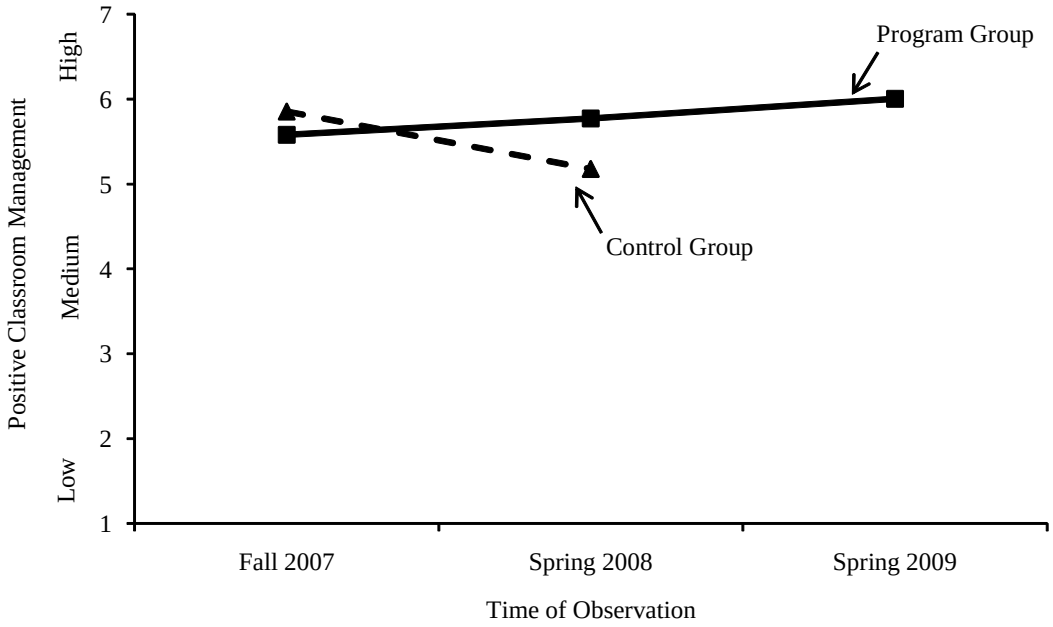
"Management of classroom time" refers to the "Productivity" dimension of the CLASS. "Use of engaging teaching methods" refers to the "Instructional learning formats" dimension of the CLASS. "Promoting understanding through conversation" refers to the "Quality of feedback" dimension of the CLASS. "Encouragement of students' language use" refers to the "Language modeling" dimension of the CLASS. For each dimension, observers rated classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

[a]"Negative climate" is reverse-coded for the composite score.

The Foundations of Learning Demonstration

Figure 4.1

Scores for Positive Classroom Management, Preschool and Kindergarten Years



SOURCES: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007, April-May 2008, and April-May 2009.

NOTES: Scores represent adjusted means that control for random assignment block. Spring 2008 and spring 2009 control for baseline fall CLASS dimensions as well.

Scores for the program group and the control group are statistically different (p-value = less than 5 percent for spring 2008).

the year earlier (although not more negative than in the control sites). These higher levels of negative climate may reflect some of the challenges that the teachers were facing in managing classrooms with fewer supports.

Figure 4.1 illustrates the effects on the summary measure of teachers' positive classroom management — the dimension on which teachers were trained in the context of FOL during the intervention year. What is clear is that teachers in the control group classrooms lost ground in their positive climate and behavior management, from fall to spring of the intervention year, during which time teachers in the FOL-assigned classrooms maintained their scores on this dimension. One year later, these teachers who had been previously trained were able to continue to sustain their practice, with scores about as high as they had been the prior spring.

Information from Teachers

Finally, qualitative interviews were conducted to obtain teachers' own perspectives on their ability to sustain practice in the year following the FOL intervention. These interviews provide information from the teachers directly about whether they felt they were using the strategies. If not, information was gathered about what contributed to challenges in using them. Among teachers who were using the strategies, the interviews focused on factors that were associated with their use in the second year. Findings from these interviews are discussed below.

First and foremost, these interviews show that, consistent with the quantitative data, teachers reported continued use of the strategies in which they had been trained the prior year, although a small number of teachers did report difficulties in continuing to use the strategies. Challenges in sustaining practice appear to have resulted from a lack of support among administrators and colleagues and from difficult relationships with co-teachers.

Second, these interviews highlight several factors that may be associated with teachers' improvements in their ability to sustain practice of the strategies learned. Three dimensions of experiences were salient for sustaining practice: (1) Teachers started off the school year differently; (2) there was more of a focus on collaboration across teaching teams; and (3) teachers changed the way that they approached children's behavior problems, throughout the year.

Starting the School Year Differently

Teachers spoke repeatedly about how they intentionally started their school year differently than in prior years. They discussed deliberately embedding classroom strategies that they had learned in FOL into their arsenal of skills from the first day of school. The first year of the intervention focused on learning the strategies, characterized by incremental steps of implementation and testing over the course of the year. In the year following, teachers who had successful experiences with the model actively changed the way that they wanted their classrooms to operate from the outset of the new school year. As one teacher stated:

> This year is when we're really implementing everything that we learned. And last year was more to try [out the strategies]. . . .
>
> There was a lot more structure [and] organization, and as a whole class I think we did a lot more of social and emotional talk. . . .We came up with rules together . . . *at the beginning to create the classroom environment.*

Team Teaching and Problem Solving

As discussed in detail in the 2009 implementation report on this project, both the lead and the assistant teachers reported that the FOL intervention was successful in facilitating dialogue about classroom practices and responsibilities that resulted in collaboration.[4] Teachers indicated that FOL provided a unique opportunity to promote this type of teamwork, since both lead and assistant teachers attended trainings together and jointly participated in regular debriefings with their CCCs. The result was that assistant teachers were more engaged in developing a shared classroom agenda, worked alongside lead teachers to develop a clearer sense of roles and responsibilities, and began to take part in decision-making about issues that affected the classroom, thereby increasing their feelings of ownership of the classroom. This stands in contrast to reports that suggested there was ongoing tension between lead and assistant teachers in Newark.

As described by teachers, the collaboration and behavior changes that FOL promoted were still evident one year later. One assistant teacher stated:

> Before they [lead teachers] [would] see us as "they," like we're not part of the team. . . . Some of the teachers still do it, but the one that I work with, I don't see that anymore.

Teachers also reported that their collaborative support resulted in clear respect for each other's views. An assistant teacher reported:

> Years before . . . we [lead teacher and I] could never compromise. Now this year we share ideas together, we're . . . closer, we're more interactive with our ideas and working as a team. . . . We're both teachers. . . . And I'm learning from that. I'm learning.

Teachers believed that these collaborative relationships translated into better classroom environments because even if children were not able to articulate what they were seeing, they could sense and feel the negativity when teachers disagreed with one another and were not on the same page about classroom management issues.

Changes in Teaching Practices

Finally, teachers reported shifts in the way that they approached classroom management philosophically, and they reported increased confidence in implementing the strategies that they had learned. By the beginning of the year following the intervention, teachers felt that they were more vigilant and more deliberate about classroom management practices, not only from the outset of the school year but also throughout the year. They discussed being actively engaged in

[4]Lloyd and Bangser (2009).

conversations with students and routinely using more positive language with children. For example, the statement below is from a teacher who reported paying more attention to children's social and emotional development:

> I changed the way I treat . . . children socially and emotionally. I can say that I understand them much better because I changed the way I was working with them [from] before with the new strategies.

Another teacher reported being much more reflective than in previous years, about the way she manages challenging behaviors:

> This project has forced me, when children have negative behaviors, it's forced me . . . to take a step back and say, "Well, why is this child acting like this?" where I never did that. What's going on and how can I change it? . . . So I ask myself those questions now, and I find that it's much easier for me to manage my children. Much easier.

Chapter 5

Conclusion and Discussion

Chapter 5 summarizes the key findings of this report on the Newark, New Jersey, site of the Foundations of Learning (FOL) demonstration. The findings are placed in the context of prior work on preschool interventions, and the chapter discusses possible next steps in research in this area.

Key Findings

As described in detail in this report, the evidence emerging from the Newark site of the FOL demonstration is somewhat promising, showing that investments in teachers' professional development can make a difference in children's experience in preschool.

First and foremost, FOL changed the primary outcome that was targeted — teachers' positive classroom management — improving the ways in which teachers managed children's behavior problems and provided an emotionally positive and supportive classroom climate. Even more important, FOL also affected a key aspect of classroom quality — teachers' use of classroom time — with FOL teachers scoring higher than their control counterparts on measures of management of classroom time, use of engaging teaching methods, and the amount of instructional time. These findings support the thinking behind this intervention: that managing children's behavior may be diverting teachers' attention from providing instruction in their preschool classrooms. That said, FOL did not affect other aspects of language instruction in preschool classrooms.

For children, FOL appears to have improved some but not all aspects of their behavior in the classroom. Observers found program group children to exhibit lower levels of problem behavior and greater engagement in classroom activities than their control group peers, but they did not observe more positive social behavior between FOL children and their classmates or teachers. Moreover, teachers did not report differences between FOL and control children in their behavior in the classroom or in their approach to learning. It is not clear whether this is due to the fact that the intervention primed teachers to see such behaviors as a result of the training.

Effects in the year following FOL show some good news and some bad news. On the one hand, benefits to children that were observed in preschool were not sustained as children moved to their kindergarten classrooms. For the most part, neutral effects were observed, with some small negative effects for the highest-risk children (those entering preschool with high levels of behavior problems) and some small positive effects for the lowest-risk children (those

entering preschool with low levels of behavior problems). While FOL appears to have benefited children during preschool, understanding how to continue its effects beyond the preschool period is critical to addressing the long-term needs of low-income children, especially those at highest risk. On the other hand, teachers *did* appear to sustain practice in the year following the intervention, showing that investments in their professional development in one year may affect successive cohorts of the children they teach in preschool.

Putting the Foundations of Learning Findings in Context

To understand the relevance of these early results from the FOL evaluation in Newark, it is worth comparing them with other studies of preschool interventions that seek to improve the school readiness of young children. A common metric, called the "effect size," is used to compare the magnitude of impacts across interventions. The effect size represents the difference between the program and control groups in standardized units, and its values range from 0 to 1.

FOL's effect sizes for classroom climate and management (approximately 0.60 to 0.75) are similar to or somewhat larger than those of other studies that have evaluated enhancements of an existing preschool curriculum. Examples of these studies include the Head Start REDI trial of the Preschool PATHS intervention, the Tools of the Mind evaluation, and earlier tests of the Incredible Years model that included both teacher and parent training components. Effect sizes range from 0.50 to 0.75 for these studies.[1] However, maybe the best measure of the FOL results are in comparison with CSRP — the project on which FOL was based — since it tested the same intervention model in a different city and was led by a different research team.[2] Results from that effort are slightly larger than, but in line with, the findings from FOL presented here.

In fact, the consistency across the findings from FOL and CSRP shows that the same model can produce relatively similar results when implemented in different cities, with different research teams, and under somewhat different organizational contexts (a university-based context in CSRP and a model that was more deeply embedded in the existing system of service delivery in FOL). Moreover, what makes this particularly noteworthy is that the effects in Newark occurred in a relatively resource-rich environment and involved a larger number of classrooms. In addition, the experience of FOL in Newark suggests that it is possible to implement an intervention like this on a modestly large scale in a variety of early education settings: Head Start centers, community-based child care centers, and public schools.

[1]Barnett et al. (2008b); Bierman et al. (2008); Diamond, Barnett, Thomas, and Munro (2007).

[2]Raver et al. (2008); Raver et al. (2009a); Raver et al. (2009b). The demonstration adopted the model used by CSRP (formerly the Chicago School Readiness Project) and adapted it slightly to fit a new policy context.

It is more difficult to compare FOL with other widely cited preschool evaluations, such as Perry Preschool and the Head Start Impact Study. These studies tested a particular form of preschool, rather than an enhancement, and they have few overlapping measures for comparison with FOL. Typically, such studies focus on effects on outcomes for children rather than on classroom-based processes as a means for assessing impacts as well. That said, FOL's impacts on children (about one-third of a standard deviation) are neither as small as those of Head Start per se (about 0.15 standard deviation) nor as large as those of Perry (on some measures, as large as 1 standard deviation).

Next Steps in Research

FOL is one of a set of new interventions emerging that show that it is possible to increase the quality of existing services being provided to children through targeted, strategic, professional development of teachers, aimed at improving children's social and emotional development as a primary outcome of interest. Current studies are testing a new generation of preschool curricula and teacher training strategies that are specifically designed to facilitate children's social-emotional competencies by (1) providing preschool classrooms with very specific hands-on activities and lessons for children to help them increase their knowledge about emotions and peer behavior; (2) providing training for teachers and parents in specific behavior strategies that support the social-emotional development of preschool children; and (3) providing children with opportunities to practice social roles while emphasizing critical skills of planning, memory, and attention. Results from FOL and from studies testing other approaches — both small-scale efficacy trials as well as the Department of Health and Human Services' Head Start Classroom-based Approaches and Resources for Emotion and Social skill promotion (CARES) national trial (managed by MDRC) — will provide critical information about how to strengthen preschools to best meet the needs of low-income children.

With this set of emerging research, the first steps in an ongoing research agenda in this area are well under way. Of concern, however, are the lack of sustained effects in kindergarten and, especially, the slightly negative effects for high-risk children. From a research perspective, understanding what kinds of kindergarten environments can sustain preschool intervention effects is critical. From an intervention perspective, interest in aligning preschool with the early grades is burgeoning, but few have figured out how to vertically align across this age period to most effectively sustain the benefits of quality improvements in preschool. This is an area that should be next on the research agenda.

In addition, supporting children's social-emotional development is only one part of strengthening preschool education. While behavior management forms the foundation on which a high-quality preschool experience develops, surprisingly little time in preschool classrooms is spent on literacy and numeracy instruction. In classrooms where teachers have

learned how to manage children's behavioral issues, the next step may be to focus on such critical preacademic skills. But early efforts at integrated models sometimes fail, as too much information is asked of teachers in a short period of time. The fact that FOL's early investments in teachers' professional development are sustained beyond the one year of intensive intervention efforts suggests that pairing this intervention with a cognitively focused curriculum in a second year might be a promising approach for enhancing the school readiness of preschool children. This may present an interesting opportunity to test whether an integrated literacy and social-emotional intervention could ensure that children get the academic skills that they need to prepare for elementary school.

Conclusion

Policymakers, educators, and the public increasingly regard early childhood programs as a promising approach to prepare children for formal schooling. The potential for a real payoff on this investment is large: High-quality model preschools have been found to return $4 to $10 in future benefits per dollar spent. As Nobel laureate economist James Heckman has argued, preschool may be the best time to intervene with children, as future gains build on prior skills. However, not all preschool models yield positive benefits, particularly when such models are delivered at scale. Targeting children's social and emotional development through effective classroom management may be one way to improve classroom quality. Findings presented in this report indicate the potential value of intervening with this focus in the preschool classroom.

Additional information on this intervention is yet to be released. Early information on the Chicago site of the FOL demonstration will follow this report. This will complement information emerging on other promising social-emotional interventions currently being tested in preschool classrooms, providing extensive information to policymakers and practitioners about where to put their attention in efforts to improve preschool quality.

Appendix A

Additional Information on the Foundations of Learning Research Design and Sampling

The Recruitment of Preschools

The recruitment effort for the full-scale phase of the Foundations of Learning (FOL) demonstration and evaluation began in the summer of 2007. Potential sites included all Abbott-funded preschools being overseen by the Newark Public School system in Newark, New Jersey. After initial phone calls to introduce the project, MDRC staff visited interested sites to meet with directors and teachers, when possible, to more fully explain the FOL model, the random assignment process, and the data collection efforts. (See Chapters 1 and 2.) Once the recruitment process reached the in-person visit stage, no sites declined to participate in the demonstration.

One classroom within each preschool, serving primarily 4-year-old children, was selected as the "study" classroom. In sites that had two or more preschool classrooms serving 4-year-olds, directors or principals nominated a teacher *prior to random assignment* for inclusion. Conducting this selection prior to random assignment was critical to ensure the equivalence of the two groups at the outset. If teachers were allowed to select into the study after knowing whether or not they were going to receive the intervention, different teachers could have selected to participate in the program and control groups, biasing the results of the study.

Consent for Research Participation

Parental consent allowed teachers to complete reports on the children and gave permission to access the children's school records. Among registered children in all classrooms, 77 percent of parents agreed to participate in the project. The treatment classrooms had a slightly higher average consent rate: 81 percent, compared with 73 percent for the control classrooms. The rates differed largely because it was difficult to gather consents in two of the 25 control group classrooms. The teachers in these two classrooms declined to actively participate in data collection. It is not clear whether or not this was due to their random assignment status, but the refusals did occur after that assignment was made. This resulted in few, if any, parental consent forms being gathered in these two classrooms. Attempts by the research team to augment the parental consent outreach with in-person visits proved fruitless. Notably, however, classroom observations and observations of individual children did not require parental consent, based on the site's agreement to participate in the demonstration, and these data were collected from all classrooms.

Although parents and teachers of consented children completed surveys at extremely high rates, as discussed in Chapter 2, there was concern about how the differences in the samples for the two sources of data — teacher reports and observations of children — could affect the results based on these sources. Sensitivity analyses were conducted to address this issue. Appendix Tables C.1 and C.2 show the impacts on the group of children in common across the two data sources (teacher surveys on children and child observations). The pattern of

these impacts mirrors the pattern of impacts shown in the tables of Chapter 3, indicating that the differing results are not due to sample differences but to differences in the source of the information.

As participants in the research demonstration, lead teachers were also asked to complete a consent form allowing collection of school records and the use of their self-report survey responses. As mentioned above, two control group teachers declined to participate in this research, and so 49 of the 51 lead teachers did agree to participate.

Data Collection and Survey Response Rates

Parent Survey

As part of the baseline data collection, parents completed a survey by telephone. Questions included a set of demographic items on the family (such as marital status, race/ethnicity, household composition) and the child (gender, age). Findings based on these data are shown in Table 2.3. In addition, parents were asked about their economic status (income, public assistance status, employment experience) and their child's exposure to a range of parental psychosocial risks. Overall, the survey was completed by 92 percent of parents who had agreed to their child's participation in the demonstration.

Teacher Surveys

A self-survey was fielded among teachers and completed prior to the first teacher training. All teachers who agreed to participate in the research effort (96 percent of all teachers) completed the self-survey at baseline and again at the spring follow-up. The survey asked teachers to report on the prevalence of challenging behaviors in their classrooms, their confidence in managing children's behavior in the classroom, how demanding they felt their job was, their control over key aspects of their job, and, finally, their own psychological distress. See Appendix D for discussion of these outcomes and the results.

Teacher Surveys on Children

At baseline and preschool follow-up, teachers completed surveys on each consented child. Completion rates for these reports were high at both time points. At baseline, reports were completed on 93 percent of consented children, overall. For the spring follow-up (2008), report completion rates remained high, at 92 percent overall, with no statistically significant difference between the rates for treatment classrooms (92.5 percent) and control classrooms (91.5 percent).

If a child attended another preschool, he or she was tracked to the new environment, and new teachers were asked to complete the spring preschool follow-up report.[1]

The same survey was administered in spring 2009, to assess the longer-term effects of FOL. Less than 10 percent of consented children were not located for this follow-up survey; nearly half of those children were reported to be living out of the country or not to be attending a formal school. Children were enrolled in more than 120 schools and 240 classrooms, primarily in kindergarten classrooms.[2] Surveys were completed for 84 percent of consented children in the demonstration. For this data collection effort in the follow-up year, the program and control groups' tracking and survey completion rates are nearly identical and, thus, are not statistically significantly different.

Classroom Observations

As discussed in Chapter 2, two types of classroom observations were collected during the intervention year. As shown in Table 2.1, CLASS observations, which assess overall classroom quality, were done at three time points during the intervention year: fall (before the first teacher training), winter, and spring. The related inCLASS measure, which assesses individual children's interactions and approaches to learning, was collected for five preselected children in each classroom at the spring follow-up.[3] As mentioned above, these observations were conducted in all classrooms, regardless of teacher participation in the data collection. This was permitted because the data collection did not interrupt the normal operations of classrooms or children's experience in such classrooms, and so it was part of the center's agreement to be in the study.

Appendix Table A.1 presents observer ratings of overall classroom quality at the time of the baseline assessment, *prior* to teacher training components of the FOL model. CLASS scores on the positive classroom management dimension are quite high — an average of 5.6 to

[1]Approximately 9 percent of consented children were no longer enrolled in their FOL classrooms by the spring follow-up period. Many had moved to a new preschool center (or to a new classroom within the same site), but some were also found in home-based care or living outside the United States. If children had moved from the FOL site in the past 30 days, the FOL teacher completed the report.

[2]At the time of the one-year follow-up survey, 24 children (less than 5 percent of the sample) were still in preschool, and 5 children (1 percent) were reportedly in first grade.

[3]For preselection, children were stratified by gender and by teacher-reported baseline behavior scores, with a goal of representing boys and girls at low, moderate, and high levels of behavior problems. In cases where baseline behavior scores were not available, teachers were asked in the fall to identify children "who could benefit from more intensive service," and those children were selected as the ones with high levels of behavior problems. Substitute child selections were also included for inCLASS coders so that five children could be observed even if a preselected child was not in attendance on the day of the observations.

Appendix Table A.1

Observed Ratings of Classroom Climate, Instructional Time, and Language Instruction at Baseline

Variable	Program Group Mean	Control Group Mean	Difference	Standard Error
Positive classroom management				
Composite[a]	5.6	5.9	-0.3	0.2
Positive climate	5.6	5.9	-0.3	0.2
Negative climate	1.3	1.2	0.1	0.1
Teacher sensitivity	4.8	5.1	-0.3	0.2
Behavior management	5.2	5.5	-0.3	0.3
Use of classroom time				
Management of classroom time	5.4	5.8	-0.4 **	0.2
Quality of language instruction				
Composite	4.1	4.0	0.0	0.2
Regard for student perspectives	4.6	4.9	-0.2	0.2
Use of engaging teaching methods	4.2	3.9	0.3 *	0.2
Promoting understanding through conversation	3.5	3.3	0.2	0.2
Encouragement of students' language use	4.0	4.1	-0.2	0.2
Sample size	26	25		

SOURCE: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.
Adjusted means control for random assignment block.
For each dimension, observers rated classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."
"Management of classroom time" refers to the "Productivity" dimension of the CLASS. "Use of engaging teaching methods" refers to the "Instructional learning formats" dimension of the CLASS. "Promoting understanding through conversation" refers to the "Quality of feedback" dimension of the CLASS. "Encouragement of students' language use" refers to the "Language modeling" dimension of the CLASS.
[a]"Negative climate" is reverse-coded for the composite score.

5.9 out of 7 on positive climate and an average of 1.2 in a range of 1 to 7 on negative climate.[4] Given the small class size and high levels of teacher credentialing as a result of the New Jersey Supreme Court decisions in the *Abbott v. Burke* class action case (see Chapter 1), these scores are not surprising. Scores on the quality of language instruction dimension are somewhat lower, as they are in most preschool classrooms. Classrooms show relatively high scores on manage-

[4]The CLASS manual's technical appendix includes average CLASS scores for prior preschool studies, including MS/SWEEP (5.3) and My Teaching Partner (5.2). Both studies found negative climate scores of 1.6.

ment of classroom time. Notably, it is against this relatively high base of quality that FOL was tested.

Obtaining baseline assessments of classroom quality also acted as a check that the randomization procedure resulted in two groups of classrooms that were roughly equivalent in quality. Appendix Table A.1 suggests that program-assigned and control group-assigned classrooms differed on only two out of nine dimensions at baseline. Classrooms in the program group had higher scores on one of the instructional dimensions (use of engaging teaching methods) but had lower scores on management of classroom time. Thus, random assignment seems to have "worked," in that neither the program group nor the control group classrooms appear to have been systematically better than the other. However, as discussed in Chapter 2, to increase the power of the impact estimates and to control for these baseline differences, all baseline ratings are included as controls in models testing the effects of FOL on spring CLASS observation scores.

As presented in Chapter 4, as part of a sustainability study, the program group teachers were located and were observed again on the CLASS measure in spring of the year following the intervention. Of the lead teachers, 23 out of 26 were still teaching in Newark schools; however, illnesses and logistical issues resulted in complete observations for only 21 of these teachers. The analysis in Chapter 4 compares their CLASS scores at the two spring observation points.

Appendix B

Supplemental Tables for Chapter 2

The Foundations of Learning Demonstration

Appendix Table B.1

Items and Factor Loadings for Classroom Assessment Scoring System Composites

Item	Composite: Positive Classroom Management	Composite: Quality of Language Instruction
Positive climate	0.9	
Negative climate	-0.6	
Teacher sensitivity	0.8	
Behavior management	0.9	
Regard for student perspectives		0.6
Use of engaging teaching methods		0.9
Promoting understanding through conversation		0.5
Encouragement of students' language use		0.8
Cronbach coefficient alpha for composite	0.8	0.7

SOURCE: MDRC calculations using Classroom Assessment Scoring System (CLASS) observations in September-October 2007.

NOTES: A promax rotation, two-factor structure was used to identify composites. Only factor items with loadings greater than or equal to |.30| are shown, based on factor analysis. Factor loadings indicate items that were used to create the respective scales. Items were included on the factors on which they most highly loaded.

"Use of engaging teaching methods" refers to the "Instructional learning formats" dimension of the CLASS. "Promoting understanding through conversation" refers to the "Quality of feedback" dimension of the CLASS. "Encouragement of students' language use" refers to the "Language modeling" dimension of the CLASS.

The Foundations of Learning Demonstration

Appendix Table B.2

Items and Factor Loadings for the Behavior Problems Index Subscales

Item in Total Scale	Internalizing Behavior	Externalizing Behavior
Is rather high strung and nervous	0.4	
Is secretive, keeps things to himself/herself	0.6	
Worries too much	0.8	
Is too fearful or anxious	0.6	
Is easily confused, seems to be in a fog	0.4	
Feels worthless or inferior	0.6	
Has difficulty getting his/her mind off certain thoughts	0.5	
Is unhappy, sad, or depressed	0.7	
Is withdrawn, does not get involved with others	0.8	
Clings to adults	0.6	
Cries too much	0.6	
Demands a lot of attention	0.5	
Is too dependent on others	0.6	
Has sudden changes in mood or feeling[a]	0.3	0.5
Hangs around with kids who get in trouble		0.6
Cheats or tells lies		0.7
Argues too much		0.8
Has difficulty concentrating, cannot pay attention for long		0.5
Bullies or is cruel or mean to others		0.8
Is disobedient		0.9
Does not seem to feel sorry after he/she misbehaves		0.8
Has trouble getting along with other children		0.7
Is impulsive, or acts without thinking		0.8
Is not liked by other children		0.5
Is restless or overly active, cannot sit still		0.7
Is stubborn, sullen, or irritable		0.7
Has a very strong temper and loses it easily		0.8
Breaks things on purpose or deliberately destroys his/her own or another's things		0.7
Feels others are out to get him/her		
Feels or complains that no one loves him/her		
Cronbach coefficient alpha for scale	0.9	0.9

SOURCE: Based on MDRC calculations from responses to teacher survey.

NOTES: A promax rotation, two-factor structure was used to identify subscales. Only factor items with loadings greater than or equal to |.30| are shown, based on factor analysis. Factor loadings indicate items that were used to create the respective scales. Items were included on the factors on which they most highly loaded.

[a]This item was grouped with the Internalizing Behavior subscale based on theory.

Appendix C

Supplemental Tables for Chapter 3

The Foundations of Learning Demonstration

Appendix Table C.1

Program Impacts on Observed Ratings of Child Behavior, Consented Children Only, Preschool Year

Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Problem behavior					
Teacher conflict	1.3	1.5	-0.2 **	0.1	-0.34
Peer conflict	1.4	1.6	-0.2 **	0.1	-0.31
Positive social behavior					
Teacher communication	2.2	2.4	-0.1	0.1	-0.15
Teacher positive engagement	3.2	3.5	-0.2	0.2	-0.30
Peer communication	2.5	2.7	-0.2	0.2	-0.21
Peer sociability	3.5	3.7	-0.2	0.1	-0.29
Peer assertiveness	2.1	2.4	-0.2	0.2	-0.26
Approach to learning					
Task engagement	4.9	4.7	0.2	0.1	0.23
Task self-reliance	3.1	3.2	-0.1	0.2	-0.10
Task behavior control	5.4	5.0	0.3 *	0.2	0.34
Sample size - students	117	95			
Sample size - classrooms	26	23			

SOURCE: Based on MDRC calculations of Individualized Classroom Assessment Scoring System (inCLASS) observations in April-May 2008.

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

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

The outcomes "Problem behavior," "Positive social behavior," and "Approach to learning" come from the inCLASS observations. "Overall classroom student engagement" comes from the CLASS. For each dimension, observers rated children and classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

The Foundations of Learning Demonstration

Appendix Table C.2

Program Impacts on Teacher-Reported Child Outcomes, Consented Children Only, Preschool Year

Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Problem behavior					
Internalizing problems	3.1	2.5	0.6	0.6	0.22
Externalizing problems	4.6	4.1	0.5	0.7	0.11
Teacher-student conflict	12.8	13.0	-0.2	1.2	-0.04
Attention problems	4.1	3.5	0.6	0.8	0.15
Positive social behavior					
Compliance with teachers' directives	4.0	3.9	0.1	0.1	0.16
Social competence	4.0	4.0	0.1	0.1	0.09
Teacher-student closeness	34.6	35.9	-1.3	1.1	-0.24
Approach to learning					
Work-related skills	4.8	4.9	-0.1	0.2	-0.10
Preacademic skills					
Language and literacy skills	35.7	33.9	1.8	1.7	0.22
Math knowledge	25.8	26.4	-0.6	1.7	-0.08
General knowledge	20.4	18.9	1.4	1.0	0.34
Sample size - students	117	95			
Sample size - classrooms	26	23			

SOURCE: Based on MDRC calculations from responses to teacher survey.

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

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

"Internalizing problems" and "Externalizing problems" refer to the internalizing and externalizing subscales of the BPI. BPI internalizing and externalizing scales were created based on factor analysis work. "Teacher-student conflict" and "Teacher-student closeness" refer to the Student-Teacher Relationship Scale conflict and closeness subscales. "Attention problems" refers to the Caregiver-Teacher Report Form attention problems subscale. "Compliance with teachers' directives" and "Social competence" refer to the Positive Behavior Scale compliance and social competence subscales. "Work-related skills" refers to the Cooper-Farran Behavioral Rating Scales work-related skills subscale. "Language and literacy skills," "Math knowledge," and "General knowledge" refer to the language and literacy, math knowledge, and general knowledge subscales of the Academic Rating Scale.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

Outcome controls for the child's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

The Foundations of Learning Demonstration

Appendix Table C.3

Program Impacts on Observed and Teacher Ratings of Child Outcomes, by Gender, Preschool Year

	Boys					Girls					
Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	H-Stars
Observations											
Problem behavior											
Teacher conflict	1.4	1.7	-0.3 **	0.1	-0.42	1.1	1.3	-0.1 *	0.1	-0.37	
Peer conflict	1.5	1.6	-0.1	0.1	-0.18	1.3	1.6	-0.3 *	0.1	-0.36	
Positive social behavior											
Teacher communication	2.2	2.3	-0.1	0.2	-0.15	2.2	2.4	-0.2	0.1	-0.23	
Teacher positive engagement	3.1	3.3	-0.2	0.2	-0.21	3.4	3.6	-0.2	0.2	-0.26	
Peer communication	2.4	2.5	-0.1	0.2	-0.11	2.5	2.6	-0.1	0.2	-0.13	
Peer sociability	3.4	3.6	-0.1	0.2	-0.18	3.6	3.6	0.0	0.2	0.03	
Peer assertiveness	2.0	2.2	-0.2	0.2	-0.21	2.2	2.3	-0.1	0.2	-0.16	
Approach to learning											
Task engagement	4.7	4.4	0.3	0.2	0.41	5.0	4.8	0.2	0.2	0.25	
Task self-reliance	3.0	3.0	0.0	0.2	0.00	3.2	3.3	-0.1	0.3	-0.10	
Task behavior control	5.1	4.8	0.3	0.2	0.34	5.7	5.3	0.4 *	0.2	0.44	
Teacher reports											
Problem behavior											
Internalizing problems	3.1	2.4	0.7	0.8	0.20	2.0	2.5	-0.5	0.5	-0.17	
Externalizing problems	5.3	4.4	0.9	0.9	0.16	2.5	3.0	-0.5	0.7	-0.11	
Teacher-student conflict	13.4	12.6	0.8	1.1	0.13	11.1	11.3	-0.1	0.7	-0.03	
Positive social behavior											
Social competence	3.9	3.9	0.0	0.1	-0.04	4.2	4.1	0.1	0.1	0.21	
Teacher-student closeness	33.9	35.2	-1.4	0.9	-0.24	35.4	36.1	-0.7	1.0	-0.15	

(continued)

Appendix Table C.3 (continued)

Outcome	Boys					Girls					H-Stars
	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	
Teacher reports (continued)											
Approach to learning											
Work-related skills	4.6	4.7	0.0	0.1	-0.04	5.1	5.0	0.1	0.1	0.14	
Preacademic skills											
Language and literacy skills	34.6	32.9	1.6	1.7	0.17	35.6	32.8	2.8	1.7	0.33	
Math knowledge	25.4	25.9	-0.4	1.5	-0.06	26.1	25.4	0.8	1.7	0.11	
Sample size - observations of students	67	59				61	62				
Sample size - teacher reports on students	153	138				130	110				
Sample size - classrooms	26	23				26	23				

SOURCES: Based on MDRC calculations of Individualized Classroom Assessment Scoring System (inCLASS) observations in April-May 2008 and on responses to teacher survey.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent. The final column lists the H-stars, which show the statistical significance of the difference between the subgroup impacts.

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

Observed outcomes "Problem behavior," "Positive social behavior," and "Approach to learning" come from the inCLASS observations. For each inCLASS dimension, observers rated children and classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

"Internalizing problems" and "Externalizing problems" refer to the internalizing and externalizing subscales of the BPI. BPI internalizing and externalizing scales were created based on factor analysis work. "Teacher-student conflict" and "Teacher-student closeness" refer to the conflict and closeness subscales of the Student-Teacher Relationship Scale. "Social competence" refers to the social competence subscale of the Positive Behavior Scale. "Work-related skills" refers to the work-related skills subscale of the Cooper-Farran Behavioral Rating Scales. "Language and literacy skills" and "Math knowledge" refer to the language and literacy and math knowledge subscales of the Academic Rating Scale.

Teacher-reported outcomes control for the student's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

The Foundations of Learning Demonstration

Appendix Table C.4

Program Impacts on Observed and Teacher Ratings of Child Outcomes, by Race/Ethnicity, Preschool Year

	Black, Non-Hispanic					Hispanic					
Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	H-Stars
Observations											
Problem behavior											
Teacher conflict	1.3	1.5	-0.2	0.1	-0.43	1.2	1.5	-0.3	0.2	-0.37	
Peer conflict	1.4	1.8	-0.4 **	0.2	-0.46	1.4	1.7	-0.3 *	0.2	-0.48	
Positive social behavior											
Teacher communication	2.2	2.2	0.1	0.2	0.07	2.3	2.5	-0.2	0.2	-0.27	
Teacher positive engagement	3.4	3.3	0.1	0.3	0.15	2.9	3.6	-0.7 **	0.2	-0.86	**
Peer communication	2.8	2.7	0.1	0.2	0.11	2.5	2.9	-0.3	0.3	-0.37	
Peer sociability	3.7	3.7	0.0	0.2	0.01	3.2	3.7	-0.6 **	0.2	-0.64	*
Peer assertiveness	2.4	2.3	0.1	0.2	0.07	1.9	2.6	-0.7 *	0.3	-0.69	*
Approach to learning											
Task engagement	4.9	4.7	0.2	0.2	0.21	5.0	4.8	0.2	0.2	0.21	
Task self-reliance	3.1	3.1	0.0	0.3	0.00	3.1	3.4	-0.3	0.3	-0.29	
Task behavior control	5.4	5.0	0.4	0.3	0.41	5.3	5.1	0.2	0.3	0.22	
Teacher reports											
Problem behavior											
Internalizing problems	3.5	2.4	1.2	0.8	0.38	2.4	2.8	-0.3	1.2	-0.08	
Externalizing problems	4.6	4.1	0.5	1.2	0.09	4.2	4.2	0.0	0.9	0.00	
Teacher-student conflict	12.1	11.6	0.5	1.5	0.08	12.4	12.9	-0.5	1.2	-0.09	
Positive social behavior											
Social competence	4.0	3.9	0.1	0.1	0.15	4.0	4.0	-0.1	0.2	-0.10	
Teacher-student closeness	34.8	36.2	-1.4	1.3	-0.23	34.5	35.6	-1.1	1.1	-0.30	

(continued)

Appendix Table C.4 (continued)

Outcome	Black, Non-Hispanic					Hispanic					H-Stars
	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size	
Teacher reports (continued)											
Approach to learning											
Work-related skills	4.7	4.7	-0.1	0.1	-0.07	4.9	4.8	0.2	0.2	0.14	
Preacademic skills											
Language and literacy skills	33.1	30.6	2.4	2.7	0.25	36.3	34.0	2.3	2.1	0.29	
Math knowledge	24.6	23.3	1.3	2.3	0.17	27.6	26.0	1.6	2.4	0.24	
Sample size - observations of students	54	51				40	31				
Sample size - teacher reports on students	124	126				89	77				
Sample size - classrooms	26	23				26	23				

SOURCES: Based on MDRC calculations of Individualized Classroom Assessment Scoring System (inCLASS) observations in April-May 2008 and on responses to teacher survey.

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent. The final column lists the H-stars, which show the statistical significance of the difference between the subgroup impacts.

Regression-adjusted means control for random assignment status and blocking, baseline Classroom Assessment Scoring System (CLASS) measures, and baseline child characteristics.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

Observed outcomes "Problem behavior," "Positive social behavior," and "Approach to learning" come from the inCLASS observations. For each inCLASS dimension, observers rated children and classrooms on a scale from 1 to 7, with 1 representing "low" and 7 representing "high."

"Internalizing problems" and "Externalizing problems" refer to the internalizing and externalizing subscales of the BPI. BPI internalizing and externalizing scales were created based on factor analysis work. "Teacher-student conflict" and "Teacher-student closeness" refer to the conflict and closeness subscales of the Student-Teacher Relationship Scale. "Social competence" refers to the social competence subscale of the Positive Behavior Scale. "Work-related skills" refers to the work-related skills subscale of the Cooper-Farran Behavioral Rating Scales. "Language and literacy skills" and "Math knowledge" refer to the language and literacy and math knowledge subscales of the Academic Rating Scale.

Teacher-reported outcomes control for the student's baseline score on a given measure, when available. These include baseline measures for the Cooper-Farran Behavioral Rating Scales, the Behavior Problems Index (BPI), and the Positive Behavior Scale.

Appendix D

Impacts on Teachers' Reports of Skills, Mental Health, and Job-Related Stressors

Along with outcomes for children and classrooms, the Foundations of Learning (FOL) evaluation also examined the effects of the intervention on teachers' perceptions of their own skills, job stress, and depression. This is an important question to address, particularly in light of concerns that teachers who feel overwhelmed by their students' behavioral difficulties may be at greater risk for being less motivated, more detached, and more burnt out when supervising their classrooms.[1] These findings are not part of the primary targets of the intervention. However, in the full analysis of the effects of FOL, it was important to understand whether the intervention, in changing teachers' ability to manage their classroom behavior, would be successful in helping teachers feel greater control in their jobs and experience less stress and burnout.

Therefore, four measures were used to assess the effects of the intervention on outcomes related to teachers. A teacher's confidence in handling the classroom and any misbehavior was evaluated using a *classroom management skills* measure (adapted from CSRP).[2] To evaluate how often stressful situations occur in the classroom and a teacher's perceived amount of control over the job, the *job stress* and *job autonomy* scales were adapted from the Childcare Worker Stress Inventory.[3] Finally, the teacher's *psychological distress* was calculated using the Kessler-6 scale (K6).[4]

Focusing on these measures, there is no evidence of impacts on these outcomes for teachers. As shown in Appendix Table D.1, no differences were found between the FOL-assigned and the control-assigned teachers in the extent to which they felt that their jobs were demanding or the extent of control that they felt they had in their jobs. Importantly, these scales

[1]Li-Grining et al. (2010).

[2]Raver et al. (2008); the demonstration slightly adapted the CSRP model to fit a new policy context. Scores are calculated by summing the responses to eight questions. Scores in this sample range from 18 to 38, though more than half of the teachers fell within the 27 to 31 range, with a larger score representing greater confidence.

[3]Curbow et al. (2000). The six-item job stress scale assesses how often ("never" to "all of the time") certain stressful situations occur in a teacher's classroom. Examples of these situations include working long hours and parents' blaming the daycare setting for their child's bad behavior. The five-item job autonomy scale assesses the perceived amount of control that a teacher has over such job situations as taking time off work when needed and the type of daily activities required. For each scale, the outcomes reported represent a summary score of all items.

[4]Kessler et al. (2003). The score is the sum of responses to six questions, on a 5-point scale, asking how often a teacher experienced symptoms of psychological distress in the past 30 days. The measure's scores can range from 0 to 24, with a score of 13 points or more as an indication of serious psychological distress. In this sample, scores ranged from 0 to 9.

The Foundations of Learning Demonstration

Appendix Table D.1

Program Impacts on Teachers' Reports of Skills, Mental Health, and Job-Related Stressors

Outcome	Program Group Mean	Control Group Mean	Difference (Impact)	Standard Error	Effect Size
Classroom management skills	29.8	29.4	0.4	1.1	0.11
Job stress	16.3	15.6	0.7	1.2	0.27
Job autonomy	18.3	17.8	0.5	0.8	0.25
Psychological distress	2.6	2.6	0.0	0.9	0.01
Sample size	25	23			

SOURCE: Based on MDRC calculations from responses to teacher surveys.

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

The table presents adjusted means that control for random assignment blocks, select baseline teacher characteristics (years teaching preschool, race/ethnicity, and degree in education), and, where available, baseline score on the given outcome.

The effect size equals the impact divided by the standard deviation of the outcome measure for the control group.

address a number of different aspects of jobs, not simply teachers' feelings of control over the classroom and the challenging behavior that children are exhibiting. Similarly, no differences were found between the two groups of teachers on the level of reported depression symptoms. It may simply be that while teachers were given concrete tools in managing one dimension of their jobs (for example, how they managed their classrooms), this did not translate to changes in other aspects of their roles and responsibilities in preschool settings.

Appendix E

Cost Estimates for the Foundations of Learning Model

As discussed in Chapter 1 of this report and, in greater detail, in the 2009 implementation report,[1] the multicomponent Foundations of Learning (FOL) intervention model was well implemented in the Newark, New Jersey, site of the demonstration. This appendix presents the costs of operating this full-scale model, as implemented in Newark by MDRC. Cost information was collected by MDRC as part of the operation and management of the intervention.

Explanation of Costs

The primary costs associated with delivering FOL are the teacher training — including the costs of the trainer, teacher compensation, and space and meals to conduct the training — and the costs of the Clinical Classroom Consultants (CCCs). As discussed further below, by far the largest costs associated with delivering FOL are the salaries of the CCCs. Costs are calculated over the one year of delivering FOL in the 26 FOL-assigned classrooms.

Appendix Table E.1 presents the costs of delivering each component of the FOL intervention, showing the per unit cost, the number of units, and the total cost.

Teacher Training

The top panel of Appendix Table E.1 lists the costs for the teacher training sessions. Two kinds of costs are associated with the trainings: *Fixed costs* did not depend on the number of teachers who took advantage of the training or who asked for child care services, while *variable costs* were affected by attendance and take-up rates. Since there was relatively high take-up of the FOL intervention, the variable costs represent nearly the highest possible costs for these activities.

All 52 teachers (including lead and assistant teachers in the 26 program group classrooms) were invited to attend the monthly Saturday sessions. Each teacher received an *Incredible Years* book for a total cost of approximately $2,000. Training sessions were conducted by a trainer who was in the process of completing certification in the Incredible Years curriculum[2] and who was assisted by a consultant who had prior experience with both the curriculum and the classroom consultation. Trainer fees and travel costs were the largest expense of the trainings: an average of $7,300 per session, for a total of $36,500. A downtown Newark location was selected, based on its centrality for teachers, with weekend rental rates of $950 per session. Parking was available on-site and was included in the rental fee. In addition, two meals were provided at each session by an outside caterer, for a total cost of just under $15,000.

[1]Lloyd and Bangser (2009).
[2]The trainer used video presentations of the FOL training sessions toward certification in the curriculum.

Additional costs varied by session, based on teacher attendance and child care needs. Teachers were well compensated for attending sessions. They were compensated for participation in the Saturday training at a rate that was commensurate with their salaries (which, as discussed in Chapter 1, were somewhat higher than preschool teachers' salaries in other urban districts). Lead teachers received $300 per session, and assistant teachers received $165 per session. The table lists the costs for payments based on the number of teachers who participated across the five sessions. In order to reduce barriers to participation in the voluntary training sessions, stipends for child care were provided on request. The combination of materials, payments, and fees for the teacher training component of the model cost approximately $126,500.

Clinical Classroom Consultation

The second panel of Appendix Table E.1 lists the costs associated with the classroom consultation component of the FOL model. The CCCs were hired to spend a full day each week in every program classroom, to model and coach the teacher in the training activities and, in the spring, to provide direct services to children. They were hired and supervised by Family Connections, a New Jersey-based family services agency. A total of seven consultants spent one day a week in each of the 26 classrooms (three or four days each in classrooms) and one day at the Family Connections office, debriefing with supervisors, completing intervention paperwork, and furthering their knowledge of the model.[3] The Family Connections budget included the 10-month salaries of the seven CCCs (a total of $300,000), clinical supervision, office space, and managerial coordination. In addition, for two of the CCCs, Family Connections paid six months of unemployment insurance payments, at an average monthly cost of $1,400. Over the course of the demonstration, MDRC paid just over $478,000 to Family Connections. This is by far the largest expense of delivering the model.

Prior to entering the classrooms in September, the CCCs participated in a three-day orientation that introduced them to the project. This included two days at the offices of MDRC and Family Connections, learning about the research design, the Incredible Years curriculum, the manual-based approach to the in-class consultation model, how to complete paperwork, and the policies and procedures of Family Connections and the research demonstration. An additional day was spent in the field, touring and learning about Newark so that the CCCs would be well versed about the community context and the sites in which they would be working. The orientation costs (training materials, catering, consultant fees) totaled approximately $1,700.

[3]Two CCCs who were in classrooms only three days each week also provided one day weekly of support on the research side of the project. Only their time spent on programmatic tasks is included here.

The Foundations of Learning Demonstration

Appendix Table E.1

Estimated Costs of the Foundations of Learning Demonstration Intervention

Component	Average Cost per Unit ($)	Number of Units	Total Cost ($)
Expenditures for Incredible Years teacher trainings			
Fixed costs			
Trainer fees and travel costs	7,309	5	36,547
Location rental	950	5	4,750
Incredible Years book	38	52	1,973
CCC attendance	2,310	5	11,550
Catering	2,956	5	14,779
Variable costs			
Child care	581	5	2,905
Lead teacher payments	300	114	34,200
Assistant teacher payments	165	101	16,583
Other expenses[a]	647	5	3,236
Total for 5 trainings			126,523
Expenditures for classroom consultation			
Family Connections			
CCC salary	41,945	7	293,614
Clinical supervision	7,691	11	84,601
Unemployment	1,434	6	8,604
Other expenses[b]	7,640	12	91,682
CCC orientation			
Incredible Years books	38	8	304
Training materials	62	1	62
Catering	195	5	975
Consultant fee	350	1	350
Stress management workshop[c]	166	26	4,325
Total for classroom consultation			484,517
Expenditures for MDRC management			58,501
Total expenditures for all components[d]			669,541
Total expenditures per classroom[e]			25,752

SOURCES: MDRC calculations based on fiscal and budget data.

NOTES: CCC = Clinical Classroom Consultant.
[a]"Other expenses" covers additional supplies purchased for training sessions.
[b]This includes variable expenses, such as rent, supplies, and travel.
[c]This includes catering, payment for teachers, an incentive for teachers to attend, and additional supplies.
Workshops were done on-site for each classroom with 26 sessions occurring.
[d]This includes all teacher training expenditures, classroom consultant expenditures, and MDRC management.
[e]This is based on 26 classrooms.

The stress management and one-on-one service components of the intervention model are also included as part of the consultation budget. The direct services to children occurred during the CCCs' scheduled time in class, so there were no additional costs for their time with regard to this component. Stress management was an ongoing area of focus for their classroom work with the teachers. In addition, the one-time stress management workshops were conducted on-site by the CCC Coordinator and were attended by the CCC, when possible. Beyond the salaried time of the CCCs and coordinator, workshop costs totaled $4,300, which includes lunch during each of the 26 sessions and a small gift to reinforce the stress reduction lessons that were learned in the workshop.[4]

Finally, staff at MDRC spent time on managerial aspects of the intervention. As discussed in the report, the FOL demonstration in Newark aimed to move the multicomponent interventions away from the university-led intervention of the Chicago School Readiness Project. Housing the CCCs in a local agency was one way to accomplish this, but MDRC's involvement and considerable programmatic management overall were evident throughout the year and were part of intervention costs. For example, the project manager, an MDRC research associate, participated in debriefing calls with the CCCs and their coordinator, served as a liaison to sites, and provided supervision regarding intervention implementation and fidelity. An administrative assistant also organized and managed logistics for the trainings and CCC orientation. This programmatic management by MDRC cost approximately $58,500.

Considerations for Replication

Notably, the costs reported here represent the costs of adding an enhancement like FOL fully *on top* of existing training and professional development in preschool settings. However, if integrated into the existing system, there are a few ways in which the incremental costs of implementing FOL may be lower. For example, trainings could be offered on existing professional development days and could be delivered on-site. And management costs might be reduced. But the largest costs of FOL are the salaries of the credentialed and skilled CCCs, for whom there are less clear cost-saving options if FOL were implemented by a district or center. Yet, given the findings on teachers' sustaining practice (Chapter 4), these costs may represent a one-year investment in teachers that has potential benefits for subsequent cohorts of children.

[4]Teachers received a $25 certificate for tickets to a local movie theater.

References

Achenbach, Thomas M. 1997. "Caregiver-Teacher Report Form (C-TRF)." Achenbach System of Empirically Based Assessment (ASEBA). Burlington: University of Vermont.

Alexander, Karl L., Doris R. Entwisle, and Nader Kabbani. 2001. "The Dropout Process in Life Course Perspective: Early Risk Factors at Home and School." *Teachers College Record* 103, 5: 760-822.

Barnett, W. Steven, Dale J. Epstein, Allison H. Friedman, Judi Stevenson Boyd, and Jason T. Hustedt. 2008a. "The State of Preschool 2008: State Preschool Yearbook." New Brunswick, NJ: National Institute for Early Education Research.

Barnett, W. Steven, Kwanghee Jung, Donald J. Yarosz, Jessica Thomas, Amy Hornbeck, Robert Stechuk, and Susan Burns. 2008b. "Educational Effects of the Tools of the Mind Curriculum: A Randomized Trial." *Early Childhood Research Quarterly* 23: 299-313.

Barrera, Manuel, Anthony Biglan, Ted K. Taylor, Barbara K. Gunn, Keith Smolkowski, Carol Black, Dennis V. Ary, and Rollen C. Fowler. 2002. "Early Elementary School Intervention to Reduce Conduct Problems: A Randomized Trial with Hispanic and Non-Hispanic Children." *Prevention Science* 3: 83-94.

Beebe-Frankenberger, Margaret, Kathleen M. Bocian, Donald L. MacMillan, and Frank M. Gresham. 2004. "Sorting Second-Grade Students: Differentiating Those Retained from Those Promoted." *Journal of Educational Psychology* 96, 2: 204-215.

Bierman, Karen L., Celene E. Domitrovich, Robert L. Nix, Scott D. Gest, Janet A. Welsh, Mark T. Greenberg, Clancy Blair, Keith E. Nelson, and Suhkdeep Gill. 2008. "Promoting Academic and Social-Emotional School Readiness: The Head Start REDI Program." *Child Development* 79, 6: 1802-1817.

Brennan, E. M., Jennifer R. Bradley, Mary Dallas Allen, and D. F. Perry. 2008. "The Evidence Base for Mental Health Consultation in Early Childhood Settings: Research Synthesis Addressing Staff and Program Outcomes." *Early Education and Development* 19, 6: 982-1022.

Brotman, Laurie Miller, Kathleen Kiely Gouley, Daniel Chesir-Teran, Tracy Dennis, Rachel G. Klein, and Patrick Shrout. 2005. "Prevention for Preschoolers at High Risk for Conduct Problems: Immediate Outcomes on Parenting Practices and Child Social Competence." *Journal of Clinical Child and Adolescent Psychology* 34, 4: 724-734.

Burchinal, Margaret, Nathan Vandergrift, Robert Pianta, and Andrew Mashburn. 2010. "Threshold Analysis of Association Between Child Care Quality and Child Outcomes for Low-Income Children in Pre-Kindergarten Programs." *Early Childhood Research Quarterly* 25, 2: 166-176.

Campbell, Susan B. 1995. "Behavior Problems in Preschool Children: A Review of Recent Research." *Journal of Child Psychology and Psychiatry* 36: 113-149.

Centers for Disease Control and Prevention, National Center for Health Statistics. 2005. *NCHS Data About Special Education*. Atlanta, GA: Centers for Disease Control and Prevention.

Chambers, Jay G., Yael Kidron, and Angeline K. Spain. 2004. *Characteristics of High-Expenditure Students with Disabilities, 1999-2000.* Washington, DC: American Institutes for Research.

Cooper, David H., and Dale C. Farran. 1991. "Behavioral Risk Factors in Kindergarten." *Early Childhood Research Quarterly* 3: 1-19.

Cryer, Debby, Thelma Harms, and Cathy Riley. 2003. *All About the ECERS-R.* Lewisville, NC: Kaplan PACT House Publishing.

Curbow, Barbara, Kai Spratt, Antoinette Ungaretti, Karen McDonnell, and Steven Breckler. 2000. "Development of the Child Care Worker Job Stress Inventory." *Early Childhood Research Quarterly* 15, 4: 515-536.

Diamond, Adele, W. Steven Barnett, Jessica Thomas, and Sarah Munro. 2007. "Supporting Online Material for Preschool Program Improves Cognitive Control." *Science* 318: 1387-1411.

Dishion, Thomas J., Doran C. French, and Gerald R. Patterson. 1995. "The Development and Ecology of Antisocial Behavior." Pages 421-471 in Dante Cicchetti and Donald J. Cohen (eds.), *Developmental Psychopathology: Risk, Disorder, and Adaptation* (vol. 2). New York: Wiley.

Downer, Jason T., Bridget K. Hamre, Robert C. Pianta, Olivia Lima, Bonnie Yoder, and Leslie Booren. 2008. "Classroom Assessment Scoring System — Child Version (CLASS-C)." Unpublished measure. Charlottesville: University of Virginia.

Dumas, Jean E., Ronald J. Prinz, Emilie Phillips Smith, and James Laughlin. 1999. "The EARLY ALLIANCE Prevention Trial: An Integrated Set of Interventions to Promote Competence and Reduce Risk for Conduct Disorder, Substance Abuse, and School Failure." *Clinical Child and Family Psychology Review* 2: 37-53.

Early, Diane M., Iheoma U. Iruka, Sharon Ritchie, Oscar A. Barbarin, Donna-Marie C. Winn, Gisele M. Crawford, Pamela M. Frome, Richard M. Clifford, Margaret Burchinal, Carollee Howes, Donna M. Bryant, and Robert C. Pianta. 2010. "How Do Pre-Kindergarteners Spend Their Time? Gender, Ethnicity, and Income as Predictors of Experiences in Pre-Kindergarten Classrooms." *Early Childhood Research Quarterly* 25: 177-193.

Entwisle, Doris R., and Leslie Alec Hayduk. 1988. "Lasting Effects of Elementary School." *Sociology of Education* 61, 3: 147-159.

Gorman-Smith, Deborah, Patrick H. Tolan, David B. Henry, Elena Quintana, Kelly Lutovsky, and Amy Leventhal. 2006. "Schools and Families Educating Children: A Preventive Intervention for Early Elementary School Children." Pages 113-135 in Patrick Tolan, José Szapocznik, and Soledad Sambrano (eds.), *Preventing Youth Substance Abuse: Science-Based Programs for Children and Adolescents.* Washington, DC: American Psychological Association.

Kaiser, Ann P., Xinsheng Cai, Terry B. Hancock, and E. Michael Foster. 2002. "Teacher-Reported Behavior Problems and Language Delays in Boys and Girls Enrolled in Head Start." *Behavioral Disorders* 28, 1: 23-39.

Kessler, Ronald C., Peggy R. Barker, Lisa J. Colpe, Joan F. Epstein, Joseph C. Gfroerer, Eva Hiripi, Mary J. Howes, Sharon-Lise T. Normand, Ronald W. Manderscheid, Ellen E. Walters, and Alan M. Zaslavsky. 2003. "Screening for Serious Mental Illness in the General Population." *Archives of General Psychiatry* 60: 184-189.

La Paro, Karen M., Robert C. Pianta, and Megan Stuhlman. 2004. "The Classroom Assessment Scoring System: Findings from the Prekindergarten Year." *Elementary School Journal* 104, 5: 409-426.

Lavigne, John V., Robert D. Gibbons, Katherine Kaufer Christoffel, Richard Arend, Diane Rosenbaum, Helen Binns, Nichole Dawson, Hollie Sobel, and Crystal Isaacs. 1996. "Prevalence Rates and Correlates of Psychiatric Disorders Among Preschool Children." *Journal of the American Academy of Child and Adolescent Psychiatry* 35: 204-214.

Lee, Valerie E., and David T. Burkam. 2002. *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School.* Washington, DC: Economic Policy Institute. Web site: http://www.asu.edu/educ/epsl/EPRU/Articles/EPRU-0603-138-OWI.pdf.

Li-Grining, Christine, C. Cybele Raver, Kina Champion, Latriese Sardin, Molly Metzger, and Stephanie M. Jones. 2010. "Understanding and Improving Classroom Emotional Climate and Behavior Management in the 'Real World': The Role of Head Start Teachers' Psychosocial Stressors." *Early Education and Development* 21, 1: 65-94.

Lloyd, Chrishana M., and Michael Bangser. 2009. *Promoting Preschool Quality Through Effective Classroom Management: Implementation Lessons from the Foundations of Learning Demonstration.* New York: MDRC.

Magnuson, Katherine A., Christopher Ruhm, and Jane Waldfogel. 2007. "Does Prekindergarten Improve School Preparation and Performance?" *Economics of Education Review* 26, 1: 33-51.

New Jersey Department of Education. 2008. "Preschool Program Implementation Guidelines." Trenton: New Jersey Department of Education.
Web site: http://www.state.nj.us/education/ece/dap/.

Patterson, Gerald R. 1982. *Coercive Family Process.* Eugene, OR: Castalia Publishing Company.

Patterson, Gerald R. 1996. "Some Characteristics of a Developmental Theory for Early Onset Delinquency." Pages 81-124 in Mark F. Lenzenweger and Jeffery J. Haugaard (eds.), *Frontiers of Developmental Psychopathology.* New York: Oxford University Press.

Patterson, Gerald R., John B. Reid, and Thomas J. Dishion. 1992. *Antisocial Boys.* Eugene, OR: Castalia Publishing Company.

Pianta, Robert C. 2001. *STRS Student-Teacher Relationship Scale: Professional Manual.* Lutz, FL: Psychological Assessment Resources.

Pianta, Robert, Carollee Howes, Margaret Burchinal, Donna Bryant, Richard Clifford, Diane Early, and Oscar Barbarin. 2005. "Features of Pre-Kindergarten Programs, Classrooms, and Teachers: Do They Predict Observed Classroom Quality and Child-Teacher Interactions?" *Applied Developmental Science* 9, 3: 144-159.

Pianta, Robert C., Karen M. LaParo, and Bridget K. Hamre. 2006. *Classroom Assessment Scoring System Manual, Preschool (Pre-K) Version.* Charlottesville: University of Virginia, Center for Advanced Study of Teaching and Learning.

Quint, Janet C., Johannes M. Bos, and Denise F. Polit. 1997. *New Chance: Final Report on a Comprehensive Program for Young Mothers in Poverty and Their Children.* New York: MDRC.

Raver, C. Cybele. 2002. "Emotions Matter: Making the Case for the Role of Young Children's Emotional Development for Early School Readiness." *Social Policy Reports* 16, 3: 3-18.

Raver, C. Cybele, Stephanie M. Jones, Christine P. Li-Grining, Molly Metzger, Kina Smallwood, and Latriese Sardin. 2008. "Improving Preschool Classroom Processes: Preliminary Findings from a Randomized Trial Implemented in Head Start Settings." *Early Childhood Research Quarterly* 63, 3: 253-255.

Raver, C. Cybele, Stephanie M. Jones, Christine P. Li-Grining, Fuhua Zhai, Kristen Bub, and Emily Pressler. 2009a. "CSRP's Impact on Low-Income Preschoolers' Pre-Academic Skills: Self-Regulation as a Mediating Mechanism." Unpublished paper.

Raver, C. Cybele, Stephanie M. Jones, Christine P. Li-Grining, Fuhua Zhai, Molly Metzger, and Bonnie Solomon. 2009b. "Targeting Children's Behavior Problems in Preschool Classrooms: A Cluster-Randomized Controlled Trial." *Journal of Consulting and Clinical Psychology* 77: 302-316.

Rimm-Kaufman, Sara, Robert C. Pianta, and Martha Cox. 2000. "Teachers' Judgments of Problems in the Transition to Kindergarten." *Early Childhood Research Quarterly* 15, 2: 147-166.

U.S. Department of Education, National Center for Education Statistics. n.d. *Academic Rating Scale.* Washington, DC: National Center for Education Statistics.

U.S. Department of Education, National Center for Education Statistics. 2001. *Entering Kindergarten: A Portrait of American Children When They Begin School: Findings from the Condition of Education 2000.* NCES 2001-035. Nicholas Zill and Jerry West. Washington, DC: U.S. Government Printing Office.

U.S. Department of Labor, Bureau of Labor Statistics. 2009. "Occupational Employment and Wages, May 2009: 25-2011 Preschool Teachers, Except Special Education."
Web site: http://www.bls.gov/oes/current/oes252011.htm.

U.S. Office of Management and Budget. 2010. President Obama's Fiscal 2010 Budget: Giving Every Child a World-Class Education.
Web site: http://www.whitehouse.gov/omb/fy2010_key_education/.

Wagner, Mary, and Jose Blackorby. 2002. *Disability Profiles of Elementary and Middle School Students with Disabilities.* Washington, DC: U.S. Department of Education, Office of Special Education Programs.

Webster-Stratton, Carolyn, M. Jamila Reid, and Mary Hammond. 2001. "Preventing Conduct Problems, Promoting Social Competence: A Parent and Teacher Training Partnership in Head Start." *Journal of Clinical Child Psychology* 30, 3: 238-302.

Zill, Nicholas, and James L. Peterson. 1986. *Behavior Problems Index.* Washington, DC: Child Trends.

Earlier MDRC Publications on the Foundations of Learning Demonstration

Can Teacher Training in Classroom Management Make a Difference for Children's Experiences in Preschool? A Preview of Findings from the Foundations of Learning Demonstration.
2009. Pamela Morris, Cybele Raver, Chrishana M. Lloyd, and Megan Millenky.

Promoting Preschool Quality Through Effective Classroom Management: Implementation Lessons from the Foundations of Learning Demonstration.
2009. Chrishana M. Lloyd and Michael Bangser with Farrah Parkes.

About MDRC

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

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

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

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

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