Impact and Implementation Findings from an Experimental Evaluation of Playworks: Effects on School Climate, Academic Learning, Student Social Skills and Behavior

MATHEMATICA Policy Research

john w. gardner center for youth and their communities

Robert Wood Johnson Foundation

EVALUATION REPORT

May 2013

Impact and Implementation Findings from an Experimental Evaluation of Playworks: Effects on School Climate, Academic Learning, Student Social Skills and Behavior

TABLE OF CONTENTS

RESEARCHERS

Mathematica Policy Research
Jane Fortson
Susanne James-Burdumy
Martha Bleeker
Nicholas Beyler

John W. Gardner Center for Youth and Their Communities, Stanford University
Rebecca A. London
Lisa Westrich
Katie Stokes-Guinan
Sebastian Castrechini

EVALUATION REPORT
MAY 2013

Executive Summary

Recess periods often lack the structure needed to support physical activity and positive social development (Robert Wood Johnson Foundation 2010). The Playworks program places full-time coaches in low-income schools to provide opportunities for organized play during recess and throughout the school day. Playworks activities are designed to engage students in physical activity, foster social skills related to cooperation and conflict resolution, improve students' ability to focus on class work, decrease behavioral problems and improve school climate.

The Robert Wood Johnson Foundation (RWJF) contracted with Mathematica Policy Research and its subcontractor, the John W. Gardner Center for Youth and Their Communities (JGC) at Stanford University, to conduct a rigorous evaluation of Playworks. Twenty-nine schools interested in implementing Playworks were randomly assigned to treatment and control groups during the 2010-2011 (cohort 1) or 2011-2012 (cohort 2) school year. During the one-year study period for each cohort, treatment schools received Playworks and control schools were not eligible to implement Playworks. We collected data from students, teachers and school staff at 25 cohort 1 schools in spring 2011 and an additional 4 cohort 2 schools in spring 2012 to document the implementation of Playworks and assess the impact of the program on key outcomes in five domains: (1) school climate, (2) conflict resolution and aggression, (3) learning and academic performance, (4) youth development and (5) student behavior.

Additional analyses on the implementation of Playworks and its impact on play, physical activity and recess outcomes will be released in two separate briefs.

Key Findings

The following significant, positive impacts of Playworks were found:

- There was a positive impact of Playworks on teachers' reports of students using positive, encouraging language; teachers' perceptions of the extent to which students felt safe at school; and teachers' perceptions of the extent to which students felt safe and included during recess. Teachers in treatment schools were also significantly more likely than teachers in control schools to report that school staff support organized play during the school day.
- Teachers in treatment schools reported less bullying and exclusionary behavior than teachers in control schools.
- Teachers in treatment schools were less likely to report difficulties in transitioning to classroom learning activities after recess and reported taking significantly less time to transition from recess to learning activities than teachers in control schools. Treatment students were also more likely than control students to report better behavior and attention in class after sports, games and play.

Teachers in treatment schools reported less bullying and exclusionary behavior than teachers in control schools.

The following key implementation findings were observed:

- Strong implementation occurred in 8 of 17 treatment schools and moderate implementation occurred in another 6 schools. Three schools had weak implementation.[1]
- Playworks implementation was stronger in schools that had recess in the past.
- Most teachers, students and principals had positive perceptions of the Playworks program.

Impact and Implementation Findings from an Experimental Evaluation of Playworks: Effects on School Climate, Academic Learning, Student Social Skills and Behavior

Background

Recess has been reduced or eliminated in up to 40 percent of school districts across the country (Zygmunt-Fillwalk and Bilello 2005), and these declines have disproportionately affected low-income minority students in urban areas (Barros, Silver and Stein 2009). In schools where recess is still offered, recess periods often lack the structure needed to support physical activity and positive social development (Robert Wood Johnson Foundation 2010). Research suggests that participating in physical activity and structured play during recess may improve academic and behavioral outcomes (Centers for Disease Control and Prevention [CDC] 2010; U.S. Department of Health and Human Services [DHHS] 2008; Ginsburg 2007; Leff, Costigan and Power 2004).

Research suggests that participating in physical activity and structured play during recess may improve academic and behavioral outcomes.

The Playworks program places full-time coaches in low-income schools to provide opportunities for organized play during recess and class time. Playworks activities are designed to engage students in physical activity, foster social skills related to cooperation and conflict resolution, improve students' ability to focus on class work, decrease behavioral problems and improve school climate. The Playworks model includes the following components, the first three of which are examined in this study:

- ***Organized Recess Activities.*** During recess, the coach teaches conflict resolution skills and fosters student play by encouraging involvement in organized, inclusive activities. The coach introduces a common set of rules to games and models conflict resolution tools such as ro-sham-bo (rock-paper-scissors), with the goal of reducing the number of conflicts that arise, enabling youth to resolve their own disputes and creating an environment that supports positive play.

- ***Class Game Time.*** Class game time is a period in which the coach meets with individual classes to lead games with the students. During this time, the coach uses games to foster team work and positive play while teaching students the rules to new games that they can play at recess. Teachers are required to be present and are encouraged to play alongside their students at class game time.
- ***Junior Coach Program.*** This program provides 4th- and 5th-grade students with an opportunity to develop leadership and conflict resolution skills so they can act as role models and facilitators during recess.
- ***After-School Activities.*** Playworks also includes an after-school program, a sports league and school staff trainings.

Past Research

The strategies and practices used in Playworks, such as engaging students in physical activity and organized recess activities, have been shown to be positively related to desirable student outcomes (CDC 2010; DHHS 2008; Ginsburg 2007). For instance, in terms of academic behaviors, physical activity among children has been associated with improvements in cognition (Ginsburg 2007; Tomporowski et al. 2008; Sibley and Etnier 2003), on-task behavior (Mahar et al. 2006; Jarrett et al. 1998), problem solving (Molloy 1989), and concentration and attentiveness (Taras 2005; Pellegrini, Huberty and Jones 1995; Evenson et al. 2009; Caterino and Polak 1999). Moreover, a comprehensive report published by the CDC (2010) reviewed eight studies that looked at academic performance and recess in elementary schools and found that children who spent time in recess appeared to have increased attention, concentration and on-task behavior in the classroom. Although recess may take away from classroom time, there is no evidence that time spent in recess is negatively associated with cognitive skills, attitudes, academic behaviors or academic achievement (CDC 2010; DHHS 2008; Trudeau and Shephard 2010; Taras 2005; Ericsson 2008; Maeda and Randall 2003; Ahamed et al. 2007; Coe et al. 2006). In fact, some evidence points to improved academic achievement as a result of increased physical activity (CDC 2010; Nelson and Gordon-Larsen 2006; Shephard 1997; Tremarche, Robinson and Graham 2007; Smith and Lounsbery 2009). Moreover, a recently published research brief found that 11 out of 14 published studies analyzing relationships between physical activity and academic performance determined that regular physical activity was associated with improved academic performance (Active Living Research 2009).

There is also some evidence suggesting that participating in play activities at recess is associated with prosocial behaviors (Ginsburg 2007). Through play at recess, students have opportunities to develop social relationships with their peers (Pellegrini and Bohn 2005; Pellegrini et al. 2002) and to experiment with social strategies such as sharing, problem solving and conflict resolution (Zygmunt-Fillwalk and Bilello 2005; Molloy 1989). Structured play during recess is also associated with decreases in aggression and bullying (Leff, Costigan and Power 2004). Another study found that increased physical activity during the school day improved classroom behavior according to teacher reports (Maeda and Randall 2003). The duration of recess appears to be less important, however. One study found that, among children who received daily recess, an increase in the length of recess was not associated with improved teacher ratings of students' classroom behavior (Barros, Silver and Stein 2009).

These findings suggest that the underlying strategies and practices that define the Playworks program—which provides opportunities for organized play during recess and class time—may have the potential to improve academic and behavioral outcomes of students. A recent non-experimental study (London et al. 2010) investigated the ways in which the Playworks program was implemented in eight schools in the San Francisco Bay area. The findings from this study suggested that when Playworks was fully implemented, recess was more structured and organized, students were more engaged during recess activities and students learned to use conflict resolution skills. However, a more rigorous evaluation of Playworks is warranted; the current study investigates the effectiveness of Playworks using a randomized, experimental design.

Current Study

RWJF contracted with Mathematica Policy Research and its subcontractor, Stanford University's John W. Gardner Center for Youth and Their Communities (JGC), to conduct a rigorous evaluation of the implementation and impact of Playworks. Twenty-nine schools interested in implementing Playworks were randomly assigned to treatment and control groups during the 2010-2011 (cohort 1) or 2011-2012 (cohort 2) school year. During the one-year study period for each cohort, treatment schools received Playworks and control schools were not eligible to implement Playworks.

Findings based on the first cohort of 25 schools (that participated during the 2010-2011 school year) showed that Playworks had positive impacts on some measures of school climate, conflict resolution and aggression, learning and academic performance, and recess experience, and showed no negative impacts of the program in any of the six domains that were assessed (Bleeker et al. 2012). In the first cohort of schools, Playworks implementation was observed to be moderate or strong in 12 of 14 study schools, with stronger implementation occurring in schools that had recess in the past or had Playworks coaches that were experienced with the program. Findings also showed that participants had positive perceptions of the program.

This study expands on these previous analyses in two ways: (1) incorporating data from four additional cohort 2 study schools that participated during the 2011-2012 school year and (2) exploring additional impacts on academic and behavioral outcomes based on

administrative records. In particular, we address the following research question relating to the program's impact:

1. What is the effect of Playworks on five outcome domains?

   (1) school climate

   (2) conflict resolution and aggression

   (3) learning and academic performance

   (4) youth development and

   (5) student behavior[2]

This brief also addresses three research questions related to program implementation:

1. How was Playworks implemented in the treatment schools?
2. In what context was Playworks implemented?
3. What were school staff and students' experiences with and perceptions of Playworks?[3]

Study Design

Twenty-nine schools (17 treatment schools, 12 control schools) from six cities across the U.S. were recruited for the Playworks evaluation. Twenty-five of these schools (14 treatment schools, 11 control schools) were drawn from five cities and participated in the study during the 2010-2011 school year (we refer to these schools as cohort 1 schools). Four additional schools (3 treatment schools, 1 control school) from one additional city participated in the study during the 2011-2012 school year (we refer to these schools as cohort 2 schools). Random assignment of schools was used to determine which study schools would implement Playworks during the study year (and which schools would be eligible to implement Playworks in the following school year). Random assignment of schools helped to ensure that there were no systematic differences between the treatment and control groups' observed and unobserved characteristics and that the differences in outcomes between the two groups could be attributed solely to the effect of Playworks. To improve the statistical precision of impact estimates and reduce the chance of differences between the treatment and control groups in the characteristics of schools, random assignment was conducted within matched groups (pairs, trios or foursomes) of schools that were similar in terms of observable characteristics (see Appendix 1 for additional details on random assignment).

Baseline comparisons of the evaluation's treatment and control schools were conducted based on data from the Common Core of Data and time-invariant characteristics of students and teachers from the student and teacher surveys (see Appendix 2, Tables 1 and 2). Only two significant differences were found between treatment and control schools, teachers and students: treatment teachers were significantly more likely to be white and significantly less likely to be African American, relative to control teachers. These two significant differences are what one might expect by chance when conducting 32 tests with a 5 percent critical value (that is, 5 percent of 32 is 1.6, which rounds to 2). We included race indicators in the impact models that use teacher survey data to account for these differences.

Outcomes and Data Sources

The evaluation's data collection activities were designed to document the implementation of Playworks, collect information on the costs of implementing the program, and measure key outcomes in the five domains listed above, as well as in the areas of play, physical activity and recess (which will be presented in a separate brief). To that end, we collected follow-up data from students, teachers and school staff at 25 schools in spring 2011 and 4 additional schools in spring 2012, roughly seven months after Playworks was first implemented in treatment schools. The data collection activities relevant for this brief include:

- ***Student Survey.*** A total of 2,331 students from 119 4th- and 5th-grade classrooms in 28 study schools participated in a survey that included items related to school climate, conflict resolution, learning and achievement and relationships with adults and peers.
- ***Teacher Survey.*** A total of 296 teachers from 29 study schools participated in a survey that included items related to school climate and students' behavior, learning, achievement and social competence.
- ***Administrative Records.*** All 29 study schools provided a list of teachers and student rosters for each classroom that was selected for participation in the study. Schools also provided information on outcomes for the study year, including achievement test results, average daily attendance, chronic absences, suspensions and expulsions.
- ***Interviews with Principals, Teachers and Playworks Coaches.*** A total of 32 principals and assistant principals, 51 teachers and 17 Playworks coaches responded to questions about opportunities for play and physical activity at school, discipline issues that arise at recess and experiences with and perceptions of Playworks.
- ***Focus Groups with Junior Coaches.*** Students from 16 treatment schools who served as Playworks junior coaches responded to questions about their experiences as junior coaches and perceptions of Playworks.
- ***Playworks Observations.*** Playworks coach involvement and strategies, student participation in Playworks games, and playground monitor and teacher activities were observed during recess and Playworks class game time at all 17 treatment schools.

Two separate study briefs will report on findings based on additional outcome and implementation data that were collected as part of the evaluation. Response rates and additional details about each of the study's data sources (including data that will be presented in other briefs) can be found in Appendix 1.

Key Findings

A. Impact Findings

For this brief, we examined the effect of Playworks on five outcome domains. Significant impacts were observed in domains covering school climate, conflict resolution and aggression, and learning and academic performance, suggesting that Playworks had positive effects. No significant impacts were detected in the other two domains addressing outcomes related to youth development and student behavior (see Appendix 1 for additional details on our approach for estimating impacts and the methods used to adjust *p*-values for multiple hypothesis testing). A subset of the impact results is summarized by domain in the exhibits below. A full set of tables that define each scale and display the impact results for each outcome is provided in Appendix 2.

School Climate. Playworks had a positive impact on four of the five teacher-reported measures of school climate but had no significant impact on the three student-reported measures of school climate (see Exhibit 1 and Appendix 2, Table 3). In particular, with regard to sense of community and feelings of safety at school, teachers in treatment schools were significantly more likely than teachers in control schools to report that students used positive, encouraging language and to report positive perceptions of students' safety and engagement in inclusive behavior at recess. Playworks had no significant impact, however, on students' feelings of safety at recess or school or about how well students and teachers treat each other within the school community. Our observations of recess in treatment schools showed Playworks coaches promoting inclusive behavior in 57 percent of recesses observed (Appendix 2, Table 4).

EXHIBIT 1.
Impacts on School Climate

Outcome (mean unless otherwise noted below)	**Treatment**	**Control**	**Difference**
School as Community			
Student-Reported Sense of School as Community Scale Score	2.8	2.7	0.1
Percentage of Teachers that "Agree" or "Strongly Agree" that Students in Their School Use Positive, Encouraging Language	53.9	27.4	26.5**
Feelings of Safety			
Student-Reported Feelings of Safety at School Scale Score	2.6	2.5	0.1
Student-Reported Feelings of Safety at Recess Scale Score	2.8	2.6	0.1
Teacher-Reported Feelings of Students' Safety at School Scale Score	3.8	3.2	0.6**
Teacher-Reported Feelings of Students' Safety/Inclusion at Recess Scale Score	3.9	3.1	0.8***
Support for Organized Play			
Teacher Support for Organized Play During the School Day Scale Score	4.2	4.0	0.1
Teacher-Reported School Staff Support for Organized Play During the School Day Scale Score	4.3	3.9	0.4**

Sources: Student (n = 2,283) and teacher surveys (n = 296) conducted in spring 2011 or spring 2012 (sample sizes may be smaller for some outcomes due to missing responses).

Note: See full table in Appendix 2, Table 3.

** Significantly different from zero at the .05 level, two-tailed test.

*** Significantly different from zero at the .01 level, two-tailed test.

Findings related to the impact of Playworks on support for organized play were mixed. Teachers in treatment schools were significantly more likely to report that, in general, staff at their school supported organized play during the school day (for activities like physical education class and Playworks) than teachers in control schools. However, when asked about the extent to which they personally supported organized play, treatment teachers were no more likely to report support for organized play than control teachers.

Teachers in treatment schools were less likely to report difficulties in transitioning to classroom learning activities after recess…

Conflict Resolution and Aggression. Teachers in treatment schools reported significantly less bullying and exclusionary behavior than teachers in control schools (Exhibit 2 and Appendix 2, Table 5). However, no significant impacts were found on teacher reports of the Behavior Assessment System for Children (BASC) aggression subscale (for example, talking back to teachers and showing off), student reports of aggressive behavior, students' beliefs about aggression or students' reports on their relationships with other students (for example, getting along well with others at recess and being able to resolve conflicts without fighting). Rather than striving to eliminate all conflict, Playworks aims to give students the tools to better manage conflicts when they arise. There was evidence that the junior coach program provided selected students in grades four and five with the opportunity to develop conflict resolution skills. Most teachers who participated in the teacher survey reported that junior coaches helped resolve conflicts (67 percent) (Appendix 2, Table 6). We observed junior coaches intervening in conflicts at 71 percent of schools and in 21 percent of recesses; these junior coaches had varying degrees of success at resolving the conflicts. When asked about conflict resolution in the focus groups, junior coaches from nearly all schools (88 percent) reported that they used ro-sham-bo at recess to resolve conflicts.

EXHIBIT 2.
Impacts on Conflict Resolution and Aggression

Outcome (mean)	**Treatment**	**Control**	**Difference**
Interactions with Other Students			
Student-Reported Relationships with Other Students Scale Score	3.1	3.1	0.1
Teacher-Reported Student Bullying/Exclusion Scale Score	0.6	1.0	-0.4**
Aggression			
Student-Reported Aggressive Behavior Scale Score	1.4	1.5	-0.1
Student-Reported Normative Beliefs About Aggression Scale Score	1.6	1.7	-0.1
Teacher-Reported Student BASC Aggression Subscale Score	6.0	6.6	-0.6

Sources: Student (n = 2,288) and teacher surveys (n = 295) conducted in spring 2011 or spring 2012 (sample sizes may be smaller for some outcomes due to missing responses).

Note: See full table in Appendix 2, Table 5.

** Significantly different from zero at the .05 level, two-tailed test.

Learning and Academic Performance. Playworks had a positive impact on both student and teacher perceptions of the transition from recess to classroom activities (Exhibit 3 and Appendix 2, Table 7). Students in both treatment and control schools were asked about the effect of sports, games and play on their behavior in class; treatment students were significantly more likely to report better behavior and attention in class after participating

in sports, games and play than control students. Similarly, teachers in treatment schools were significantly less likely to report difficulties in transitioning to learning activities after recess and reported significantly less time taken to transition from recess to learning activities than teachers in control schools (a difference of 3.5 minutes on the most recent day in which students participated in recess). There were no significant differences on six additional outcome measures that assessed student engagement with classroom activities and academic performance, homework completion and motivation to succeed academically. Finally, there were no significant differences between treatment and control schools in the percentage of 3rd through 5th grade students showing proficiency in reading and math.

EXHIBIT 3.
Impacts on Learning and Academic Performance

Outcome (mean unless otherwise noted below)	**Treatment**	**Control**	**Difference**
Transition from Recess to Classroom Activities			
Student-Reported Effect of Recess on Behavior in Class Scale Score	2.5	2.4	0.1
Student-Reported Effect of Sports, Games and Play on Behavior in Class Scale Score	2.7	2.6	0.2**
Percentage of Students That Report That It Is "Somewhat True" or "Very True" That Beginning Class Work After Recess Is Easy	59.3	54.2	5.1
Teacher-Reported Number of Minutes to Transition from Recess to Learning Activities	6.8	10.3	-3.5*
Teacher-Reported Difficult Transition to Learning After Recess Scale Score	2.4	3.1	-0.7***

Sources: Student (n = 2,279) and teacher surveys (n = 293) conducted in spring 2011 or spring 2012 (sample sizes may be smaller for some outcomes due to missing responses).

Note: See full table in Appendix 2, Table 7.

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

*** Significantly different from zero at the .01 level, two-tailed test.

During our interviews, we asked treatment teachers an open-ended question about how Playworks was related to students' behavior in their classroom. Several themes emerged; for instance, 29 percent of teachers reported that students were now more likely to come to class ready to learn (compared with teachers' prior experiences) because fewer conflicts carried over from recess, 37 percent of teachers reported that Playworks resulted in students using ro-sham-bo in class to resolve conflicts or make decisions and 14 percent of teachers reported improvements in students' teamwork and inclusiveness in class (compared with teachers' prior experiences; Appendix 2, Table 8). Some teachers reported that Playworks served as an incentive to positively influence students' class performance because students did not want to lose the opportunity to participate in Playworks activities. Less than a quarter of teachers reported that Playworks positively affected their practices in the classroom; in particular, 22 percent reported using Playworks games on their own, 18 percent reported using group facilitation techniques and management strategies learned from Playworks and 16 percent reported spending less time dealing with conflict in the classroom (Appendix 2, Table 8). Finally, very few teachers reported that the junior coach program, which in some schools required students to miss class time, was a detriment to the academic performance of the junior coaches.

Youth Development. There were no significant impacts of Playworks on eight measures of youth development. In particular, students in treatment and control schools had similar reports on a six-item scale that measured feelings about adult interactions (such as "At my school, there is an adult who listens to me when I have something to say"). In addition, a similar percentage of treatment and control students reported getting along well with other students. There was also no significant difference on a scale that included eight items asking students to indicate their effectiveness at interacting with peers in conflict situations, such as their ability to tell kids to stop teasing a friend. Teachers in treatment and control schools also reported similar perceptions of students' abilities to regulate their emotions, act responsibly and engage in prosocial and altruistic behavior (Appendix 2, Table 9).

Student Behavior. Despite the fact that most treatment teachers who responded to the survey felt that Playworks reinforced positive behavior during recess (96 percent) and resulted in fewer students getting into trouble (89 percent) (Appendix 2, Table 10), there were no significant impacts of Playworks on multiple indicators of student behavior. Treatment and control group students who took the student survey reported similar levels of disruptive behavior in class and behavioral problems at school (Appendix 2, Table 11). Teachers in treatment and control schools reported similar amounts of student misbehavior, absences, tardiness, suspensions and detentions among their students. Administrative records showed no significant differences between treatment and control schools in average daily attendance, the percentage of students who were chronically absent, the suspension rate, or the expulsion rate. The number of disciplinary incidents in the treatment and control schools, measured via discipline referral data gathered from principals, was also not significantly different overall, by setting (for example, at recess), or by reason (for fighting, profanity and so on). One caveat with respect to the findings based on the administrative records and discipline referral data is that the findings are based on a smaller sample size (between 24 and 29 schools, depending on the outcome).

Discussion of Impact Findings. This study draws on many data sources to present a comprehensive picture of the impacts of Playworks from a variety of perspectives. One interesting pattern in the impact findings is that most of the positive impacts we observed were found on outcomes reported by teachers (seven of eight significant impact findings), as opposed to students (one of eight significant impact findings). We explored several potential explanations for this pattern (for a more in-depth discussion, see Appendix 3). Though we can rule out many possible explanations, we cannot rule out the possibility that these differences simply represent true differences in the experiences or perceptions of teachers and students or differences in the way that survey questions were framed. Nonetheless, the pattern of findings suggests that Playworks led to improvements on some measures of school climate, conflict resolution/aggression and learning/academic performance, particularly on measures reported by teachers.

B. Cost Analysis

As described above, the impact analysis showed that Playworks had a positive impact on outcomes in the school climate, conflict resolution/aggression and learning/academic performance domains. These impacts suggest that Playworks was beneficial to schools, teachers and students along multiple dimensions. To put these findings into context for school administrators and policymakers who are comparing Playworks to other programs, we conducted a cost analysis. Our estimates of Playworks' cost can be used with the impact estimates to provide an indication of the magnitude of impacts relative to the program's cost. When compared to other cost and impact estimates, administrators and policymakers can choose between different programs with different benefits and costs. In Appendix 3, we describe the cost analysis and provide information that could be used to conduct a comparison to other programs.

C. Implementation Findings

The implementation component of the evaluation assessed key program goals, how Playworks was implemented in treatment schools, the context within which the program was implemented and student and staff perceptions of Playworks. Key findings in each of these areas are described below. A full set of tables that define each measure and display all implementation findings is provided in Appendix 2.

Principals Described the Key Program Goals. Principals were the main driving force behind bringing the Playworks program to schools. According to principals, key goals for Playworks were to (1) organize recess, (2) improve overall school climate and help students work together, (3) improve school safety and reduce conflicts and (4) increase student physical activity levels. Less frequently cited goals were to support the introduction of recess and to promote student leadership (Appendix 2, Table 12).

Site Visits Suggest Strong Implementation Occurred in Eight Study Schools and Moderate Implementation Occurred in Six Schools. Overall, we observed strong implementation of the Playworks program in 8 of the 17 treatment schools, with moderate implementation observed in 6 schools and weak implementation observed in 3 schools.[4] We defined schools as having "strong" implementation if the following were observed during site visits:

- Recess was structured and organized, students were engaged in games and other play activities, coaches were engaged with students, junior coaches were doing their jobs, positive and inclusive language was being used and conflicts were resolved quickly.
- Teachers, principals and other staff were knowledgeable about Playworks and supportive of its values and goals.
- Principals were willing to schedule regular class game times and accommodate junior coaches to work at younger students' recesses.
- School policies and structures supported Playworks activities and goals.

Schools were categorized as having "moderate" implementation when we observed most program components to be in place and commitment was strong from some, but not all, staff members. We categorized schools as having "weak" implementation when staff commitment to the program was not strong and key components were not implemented as intended. In the three schools with weak implementation, contextual factors—such as whether the school was accustomed to providing students with recess and the principal's buy-in for the program—made it difficult for the Playworks coach to implement the program as intended.

In addition to looking at overall program implementation, we examined implementation of specific key components of Playworks and found the following:

- **Program strategies were modeled by coaches and used by students during the majority of recess periods at treatment schools.** Coaches used positive messaging such as "good job, nice try" at an average of 67 percent of recesses observed at treatment schools. They promoted inclusive behavior, encouraging students to join games and participate in activities at an average of 57 percent of recesses and were observed playing with students at an average of 67 percent of recesses. When coaches were not playing, they typically moved around the playground to supervise games or manage conflicts. In 13 of 17 schools and an average of 33 percent of recesses, we observed students using ro-sham-bo to resolve minor conflicts at recess, either on their own or with encouragement from an adult (Appendix 2, Table 4). Among student respondents to the survey, an average of 68 percent reported using ro-sham-bo at recess (Appendix 2, Table 14).
- **Class game time provided a fun opportunity for coaches to model Playworks strategies and techniques in a smaller group setting.** Coaches used positive language at an average of 92 percent of class game time periods observed in treatment schools, and teachers played with their students at an average of 39 percent of class game time periods (Appendix 2, Table 4). Most teachers who participated in the teacher survey reported that class game time was fun for students (92 percent), provided students with good exercise (88 percent) and helped students learn new games (90 percent). Most teachers (85 percent) reported in the survey that they viewed class game time as an important part of Playworks (Appendix 2, Table 13). In interviews, teachers who were not as supportive of class game time mentioned several challenges, including scheduling problems, the inability of the coach to work effectively with students or teachers and a concern—especially among 5th-grade teachers—that it interfered with instructional time.

- **The junior coach program provided students with the opportunity to gain leadership skills.** Junior coaches at all schools were scheduled to support at least one recess per week and had an opportunity to work with younger students at recess. In five schools, upper and lower grades had combined or overlapping recess periods, allowing junior coaches to act as role models for younger students during their own recess time. Students at seven schools were allowed to miss some class time to work at younger students' recesses, though not all younger students' recesses had a junior coach. Junior coaches at five schools ate quickly and worked their Playworks shifts during their regular lunchtime. Nearly all teachers felt that students who served as junior coaches gained leadership skills (89 percent), taught other students games (81 percent) and enjoyed their role at recess (88 percent) (Appendix 2, Table 6).
- **Coaches cited several challenges to implementing the junior coach program.** Frequently cited issues included problems selecting the right students for the program, students missing their shift because of academic or behavioral issues and students forgetting to come to their recess shift. Most Playworks coaches also reported that at least a few junior coaches had to be removed from the program because of academic or behavioral concerns.

Playworks Implementation Varied by School Context. Our implementation site visits suggested that several contextual factors were associated with the implementation of Playworks:

- **Principals in schools at risk of failing to meet Adequate Yearly Progress student achievement targets (44 percent of treatment schools had not met these targets in the year prior to Playworks implementation) either saw Playworks as part of their overall strategy for improving achievement or were concerned that the program would take time away from academics.** Principals in the first group reported that they explicitly used Playworks as part of an overall strategy for improving school climate (e.g., student behavior and safety) because they felt the program would ultimately improve student achievement. Principals in the second group diminished the time allotted to Playworks and the importance given to recess, including limiting class game time and the availability of junior coaches at younger students' recesses.

- **A history of recess prior to Playworks implementation was associated with the quality of the program's implementation.** Two of the treatment schools did not have a history of recess prior to implementation of the Playworks program. Students in these schools were not always released for recess (or for the full recess period), so coaches were not able to implement the program consistently from day to day.

- **Principals' years of experience at the schools and coaches' experience with the program did not seem to be associated with implementation strength.** Although 38 percent of principals in treatment schools were new to their schools (Appendix 2, Table 12), schools with first-year principals were not observed to have a different quality of implementation. Prior coach experience also was not a factor, as coaches who were experienced with the program were no more likely to have a stronger implementation of Playworks.[5]

Most Teachers, Students and Principals Had Favorable Impressions of Playworks. Surveys and interviews conducted in treatment schools showed that teachers, students and principals generally had positive perceptions of the Playworks program.

- **Most teacher survey respondents in treatment schools reported a positive relationship and good communication with the Playworks coach.** Teachers felt their coaches were adequately trained (96 percent) and used appropriate techniques when working with students (94 percent) (Appendix 2, Table 6). Teachers also reported that Playworks was highly valued by students (95 percent), staff (86 percent) and, to a lesser extent, parents (59 percent) (Appendix 2, Table 10). In fact, 98 percent of interviewed teachers and 97 percent of surveyed teachers reported that they wanted Playworks in their school again the following year.

- **Teachers in treatment schools viewed Playworks as benefiting students in multiple ways, including providing their students with a positive recess experience.** Most teachers agreed that the program addressed important student needs (86 percent), reinforced positive behavior on the playground (96 percent) and helped students stay out of trouble (89 percent). They also felt that students had learned new games (97 percent) and recess rules (95 percent) (Appendix 2, Table 10).

- **Students in treatment schools reported being engaged with Playworks.** Data from the student survey found that in the two weeks prior to the survey, most students reported having participated in Playworks activities (73 percent), and the vast majority reported enjoying activities at recess (89 percent) and class game time (90 percent) (Appendix 2, Table 14).

- **Principals in all but one of the treatment schools reported that their schools needed Playworks again in the following year.** Principals were concerned about continued program funding and resulting sustainability. Playworks was funded mainly by schools and school districts. Just three schools (two treatment and one control) used—or planned to use, in the case of the control school—external grants to support the program. This speaks to schools' commitment to the program but also highlights the issue of long-term sustainability in an era of shrinking school budgets.

Summary of Findings

The current evaluation found positive impacts of the Playworks program on some measures of school climate, conflict resolution and aggression, and learning and academic performance, and showed no negative impacts of the program in any of the five domains that were assessed. In particular, our impact analyses showed the following:

- There was a significant positive impact of Playworks on teachers' reports of students using positive, encouraging language and teachers' perceptions of students' safety and engagement in inclusive behavior at recess, but no significant impact was found on three student-reported measures of school climate.
- Teachers in treatment schools were significantly more likely than teachers in control schools to report that school staff support organized play during the school day. However, treatment teachers were no more likely than control teachers to report their own support for organized play.
- Teachers in treatment schools reported less bullying and exclusionary behavior than teachers in control schools. No impacts were found, however, on teacher and student reports of aggressive behavior, students' beliefs about aggression or students' reports on their relationships with other students.
- Teachers in treatment schools were significantly less likely to report difficulties in transitioning to learning activities after recess and reported significantly less time to transition from recess to learning activities than teachers in control schools. Treatment students also were more likely than control students to report better behavior and attention in class after sports, games and play. We found no impacts of Playworks, however, on academic performance or student engagement with classroom activities.
- We found no significant impacts of Playworks on measures of youth development, such as students' feelings about interactions with adults or peers, and teachers' perceptions of students' abilities to regulate their emotions and engage in positive social behaviors.
- There were no significant impacts of Playworks on multiple indicators of student behavior, including school-level indicators based on administrative records, such as chronic absences and suspensions.

The implementation component of the evaluation provided additional insight into the school context at each study school, the degree to which each component of the program was carried out and student and staff perceptions of Playworks. In particular, our implementation findings suggested the following:

- Strong implementation occurred in 8 of 17 treatment schools and moderate implementation occurred in another 6 schools. Program strategies such as positive messaging (such as saying "good job, nice try"), promotion of inclusive behavior and conflict resolution strategies were modeled by coaches during the majority of recess periods observed at treatment schools. Most teachers reported that class game time was fun for students and provided them with good exercise and an opportunity to learn new games. Although coaches cited several challenges to implementing the junior coach program, junior coaches at all schools were scheduled to support at least one recess per week and had an opportunity to gain leadership skills by working with younger students at recess.
- Playworks implementation was stronger in schools that had recess in the past. Principal experience at the school and prior coach experience were not associated with the strength of implementation.
- Most teachers, students and principals had positive perceptions of the Playworks program. Teachers reported that recess was more organized and students were more engaged. Teachers also reported positive relationships with the coach, felt coaches were well prepared and believed that the program addressed important student needs, such as reinforcing positive behavior on the playground and helping students stay out of trouble. The majority of students reported enjoying recess and class game time activities, and principals in all but one treatment school reported that their schools needed Playworks again the following year.

The majority of students [in Playworks schools] reported enjoying recess and class game time activities…

References

Active Living Research. "Active Education: Physical Education, Physical Activity and Academic Performance." Research brief. San Diego, CA: Active Living Research, 2009. Available at *http://www.activelivingresearch.org/files/ALR_Brief_ActiveEducation_Summer2009.pdf*. Accessed December 1, 2011.

Ahamed Y, MacDonald H, Reed K, et al. "School-based Physical Activity Does Not Compromise Children's Academic Performance." *Medicine and Science in Sports and Exercise*, 39(2):371–376, 2007.

Barros RM, Silver EJ and Stein RE. "School Recess and Group Classroom Behavior." *Pediatrics*, 123(2):431–436, February 2009.

Bleeker M, James-Burdumy S, Beyler N, et al. "Findings from a Randomized Experiment of Playworks: Results from Cohort 1." Report submitted to the Robert Wood Johnson Foundation. Princeton, NJ: Mathematica Policy Research, April 2012.

Caterino M and Polak E. "Effects of Two Types of Activity on the Performance of Second-, Third-, and Fourth-Grade Students on a Test of Concentration." *Perceptual and Motor Skills,* 89(1), 245–248, August 1999.

Centers for Disease Control and Prevention. "The Association Between School-based Physical Activity, Including Physical Education, and Academic Performance." Atlanta, GA: U.S. Department of Health and Human Services, 2010.

Coe D, Pivarnik J, Womak C, et al. "Effect of Physical Education and Activity Levels on Academic Achievement in Children." *Medicine and Science in Sports and Exercise*, 38(8):1515–1519, August 2006.

Ericsson I. "Motor Skills, Attention and Academic Achievements: An Intervention Study in School Years 1-3." *British Educational Research Journal,* 34(3):301–313, 2008.

Evenson K, Ballard K, Lee G, et al. "Implementation of a School-based State Policy to Increase Physical Activity." *Journal of School Health,* 79(5):231–233, May 2009.

Ginsburg KR, the Committee on Communications and the Committee on Psychosocial Aspects of Child and Family Health. "The Importance of Play in Promoting Healthy Child Development and Maintaining Strong Parent-Child Bonds." *Pediatrics,* 119:182–191, 2007.

Hothorn T, Bretz F and Westfall P. "Simultaneous Inference in General Parametric Models." *Biometrical Journal,* 50(3):346–363, 2008.

Jarrett O, Maxwell D, Dickerson C, et al. "Impact of Recess on Classroom Behavior: Group Effects and Individual Differences." *The Journal of Educational Research,* 92(2):121–126, November 1998.

Leff SS, Costigan T and Power TJ. "Using Participator Research to Develop a Playground-based Prevention Program." *Journal of School Psychology,* 42:3–21, 2004.

London RA, Mallonee N, Stokes-Guinan K, et al. "Playworks Implementation in Eight Bay Area Elementary Schools: Final Report." Stanford, CA: John W. Gardner Center, 2010. Available at *http://www.rwjf.org/content/dam/farm/reports/evaluations/2010/rwjf69093*. Accessed December 14, 2010.

Maeda JK and Randall LM. "Can Academic Success Come from Five Minutes of Physical Activity?" *Brock Education,* 13(1):14–22, 2003.

Mahar MT, Murphy SK, Rowe DA, et al. "Effects of a Classroom-based Program on Physical Activity and On-Task Behavior." *Medicine and Science in Sports and Exercise,* 38(12):2086–2094, 2006.

Molloy GN. "Chemicals, Exercise and Hyperactivity: A Short Report." *International Journal of Disability, Development and Education,* 36(1):57–61, 1989.

Nelson M and Gordon-Larsen P. "Physical Activity and Sedentary Behavior Patterns Are Associated with Selected Adolescent Health Risk Behaviors." *Pediatrics,* 117(4):1281–1290, April 2006.

Pellegrini AD and Bohn CM. "The Role of Recess in Children's Cognitive Performance and School Adjustment." *Educational Researcher,* 34(1):13–19, 2005.

Pellegrini AD, Huberty PD and Jones I. "The Effects of Recess Timing on Children's Playground and Classroom Behaviors." *American Educational Research Journal,* 32(4):845–864, December 1995.

Pellegrini AD, Kato K, Blatchford P, et al. "A Short-Term Longitudinal Study of Children's Playground Games Across the First Year of School: Implications for Social Competence and Adjustment to School." *American Educational Research Journal,* 39(4):991–1015, 2002.

Robert Wood Johnson Foundation. "The State of Play: Gallup Survey of Principals on School Recess." Princeton, NJ: RWJF, 2010. Available at *http://www.playworks.org/files/StateOfPlayFeb2010.pdf*. Accessed November 18, 2011.

Shephard RJ. "Curricular Physical Activity and Academic Performance." *Pediatric Exercise Science,* 9(2):113–126, 1997.

Sibley BA and Etnier JL. "The Relationship Between Physical Activity and Cognition in Children: A Meta-Analysis." *Pediatric Exercise Science,* 15(3):243–256, 2003.

Taras H. "Physical Activity and Student Performance at School." *Journal of School Health,* 75(6):214–218, 2005.

Smith NJ and Lounsbery M. "Promoting Physical Education: The Link to Academic Achievement." *Journal of Physical Education, Recreation, and Dance,* 80(1):1–60, January 2009.

Tomporowski PD, Davis CL, Miller PH, et al. "Exercise and Children's Intelligence, Cognition, and Academic Achievement." *Educational Psychology Review,* 20(2):111–131, 2008.

Tremarche PV, Robinson EM and Graham LB. "Physical Education and Its Effect on Elementary Testing Results." *Physical Educator,* 64(2):58–64, 2007.

Trudeau F and Shephard RJ. "Relationships of Physical Activity to Brain Health and the Academic Performance of Schoolchildren." *American Journal of Lifestyle Medicine,* 4(2):138–150, 2010.

U.S. Department of Health and Human Services. "Physical Activity Guidelines Advisory Committee Report." Washington, DC: U.S. Department of Health and Human Services, 2008.

Zygmunt-Fillwalk E and Bilello TE. "Parents' Victory in Reclaiming Recess for Their Children." *Childhood Education,* 19–23, Fall 2005.

ENDNOTES

1. The measure of implementation strength used on this study reflects the number of program components in place and staff commitment to the program. Differences in implementation strength are described in the Implementation Findings section.
2. The impact of Playworks on students' play, physical activity and recess will be addressed in a separate brief.
3. Opportunities for physical activity and play and the recess environment in both treatment and control schools will be addressed in a separate brief.
4. We might also expect impacts to be larger in schools with stronger implementation. Our ability to measure statistically significant differences in impacts by implementation strength is limited because the study was not designed to measure differences by implementation strength. Nevertheless, we estimated the relationship between implementation strength and the magnitude of impacts for the eight outcomes with significant findings. Implementation strength was not significantly related to the magnitude of impacts for six of those outcomes. For two outcomes—teachers' reports of the use of positive, encouraging language and teachers' reports of students' safety/inclusion at recess (both from Table 3 of Appendix 2)—we found evidence suggesting that schools that had moderate or strong implementation experienced larger positive impacts of Playworks.
5. This finding differs from previously reported results (based only on cohort 1 schools) that showed coach tenure with Playworks was associated with implementation quality (Bleeker et al. 2012).

Appendix 1
Description of Study Design and Data Sources

Random Assignment Design

Random assignment of schools was used to help ensure that there were no systematic differences between the treatment and control groups, and so that the observed differences in outcomes between the two groups could be attributed solely to the effect of Playworks. The 29 schools from six cities that participated in the study were matched into blocks within each city prior to random assignment, with the goal of reducing the probability of chance differences between groups and improving the precision of the impact estimates. (In two cities, all study schools were in the same block.) Data from the U.S. Department of Education's Common Core of Data (CCD) from 2007–2008 were used to create the blocks. The CCD variables used included the highest grade in the school; school size (number of students); the percentage of black, Hispanic, and/or white students in the school; and the percentage of students eligible for free or reduced-price lunch. Three of the six cities had two blocks of matched schools, one had four blocks of matched schools and two had a single block of schools that required no matching. In total, there were 12 blocks of matched schools, one of which was a foursome, 3 of which were trios, and 8 of which were pairs. For a block of paired schools, one school was randomly assigned to the treatment group and one school to the control group. For blocks of 3 schools, 2 were randomly assigned to the treatment group and one to the control group. For the block of 4 schools, 3 were randomly assigned to the treatment group and one to the control group. Under this design, 17 schools were randomly assigned to the treatment group and 12 schools were randomly assigned to the control group.

Approach for Estimating Impacts

The impacts of Playworks on students, teachers (or classrooms) and schools were determined by comparing the average outcomes in treatment and control group schools using regression models that were customized to the unit of analysis (for example, school, teacher/classroom and student). For outcomes based on school-level data, we estimated the impact of Playworks with the following model:

$$Y_s = \alpha + \gamma T_s + \varepsilon_s,$$

where Y_s is the outcome for school s, α is a vector of indicator variables denoting the random assignment block in which the school was located, T_s indicates whether the school was assigned to the treatment group, ε_s is a school-level random error term, and γ is the parameter to be estimated from the model (γ represents the impact of Playworks on the school-level outcome).

For outcomes based on teacher-level (or classroom-level) data, we estimated the following model:

$$Y_{js} = \alpha + \beta X_{js} + \gamma T_s + \mu_s + \varepsilon_{js},$$

where Y_{js} is the outcome for classroom (or teacher) j in school s, α is a vector of indicator variables denoting the random assignment block in which the school was located, X_{js} is a vector of indicators for teacher race, T_s indicates whether the school in which the classroom (or teacher) was located was assigned to the treatment group, μ_s is a school-specific random error term, ε_{js} is a classroom-level (or teacher-level) random error term, and β and γ are parameters to be estimated. Indicators for teacher race were included as baseline characteristics (X) in the teacher-level (or classroom-level) impact models to account for the significant baseline differences in teacher race observed between the treatment and control groups.

For outcomes based on student-level data, we estimated the following model:

$$Y_{ijs} = \alpha + \gamma T_s + \mu_s + \varepsilon_{ijs},$$

where Y_{ijs} is the outcome for student i in classroom j in school s, α is a vector of indicator variables denoting the random assignment block in which the school was located, T_s indicates whether the school in which the student was enrolled was assigned to the treatment group, μ_s is a school-specific random error term, ε_{ijs} is a student-level random error term, and γ is the parameter to be estimated. Student baseline characteristics were not included in the student-level impact model because we did not observe significant baseline differences in student characteristics between the treatment and control groups (Appendix 2, Table 2).

Models for continuous outcome variables were estimated using least squares estimation, and models of binary outcome variables were estimated using logistic regression estimation. Standard errors for the estimated impacts on teacher- and student-level outcomes accounted for clustering at the school level using a generalized estimating equations approach. Outcomes were grouped into domains for the purpose of estimating impacts while accounting for multiple hypothesis testing (MHT). Each outcome was included in a single domain. We used our best judgment when grouping outcomes into domains, realizing that some outcomes may be appropriate for multiple domains. All statistically significant impacts discussed in the brief and presented in Appendix 2 are based on the MHT adjusted *p*-values. For the adjustments, we calculated statistical significance tests based on critical values from the multivariate t-distribution, taking into account correlations among the tests. Accounting for correlations among tests reduces the magnitude of the MHT adjustment, thereby increasing statistical power while still controlling the probability of finding a false impact (Hothorn, Bretz and Westfall 2008).

Sampling weights were used for estimating the impacts of teacher- and student-level outcomes to account for sampling of teachers and students within schools and attrition (nonresponse) occurring after sampling. The sampling weights were constructed so that teachers and students used in the impact analysis represented all eligible teachers and students, respectively, in the participating schools. That is, teachers and students were weighted so that larger schools were given more overall weight than smaller schools to account for the fact that the larger schools had more eligible teachers and students. In a sensitivity analysis, we confirmed that the impact estimates based on weighted teacher-

and student-level observations yielded similar results to the impact estimates based on unweighted observations, where teachers and students were all given equal weight across schools.

Data Sources

To address the study's primary research questions, we obtained data from both treatment and control schools from a variety of sources near the end of the school year (spring 2011 for cohort 1 schools and spring 2012 for cohort 2 schools). Data collection activities for the impact study included administration of student and teacher surveys and collection of administrative records. The implementation study included interviews with principals, teachers and Playworks coaches; focus groups with Playworks junior coaches; and observations of Playworks class game time and recess. The data collection activities that are the focus of this brief are described below.

- **Student Survey.** A total of 2,331 students from 119 4th- and 5th-grade classrooms in 28 study schools participated in a survey during the regular school day. A team of experienced survey administration staff from Mathematica conducted the 30-minute survey in each classroom. The survey captured information about students' perceptions of school climate, conflict resolution, learning and achievement, recess experience and relationships with adults and peers. In schools with five or fewer 4th- and 5th-grade classrooms, all 4th- and 5th-grade classrooms were asked to participate in the survey. In schools with more than five classrooms, we selected a random sample of five classrooms, balanced across the 4th and 5th grades. Students from one study school did not participate in the student survey because the school did not have any separate 4th- or 5th-grade classrooms (these students were combined with lower and higher grade level classrooms in the school); this school and the school it was matched with during random assignment were dropped from the student survey data analysis, leaving 27 schools. The response rate for the student survey was 83 percent (treatment schools: 82 percent; control schools: 83 percent).

- **Teacher Survey.** A total of 296 teachers from 29 study schools completed a 50-minute, self-administered, hard-copy instrument. The first half of the survey asked teachers to report on school climate (perceptions of safety, overall school environment and school support for organized play activities) as well as their perceptions of students' recess experience. The second half asked teachers about a random sample of five students in their classroom; teachers were asked to report on these students' behavior at school, learning and achievement and social competence. In schools with fewer than 15 teachers, all teachers were asked to participate in the survey. In schools with more than 15 teachers, we selected a random sample of 15 teachers to complete the survey, balanced across grade levels (grades one through five). The response rate for the teacher survey was 83 percent (treatment schools: 85 percent; control schools: 80 percent).

- **Teacher Survey Addendum.** Separately, we asked treatment group teachers responding to the teacher survey to provide information on changes (due to Playworks) in their time spent supervising recess, leading physical education, participating in professional development activities or training related to Playworks, addressing behavioral or disciplinary issues and engaging in other tasks. Teachers were not asked about the value of their time, only the increase or decrease in the number of hours

they spent on these activities due to Playworks. The teacher survey addendum was administered to 175 teachers in the treatment group. We used these data to construct estimates of the average change in time spent by category.

- **Administrative Records.** All 29 study schools provided a list of teachers to Mathematica. Schools then provided student rosters for each classroom that was selected for participation in the study. After the study year, schools were also asked to provide information about demographic characteristics, attendance, disciplinary events and academic performance. All 29 schools provided this information, though some items are missing for some schools.

- **Interviews with Principals.** JGC staff interviewed one principal from each of the 29 study schools during the school day for about 60 to 90 minutes each. Assistant principals were also interviewed in 3 schools. Interviews at both treatment and control schools were designed to collect information about non-Playworks opportunities for play and physical activity; reasons for wanting to bring Playworks to the school; typical recess experiences of students and teachers; school context and student population; and the principals' views of play. At treatment schools, interviews also included questions about issues such as Playworks rollout at the school; integration of the Playworks coach into the school; views of the Playworks model and its effects on recess, physical activity, discipline, class behavior and learning; challenges faced; and costs of implementing Playworks. Principals at all study schools were also asked to report on discipline referrals to the principal's office that occurred over the course of the week prior to the interview. One school did not provide discipline referral data; this school was part of a trio of schools for randomization, so all 3 schools were dropped from the discipline referral data analysis, leaving 26 schools.

- **Interviews with Teachers.** JGC staff interviewed a total of 51 teachers from treatment schools for about 30 minutes each. We sampled one teacher from grade five, one teacher from grade three or four and one teacher from grade one or two in each study school.[1] Teacher interviews focused on topics such as the typical recess experiences for students and teachers; Playworks rollout at the school, including individual components; staff training and experiences; relationships with the Playworks coach; views of the Playworks model and its effects on students; and challenges faced.

- **Interviews with Playworks Coaches.** JGC staff interviewed the Playworks coach in each of the 17 treatment schools for about 60 minutes each. Interview topics included reasons for working with Playworks; previous experience and training; Playworks rollout at the school (including individual components); relationships with principals and teachers and integration of the Playworks coach into the school; views of the Playworks model and its effects on recess, physical activity and students; and challenges faced.

- **Focus Groups with Junior Coaches.** JGC staff conducted focus groups with students who were junior coaches at 16 treatment schools. Focus groups took place after school in a secure room without Playworks staff present and lasted about 90 minutes each. Students were asked to describe reasons for wanting to become a junior coach; the training they received; experiences as a junior coach; other students' perceptions of Playworks; and challenges faced.

1 One school did not include 5th grade; in that school, JGC staff interviewed a 1st grade teacher, a 2nd grade teacher, and a 3rd-4th grade combination class teacher.

- **Playworks Observations.** JGC staff conducted recess observations in all 17 treatment schools to assess Playworks coach involvement and strategies, student participation in Playworks games, students' use of Playworks strategies and language, playground monitor and teacher activities and junior coach participation. Staff also observed class game time in order to assess the coaches' relationships with students in smaller groups and examine teacher and coach interactions and discipline styles.

Additional Study Briefs

Two separate study briefs will report findings related to Playworks implementation and the impact of the program on play, physical activity and recess outcomes. The implementation brief will describe Playworks implementation in the 17 treatment schools, focusing on the ways that Playworks supports recess quality and the contextual factors that influence recess functioning and program implementation quality. The brief focuses specifically on Playworks recess, junior coaches and class game time and how these components support the overall program goals.

The brief focused on play, physical activity and recess outcomes will draw on data from the following data collection activities:

- **Recess Observations**. A team of trained observers from Mathematica measured students' physical activity and active participation in organized games during six recess periods at each school using a structured observation protocol. Mathematica staff also measured negative interactions (such as teasing, verbal abuse or aggression) and positive interactions (such as supportive language or use of conflict resolution skills) among students. JGC staff also conducted recess observations in treatment and control schools; during these observations, JGC assessed organization of recess, engagement in games and play, conflicts and resolution, inclusiveness and physical activity.
- **Objective Physical Activity Data from Accelerometers.** Samples of 4th- and 5th-grade students in each school were asked to wear accelerometers for one full school day to measure their physical activity. Accelerometers are movement monitors similar to pedometers. They are recognized as one of the most effective ways to record frequency, intensity and duration of physical activity with minimal burden on participants.
- **Physical Activity and Recess Data from Student and Teacher Surveys.** The student survey described in this brief included a section that asked students to report on their extent and enjoyment of physical activity, confidence in physical skills and capabilities and physical activity outside of school. The teacher and student surveys also measured recess activities, students' perceptions of recess, conflict and behavior during recess and teacher perceptions of students' feelings about recess.

Appendix 2
Tables

TABLE 1.
Characteristics of Schools in the Study

Characteristic (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Percentage of Schools Receiving Title I	28 (CCD)				
Title I-eligible school		92.9	92.9	0.0	1.00
Schoolwide Title I		86.7	84.2	2.5	1.00
Percentage of Schools in the Following Areas:	28 (CCD)				
Urban		100.0	100.0	0.0	1.00
Suburban		0.0	0.0	0.0	1.00
Town		0.0	0.0	0.0	1.00
Rural		0.0	0.0	0.0	1.00
Number of Students Per Teacher	28 (CCD)	16.3	16.3	0.0	1.00
Number of Students Per School	28 (CCD)	494.0	562.3	-68.3	0.97
Percentage of Students Eligible for Free or Reduced-Price Lunch	28 (CCD)	81.0	83.1	-2.1	1.00
Percentage of Students that Are the Following Race/Ethnicity:[a]	28 (CCD)				
Black		40.7	38.3	2.4	1.00
Hispanic		25.6	32.3	-6.7	0.99
White		17.0	12.9	4.1	0.92
Asian		14.2	7.7	6.5	0.62
Native American		0.8	0.9	-0.1	0.99

Sources: Common Core of Data (CCD) from the 2009-2010 school year (25 schools) and 2010-2011 school year (3 schools).

Note: Random assignment block indicators were included as covariates in all difference models. The *p*-values reported in this table account for clustering of students and teachers within schools and for multiple hypothesis testing (MHT) to control the probability of finding any falsely significant impacts (the family-wise error rate) at 5 percent. The adjustment for MHT is based on the multivariate t-distribution and takes into account correlations among test statistics. The adjustment accounts for the tests presented in this table, but not for tests presented in other tables. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding. CCD information was not available for one cohort 2 school that was new in 2011–2012.

a These percentages do not necessarily sum to 100 because they are calculated by averaging school-level percentages.

TABLE 2.
Characteristics of Students and Teachers in the Study

Characteristic (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Student Characteristics					
Percentage of Students that Are Female	2,305 (student survey)	52.5	50.5	2.0	0.75
Percentage of Students that Are in the Following Grades:	2,305 (administrative records)				
4th		53.3	52.2	1.1	1.00
5th		46.7	47.8	-1.1	1.00
Percentage of Students that Are the Following Race/Ethnicity:[a]	2,254 (student survey)				
Black or African American		31.8	30.4	1.3	1.00
Hispanic		33.2	47.3	-14.1	0.70
White		27.1	22.0	5.1	0.62
Asian, Native Hawaiian, or Other Pacific Islander		23.6	13.0	10.6	0.65
American Indian or Alaskan Native		9.0	6.3	2.7	0.70
Teacher Characteristics					
Percentage of Teachers that Are Female	297 (teacher survey)	90.1	88.0	2.1	1.00
Percentage of Teachers that Are Hispanic or Latino	293 (teacher survey)	8.4	14.6	-6.2	0.61
Percentage of Teachers that Are the Following Race:[b]	273 (teacher survey)				
White		86.2	73.1	13.0**	0.02
African American		8.1	18.6	-10.5**	0.02
Other race[c]		6.3	10.2	-3.9	0.91
Percentage of Teachers with the Following Highest Level of Education:	294 (teacher survey)				
Bachelor's degree		35.8	40.6	-4.7	1.00
Master's degree		54.9	53.6	1.4	1.00
Other degree		9.3	5.8	3.4	0.95
Number of Years Teaching Experience	291 (teacher survey)	11.7	11.5	0.2	1.00
Number of Years Teaching at the Current School	291 (teacher survey)	5.6	6.2	-0.6	1.00

Sources: Student and teacher surveys conducted in spring 2011 and spring 2012 and administrative records data collected from schools.

Note: Random assignment block indicators were included as covariates in all difference models. The *p*-values reported in this table account for clustering of students and teachers within schools and for multiple hypothesis testing (MHT) to control the probability of finding any falsely significant impacts (the family-wise error rate) at 5 percent. The adjustment for MHT is based on the multivariate t-distribution and takes into account correlations among test statistics. The adjustment accounts for the tests presented in this table, but not for tests presented in other tables. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding.

a These percentages can sum to more than 100 because students could report more than one race or ethnicity.

b These percentages can sum to more than 100 because teachers could report more than one race.

c This includes Asian, Native Hawaiian, Other Pacific Islander, American Indian, and Alaskan Native.

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

*** Significantly different from zero at the .01 level, two-tailed test.

TABLE 3.
Impacts on School Climate

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
School as Community					
Sense of School as Community Scale Score[a]	2,281 (student survey)	2.8	2.7	0.1	0.62
Percentage of Teachers that "Agree" or "Strongly Agree" that Students in Their School Use Positive, Encouraging Language	291 (teacher survey)	53.9	27.4	26.5**	0.03
Feelings of Safety					
Student-Reported Feelings of Safety at School Scale Score[b]	2,283 (student survey)	2.6	2.5	0.1	0.57
Student-Reported Feelings of Safety at Recess Scale Score[c]	2,282 (student survey)	2.8	2.6	0.1	0.23
Teacher-Reported Feelings of Students' Safety at School Scale Score[d]	295 (teacher survey)	3.8	3.2	0.6**	0.05
Teacher-Reported Feelings of Students' Safety/Inclusion at Recess Scale Score[e]	294 (teacher survey)	3.9	3.1	0.8***	0.00
Support for Organized Play					
Teacher Support for Organized Play During the School Day Scale Score[f]	287 (teacher survey)	4.2	4.0	0.1	0.41
School Staff Support for Organized Play During the School Day Scale Score[g]	296 (teacher survey)	4.3	3.9	0.4**	0.02

Sources: Student and teacher surveys conducted in spring 2011 and spring 2012.

Note: Random assignment block indicators were included as covariates in all impact models; indicators for teachers' race were included as covariates in the impact models for teacher survey outcomes. The *p*-values reported in this table account for clustering of students and teachers within schools and for multiple hypothesis testing (MHT) to control the probability of finding any falsely significant impacts (the family-wise error rate) at 5 percent. The adjustment for MHT is based on the multivariate t-distribution and takes into account correlations among test statistics. The adjustment accounts for the tests presented in this table, but not for tests presented in other tables. Sample sizes based on the same data source might be different due to missing responses. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding.

a The Sense of School as Community Scale averages student responses to 13 items from the student survey: (A1) "Students at this school really care about each other."; (A2) "Students at this school are willing to go out of their way to help someone."; (A3) "When I'm having a problem, some other student will help me."; (A4) "Teachers and students treat each other with respect in this school."; (A5) "People care about each other in this school."; (A6) "Students at this school work together to solve problems."; (A7) "Students in this school don't seem to like each other very well."; (A8) "Students in this school are just looking out for themselves."; (A9) "Students in this school treat each other with respect."; (A10) "The students in this school don't really care about each other."; (A11) "I feel that I can talk to the teachers in this school about things that are bothering me."; (A12) "Teachers and students in this school don't seem to like each other."; and (A13) "Students in this school help each other, even if they are not friends." Responses are coded on a 4-point scale ranging from 1 (agree a lot) to 4 (disagree a lot) for A7, A8, A10, and A12 and from 1 (disagree a lot) to 4 (agree a lot) for all other items. Higher values on the scale indicate more positive student feelings about their sense of school as community. The scale is coded as missing if responses were missing for four or more items.

b The Student-Reported Feelings of Safety at School Scale averages student responses to four items from the student survey: (A14) "Students feel afraid that someone will bully them at school."; (A15) "Students feel afraid that someone will hurt them at school."; (A16) "Students feel afraid that someone will tease them at school."; and (A17) "Students feel safe at this school." Responses are coded on a 4-point scale ranging from 1 (disagree a lot) to 4 (agree a lot) for A17 and from 1 (agree a lot) to 4 (disagree a lot) for the other three items. Higher values on the scale indicate more positive student feelings about safety at school. The scale is coded as missing if responses were missing for two or more items.

c The Student-Reported Feelings of Safety at Recess Scale averages student responses to four items from the student survey: (A18) “Students feel afraid that someone will bully them at recess.”; (A19) “Students feel afraid that someone will hurt them at recess.”; (A20) “Students feel afraid that someone will tease them at recess.”; and (A21) “Students feel safe during recess.” Responses are coded on a 4-point scale ranging from 1 (disagree a lot) to 4 (agree a lot) for A21 and from 1 (agree a lot) to 4 (disagree a lot) for the other three items. Higher values on the scale indicate more positive student feelings about safety at recess. The scale is coded as missing if responses were missing for two or more items.

d The Teacher-Reported Feelings of Students’ Safety at School Scale averages teacher responses to two items from the teacher survey: (B5) “Students feel safe at this school.” and (B6) “Students feel afraid that someone will hurt them at school.” Responses are coded on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree) for B5 and from 1 (strongly agree) to 5 (strongly disagree) for B6. Higher values on the scale indicate more positive teacher feelings about safety at school. The scale is coded as missing if responses were missing for either item.

e The Teacher-Reported Feelings of Students’ Safety/Inclusion at Recess Scale averages teacher responses to three items from the teacher survey: (B7) “Students feel safe at recess.”; (B8) “Students feel included at recess.”; and (B9) “Students feel afraid that someone will hurt them at recess.” Responses are coded on a 5-point scale ranging from 1 (strongly agree) to 5 (strongly disagree) for B9 and from 1 (strongly disagree) to 5 (strongly agree) for B7 and B8. Higher values on the scale indicate more positive teacher feelings about safety/inclusion at recess. The scale is coded as missing if responses were missing for one or more items.

f The Teacher Support for Organized Play During the School Day Scale averages teacher responses to eight items from the teacher survey: (A19) “The transition back to class after recess is shortened if students have participated in organized play/activities during recess.”; (A20) “Conflict in the classroom is reduced if students have participated in organized play/activities during recess.”; (A21) “Students are more likely to feel included if they participate in organized play/activities during recess.”; (A22) “Participating in organized play/activities during recess helps increase students’ physical activity levels.”; (A23) “When there are organized play/activities during recess, kids are less likely to get involved in arguments or fights.”; (A24) “Scheduling physical activity programs during the school day takes away important time that my students need to focus on their academic achievement.”; (A25) “Participating in play/activities organized by adults during recess takes away important time that children have for unstructured play.”; and (B3) “It is important for students to have the opportunity to play during the school day.” Responses are coded on a 5-point scale ranging from 1 (strongly agree) to 5 (strongly disagree) for A24 and A25 and from 1 (strongly disagree) to 5 (strongly agree) for all other items. Higher values on the scale indicate higher levels of teacher support for organized play activities during the school day. The scale is coded as missing if responses were missing for three or more items.

g The School Staff Support for Organized Play During the School Day Scale averages teacher responses to two items from the teacher survey: (B1) “Staff at our school think it is important to provide students with an opportunity to play during the school day.” and (B2) “Staff at our school think it is important to have recess for students.” Responses are coded on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree) for both items. Higher values on the scale indicate higher levels of staff support for organized play during the school day. The scale is coded as missing if responses were missing for either item.

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

*** Significantly different from zero at the .01 level, two-tailed test.

TABLE 4.
Implementation of Key Program Components in Playworks Schools

	Sample Size (data source)	
Recess Implementation		
Percentage of Recesses in Which the Following Was Observed:	98 (recess observations)	
Use of positive messaging by the coach		67.3
Use of positive messaging by other adults		31.6
Use of inclusive behavior by the coach		57.1
Use of inclusive behavior by other adults		19.4
Coach played with students		67.3
Other adults played with students		43.9
Use of ro-sham-bo by students		32.7
Class Game Time Implementation		
Percentage of Class Game Times in Which the Following Was Observed:	61 (class game time observations)	
Use of positive messaging by the coach		91.8
Use of ro-sham-bo by students		34.4
Teacher played with students		39.3
Use of positive messaging by the teacher		29.5
Coach addressed negative student behavior		50.8
Teacher addressed negative student behavior		50.8
Junior Coach Program		
Percentage of Junior Coaches that Reported Using Ro-Sham-Bo at Recess	16 (junior coach focus groups)	87.5
Percentage of Recesses in Which Junior Coaches Were Observed Intervening in Conflict	98 (recess observations)	21.4
Percentage of Schools with 19 or Fewer Junior Coaches	17 (Playworks coach interviews)	52.9
Percentage of Schools with 20 or More Junior Coaches	17 (Playworks coach interviews)	47.1

Sources: Recess observations, class game time observations, focus groups, and interviews with Playworks coaches conducted in treatment schools in spring 2011 and spring 2012.

TABLE 5.
Impacts on Conflict Resolution and Aggression

Outcome (mean)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Interactions with Other Students					
Relationships with Other Students Scale Score[a]	2,269 (student survey)	3.1	3.1	0.1	0.81
Student Bullying/Exclusion Scale Score[b]	295 (teacher survey)	0.6	1.0	-0.4**	0.01
Aggression					
Aggressive Behavior Scale Score[c]	2,288 (student survey)	1.4	1.5	-0.1	0.55
Normative Beliefs About Aggression Scale Score[d]	2,285 (student survey)	1.6	1.7	-0.1	0.41
BASC Aggression Subscale Score[e]	1,446 (teacher survey, student-specific portion)	6.0	6.6	-0.6	0.86

Sources: Student and teacher surveys conducted in spring 2011 and spring 2012.

Note: Random assignment block indicators were included as covariates in all impact models; indicators for teachers' race were included as covariates in the impact models for teacher survey outcomes. The *p*-values reported in this table account for clustering of students and teachers within schools and for multiple hypothesis testing (MHT) to control the probability of finding any falsely significant impacts (the family-wise error rate) at 5 percent. The adjustment for MHT is based on the multivariate t-distribution and takes into account correlations among test statistics. The adjustment accounts for the tests presented in this table, but not for tests presented in other tables. Sample sizes based on the same data source might be different due to missing responses. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding.

a The Relationships with Other Students Scale averages student responses to three items from the student survey: (B7) "I get along well with other kids during recess."; (B8) "If a disagreement with another kid happens at school, I know how to work things out without getting into a fight."; and (B9) "If a disagreement with another kid happens during recess, I know how to work things out without getting into a fight." Responses are coded on a 4-point scale ranging from 1 (disagree a lot) to 4 (agree a lot) for all items. Higher values on the scale indicate more positive views of relationships with other students. The scale is coded as missing if responses were missing for one or more items.

b The Student Bullying/Exclusion Scale averages teacher responses to seven items from the teacher survey: (B10) "In the past 30 days, how often have students reported to you that they have been a victim of name-calling during recess?"; (B11) "In the past 30 days, how often have students reported to you that they have been hit or pushed by another student during recess?"; (B12) "In the past 30 days, how often have students reported that they have been isolated from their normal peer group during recess?"; (B13) "In the past 30 days, how often have students reported to you that they have been bossed or coerced to do something they didn't want to do during recess?"; (B14) "In the past 30 days, how often have students indicated that they are afraid to come to school because of the fear of being bullied?"; (B15) "In the past 30 days, how often have students indicated that they are afraid to go to recess because of the fear of being bullied?"; and (B16) "In the past 30 days, how often has a parent indicated that their child is afraid to come to school because of the fear of being bullied?" Responses are coded on a 4-point scale ranging from 0 (never) to 3 (5 or more times) for all items. Higher values on the scale indicate higher levels of bullying/exclusion. The scale is coded as missing if responses were missing for two or more items.

c The Aggressive Behavior Scale averages student responses to six items from the student survey: (B1) "In the past two weeks I teased a kid at school."; (B2) "In the past two weeks I pushed, shoved, or hit a kid at school."; (B3) "In the past two weeks I called a kid at school a bad name."; (B4) "In the past two weeks I said that I would hit a kid at school."; (B5) "In the past two weeks I left out another kid on purpose."; and (B6) "In the past two weeks I made up something about other students to make other kids not like them anymore." Responses are coded on a 4-point scale ranging from 1 (never) to 4 (many times) for all items. Higher values on the scale indicate higher frequency of aggressive behavior. The scale is coded as missing if responses were missing for two or more items.

d The Normative Beliefs About Aggression Scale averages student responses to eight items from the student survey: (B10) "It is OK to take your anger out on others by using physical force."; (B11) "If you're angry, it is OK to say mean things to other people."; (B12) "It is OK to yell at others and say bad things."; (B13) "It is OK to punch or shove other people around if you're mad."; (B14) "It is wrong to insult (that is put down or make fun of) other people."; (B15) "It is wrong to take it out on others by saying mean things when you're mad."; (B16) "It is wrong to get into physical fights with other people."; and (B17) "It is wrong to hit other people." Response are coded

on a 4-point scale ranging from 1 (disagree a lot) to 4 (agree a lot) for B10 to B13 and from 1 (agree a lot) to 4 (disagree a lot) for B14 to B17. Higher values on the scale indicate higher general approval of aggression. The scale is coded as missing if responses were missing for three or more items.

e The BASC Aggression Subscale is created by summing teacher responses to 14 items from the student-specific portion of the teacher survey. Higher values on the scale indicate higher levels of aggression in the student. The scale is coded as missing if responses were missing for three or more items.

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

*** Significantly different from zero at the .01 level, two-tailed test.

BASC = Behavioral Assessment System for Children.

TABLE 6.
Teacher Perceptions of Junior Coaches and Playworks Coaches

	Sample Size (data source)	Treatment
Junior Coaches		
Percentage of Teachers Reporting Having One or More Playworks Junior Coaches in Their Class	170 (teacher survey)	29.8
Percentage of Teachers that "Agree" or "Strongly Agree" that Junior Coaches:	135 (teacher survey)	
Enjoy their role at recess		87.9
Have gained leadership skills through their participation		89.0
Have improved their own recess conduct		69.6
Teach other students games		81.1
Help resolve conflicts at recess		67.4
Have reduced their own incidents of conflict with others		66.0
Include others at recess		78.1
Are good role models		75.7
Are eager to be junior coaches		82.1
Playworks Coaches		
Percentage of Teachers that "Agree" or "Strongly Agree" that the Playworks Coach at Their School:	175 (teacher survey)	
Is adequately trained		95.8
Gets along well with students		99.0
Is successful at including all students in organized activities		93.6
Gets along well with school staff		93.7
Communicates well with teachers		88.4
Communicates well with students		94.2
Uses appropriate techniques when working with students		93.6
Percentage of Teachers that "Agree" or "Strongly Agree" that Students Feel Connected to the Playworks Coach	175 (teacher survey)	91.2
Percentage of Teachers that "Agree" or "Strongly Agree" that They Have a Positive Relationship with the Playworks Coach	175 (teacher survey)	89.7

Source: Teacher surveys conducted in spring 2011 and spring 2012.

Note: Sample sizes based on the same data source might be different due to missing responses.

TABLE 7.
Impacts on Learning and Academic Performance

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Transition from Recess to Classroom Activities					
Effect of Recess on Behavior in Class Scale Score[a]	2,247 (student survey)	2.5	2.4	0.1	0.86
Effect of Sports, Games and Play on Behavior in Class Scale Score[b]	2,241 (student survey)	2.7	2.6	0.2**	0.03
Percentage of Students that Report that It Is "Somewhat True" or "Very True" that Beginning Class Work After Recess Is Easy	2,279 (student survey)	59.3	54.2	5.1	0.87
Number of Minutes to Transition from Recess to Learning Activities[c]	286 (teacher survey)	6.8	10.3	-3.5*	0.10
Difficult Transition to Learning After Recess Scale Score[d]	293 (teacher survey)	2.4	3.1	-0.7***	0.00
Engagement with Classroom Activities					
Engagement Versus Disaffection with Learning Scale Score[e]	2,279 (student survey)	3.3	3.2	0.1	0.28
Percentage of Teachers that "Agree" or "Strongly Agree" that Their Students:[f]	1,445 (teacher survey, student-specific portion)				
Are attentive in class		59.3	56.7	2.6	1.00
Participate in class		77.8	69.8	7.9	0.47
Academic Performance					
Percentage of Teachers that Report that Their Students "Often" or "Always or Almost Always" Complete Their Homework[f]	1,460 (teacher survey, student-specific portion)	77.5	75.7	1.9	1.00
Percentage of Teachers that Report that Their Students' Academic Performance Is "Somewhat" or "Far" Above Grade Level[f]	1,459 (teacher survey, student-specific portion)	8.8	13.3	-4.5	0.97
Percentage of Teachers that Report that Their Students' Motivation to Succeed Academically Is "High" or "Extremely High"[f]	1,459 (teacher survey, student-specific portion)	34.5	30.7	3.8	1.00
Percentage of 3rd Grade Students Proficient/Advanced in:[g]	29 (administrative records)				
Reading		57.2	58.5	-1.3	1.00
Math		61.8	62.9	-1.0	1.00

TABLE 7. (continued)
Impacts on Learning and Academic Performance

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Percentage of 4th Grade Students Proficient/Advanced in:[g]	29 (administrative records)				
Reading		63.7	54.3	9.3	0.73
Math		64.2	59.9	4.3	1.00
Percentage of 5th Grade Students Proficient/Advanced in:[g]	28 (administrative records)				
Reading		57.3	55.0	2.3	1.00
Math		52.2	55.1	-2.9	1.00

Sources: Student and teacher surveys conducted in spring 2011 and spring 2012 and administrative records.

Note: Random assignment block indicators were included as covariates in all impact models; indicators for teachers' race were included as covariates in the impact models for teacher survey outcomes. The *p*-values reported in this table account for clustering of students and teachers within schools and for multiple hypothesis testing (MHT) to control the probability of finding any falsely significant impacts (the family-wise error rate) at 5 percent. The adjustment for MHT is based on the multivariate t-distribution and takes into account correlations among test statistics. The adjustment accounts for the tests presented in this table, but not for tests presented in other tables. Sample sizes based on the same data source might be different due to missing responses. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding.

a The Effect of Recess on Behavior in Class Scale averages student responses to three items from the student survey: (C1) "Problems that happened during recess make it hard for my teacher to start lessons after recess."; (C3) "It is hard to get settled down after recess."; and (C7) "I sometimes make it hard for my teacher to start lessons after recess." Responses are coded on a 4-point scale ranging from 1 (very true) to 4 (not at all true) for all items. Higher values on the scale indicate more positive effects from recess on class. The scale is coded as missing if responses were missing for one or more items.

b The Effect of Sports, Games and Play on Behavior in Class Scale averages student responses to three items from the student survey: (C4) "It is easier to pay attention in class on days when I play and run around than on days when I don't play and run around."; (C5) "Participating in sports and games at recess helps me pay attention in class."; and (C6) "Participating in sports and games at recess helps me behave better in class." Responses are coded on a 4-point scale ranging from 1 (not at all true) to 4 (very true) for all items. Higher values on the scale indicate more positive effects from playing sports on class. The scale is coded as missing if responses were missing for one or more items.

c This outcome averages teacher responses to the following question from the teacher survey: (A15) "On the most recent school day in which your students participated in recess, approximately how many minutes did it take for the majority of students to become engaged in their first classroom activity after recess?"

d The Difficult Transition to Learning After Recess Scale averages teacher responses to three items from the teacher survey: (A16) "Think about the most recent school day in which your students participated in recess. In the 15 minutes just after recess, some students became restless and began to lose focus on their tasks."; (A17) "Think about the most recent school day in which your students participated in recess. In the 15 minutes just after recess, there were incidents of negative student behavior toward peers or the teacher (e.g., teasing, name-calling, aggression, or exclusionary behavior)."; and (A18) "Think about the most recent school day in which your students participated in recess. In the 15 minutes just after recess, I spent more time than I would have liked redirecting student misbehavior." Responses are coded on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher values on the scale indicate a rougher transition to learning after recess. The scale is coded as missing if responses were missing for one or more items.

e The Engagement Versus Disaffection with Learning Scale averages student responses to 10 items from the student survey: (C13) "I pay attention in class."; (C14) "When I'm in class, I join in on class discussions."; (C15) "I try hard to do well in school."; (C16) "In class, I work as hard as I can."; (C17) "When I am in class, I listen very carefully."; (C18) "When I'm in class, I just pretend like I'm working."; (C19) "I don't try very hard at school."; (C20) "In class, I only work as much as I have to so I don't get in trouble."; (C21) "When I'm in class, I think about other things."; and (C22) "When I'm in class, my mind wanders." Responses are coded on a 4-point scale ranging from 1 (disagree a lot) to 4 (agree a lot) for questions C13, C14, C15, C16, and C17 and from 1 (agree a lot) to 4 (disagree a lot) for all other items. Higher values on the scale indicate higher levels of engagement with learning. The scale is coded as missing if responses were missing for two or more items from C13 to C17 and/or two or more items from C18 to C22.

f Teachers reported on up to five students. For each teacher, we constructed a weighted average of his or her students' outcomes. Impacts are estimated on the teacher-level weighted average. The reported sample size is the number of students about whom teachers reported.

g Percentage proficient/advanced is the percentage of students meeting or exceeding proficiency standards on state achievement tests based on school administrative records. Within random assignment block, all schools administered the same test and applied the same proficiency standard. Only 28 schools are included in the 5th grade tests score analysis because one school did not have 5th grade students during the study year; however, there were both treatment and control students in the random assignment block, so the rest of the block is included.

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

*** Significantly different from zero at the .01 level, two-tailed test.

TABLE 8.
Playworks' Role in Class Environment

	Sample Size (data source)	
Percentage of Teachers that Reported that Playworks Had the Following Effects on Students in Their Classrooms:	51 (teacher interviews)	
Students used ro-sham-bo		37.3
Students showed an increase in teamwork/inclusiveness, compared with previous experience		13.7
Students were more ready to learn, compared with previous experience		29.4
Students participated more in class, compared with previous experience		5.9
Students behaved better, compared with previous experience		5.9
Junior coaches' academics were negatively affected by missing class time		3.9
Percentage of Teachers Who Reported that Playworks Had the Following Effects on Them as Teachers:	51 (teacher interviews)	
Used games learned during class game time in classroom		21.6
Used Playworks' facilitation and management strategies		17.6
Spent less time dealing with conflict in classroom		15.7

Source: Interviews conducted with teachers in treatment schools in spring 2011 and spring 2012.

TABLE 9.
Impacts on Youth Development

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Interactions with Peers and Adults					
Student Interactions with Adults at School Scale Score[a]	2,280 (student survey)	3.3	3.2	0.1	0.85
Percentage of Students that Agree "A Little" or "A Lot" that They Get Along Well with Other Students	2,280 (student survey)	83.5	82.0	1.4	0.99
Self-Efficacy for Peer Interaction – Conflict Subscale Score[b]	2,275 (student survey)	2.1	2.0	0.0	0.96
Self-Efficacy for Peer Interaction – Non-Conflict Subscale Score[c]	2,267 (student survey)	1.7	1.7	0.0	1.00
Social Competence					
Social Competence – Emotional Regulation Subscale Score[d]	1,450 (teacher survey, student-specific portion)	2.1	2.0	0.1	0.54
Social Competence – Prosocial Behavior Subscale Score[e]	1,449 (teacher survey, student-specific portion)	2.1	2.0	0.1	0.90
Responsibility and Altruism					
Responsibility Scale Score[f]	1,436 (teacher survey, student-specific portion)	2.0	1.9	0.1	0.56
Altruism Scale Score[g]	1,419 (teacher survey, student-specific portion)	1.3	1.3	0.0	1.00

Sources: Student and teacher surveys conducted in spring 2011 and spring 2012.

Note: Random assignment block indicators were included as covariates in all impact models; indicators for teacher's race were included as covariates in the impact models for teacher survey outcomes. The *p*-values reported in this table account for clustering of students and teachers within schools and for multiple hypothesis testing (MHT) to control the probability of finding any falsely significant impacts (the family-wise error rate) at 5 percent. The adjustment for MHT is based on the multivariate t-distribution and takes into account correlations among test statistics. The adjustment accounts for the tests presented in this table, but not for tests presented in other tables. Sample sizes based on the same data source might be different due to missing responses. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding.

a The Student Interactions with Adults at School Scale averages student responses to six items from the student survey: (F1) "At my school, there is an adult who really cares about me."; (F2) "At my school, there is an adult who tells me when I do a good job."; (F3) "At my school, there is an adult who always wants me to do my best."; (F4) "At my school, there is an adult who listens to me when I have something to say."; (F5) "During recess, there is an adult who often plays or participates in activities with us."; and (F6) "During recess, there is an adult who would be happy to help me if I had a problem." Responses are coded on a 4-point scale ranging from 1 (disagree a lot) to 4 (agree a lot) for all items. Higher values on the scale indicate more positive student feelings about their interactions with adults. The scale is coded as missing if responses were missing for two or more items.

b The Self-Efficacy for Peer Interaction – Conflict Subscale averages student responses to eight items from the student survey: (G2) "Some kids are teasing your friend. How easy or hard would it be for you to tell them to stop?"; (G3) "A kid cuts in front of you in line. How easy or hard would it be for you to tell the kid not to cut in front of you?"; (G4) "A kid wants to do something that will get you into trouble. How easy or hard would it be for you to ask the kid to do something else?"; (G5) "Some kids are making fun of someone in your classroom. How easy or hard would it be for you to tell them to stop?"; (G6) "A kid wants to be first when you play a game. How easy or

hard would it be for you to tell the kid that you are going first?"; (G7) "A kid does not like your friend. How easy or hard would it be for you to tell the kid to be nice to your friend?"; (G8) "A group of kids wants to play a game that you don't like. How easy or hard would it be for you to ask them to play a game that you like?"; and (G9) "A kid is yelling at you. How easy or hard would it be for you to tell the kid to stop?" Responses are coded on a 4-point scale ranging from 1 (really hard) to 4 (really easy) for all items. Higher values on the scale indicate more effective interacting with peers in conflict situations. The scale is coded as missing if responses were missing for three or more items.

c The Self Efficacy for Peer Interaction – Non-Conflict Subscale averages student responses to four items from the student survey: (G10) "Some kids are deciding what game to play. How easy or hard would it be for you to tell them about a game you like?"; (G11) "Some kids need more people to be on their teams. How easy or hard would it be for you to ask to be on their team?"; (G12) "Your class is going on a trip and everyone needs a partner. How easy or hard would it be for you to ask someone to be your partner?"; and (G13) "Some kids are going to lunch. How easy or hard would it be for you to ask if you can sit with them?" Responses are coded on a 4-point scale ranging from 1 (really hard) to 4 (really easy) for all items. Higher values on the scale indicate more effective interacting with peers in nonconflict situations. The scale is coded as missing if responses were missing for two or more items.

d The Social Competence – Emotional Regulation Subscale averages teacher responses to six items from the student-specific portion of the teacher survey: (H4) "This student can accept things not going his/her way."; (H5) "This student copes well with failure."; (H6) "This student accepts legitimate imposed limits."; (H7) "This student thinks before acting."; (H8) "This student can calm down when excited or all wound up."; and (H9) "This student plays by the rules of the game." Responses are coded on a 4-point scale ranging from 0 (never) to 3 (almost always) for all items. Higher values on the scale indicate higher student emotional regulation. The scale is coded as missing if responses were missing for four or more items. Teachers reported on up to five students. A scale was constructed for each student; for each teacher, we then constructed a weighted average of his or her students' scales. Impacts are estimated on the teacher-level weighted average. The reported sample size is the number of students about whom teachers reported.

e The Social Competence – Prosocial Behavior Subscale averages teacher responses to 12 items from the student-specific portion of the teacher survey: (H10) "This student expresses needs and feelings appropriately."; (H11) "This student resolves peer problems on his/her own."; (H12) "This student is very good at understanding other people's feelings."; (H13) "This student is aware of the effect of his/her behavior on others."; (H14) "This student works well in a group."; (H15) "This student controls his/her temper when there is a disagreement."; (H16) "This student shares materials with others."; (H17) "This student cooperates with peers without prompting."; (H18) "This student is helpful to others."; (H19) "This student listens to others' points of view."; (H20) "This student can give suggestions or opinions without being bossy."; and (H21) "This student acts friendly toward others." Responses are coded on a 4-point scale ranging from 0 (never) to 3 (almost always) for all items. Higher values on the scale indicate higher student prosocial skills. The scale is coded as missing if responses were missing for seven or more items. Teachers reported on up to five students. A scale was constructed for each student; for each teacher, we then constructed a weighted average of his or her students' scales. Impacts are estimated on the teacher-level weighted average. The reported sample size is the number of students about whom teachers reported.

f The Responsibility Scale averages teacher responses to three items from the student-specific portion of the teacher survey: (H1) "This student apologizes when he/she has done something wrong."; (H2) "This student can wait in line patiently when necessary."; and (H3) "This student takes responsibility for his/her own actions." Responses are coded on a 4-point scale ranging from 0 (never) to 3 (almost always) for all items. Higher values on the scale indicate more student responsibility. The scale is coded as missing if responses were missing for one or more items. Teachers reported on up to five students. A scale was constructed for each student; for each teacher, we then constructed a weighted average of his or her students' scales. Impacts are estimated on the teacher-level weighted average. The reported sample size is the number of students about whom teachers reported.

g The Altruism Scale averages teacher responses to two items from the student-specific portion of the teacher survey: (H22) "This student helped someone who was being picked on."; and (H23) "This student stopped a child from hurting another child." Responses are coded on a 4-point scale ranging from 0 (never) to 3 (often). Higher values indicate higher levels of student altruism. The scale is coded as missing if responses were missing for either item. Teachers reported on up to five students. A scale was constructed for each student; for each teacher, we then constructed a weighted average of his or her students' scales. Impacts are estimated on the teacher-level weighted average. The reported sample size is the number of students about whom teachers reported.

TABLE 10.
Teacher Perceptions of Playworks

	Sample Size (data source)	**Treatment**
Percentage of Teachers that "Agree" or "Strongly Agree" that the Playworks Program:	175 (teacher survey)	
Helps students stay out of trouble		89.0
Provides positive experiences for students during recess		98.9
Reinforces positive behavior during recess		95.8
Addresses important student needs at their school		86.1
Takes away important time that children have for unstructured play		6.0
Is valued by the staff at their school		86.4
Is valued by the students at their school		95.2
Is valued by the parent community at their school		59.1
Percentage of Teachers that "Agree" or "Strongly Agree" that Playworks Recess Activities:	174 (teacher survey)	
Help their students learn new games		96.8
Help them learn new games		83.6
Help their students learn recess rules		94.7
Are fun for their students		98.9
Provide good exercise for their students		97.2
Take away from students' academic learning		5.7
Allow them to play with their students		58.1
Are an important part of Playworks		96.0
Percentage of Teachers that Report Hoping that Playworks Is Implemented in the Future at Their School	173 (teacher survey)	96.9

Source: Teacher surveys conducted in spring 2011 and spring 2012.

Note: Sample sizes based on the same data source might be different due to missing responses.

TABLE 11.
Impacts on Student Behavior

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Classroom Behavior					
Disruptive Behavior in Class Scale Score[a]	2,277 (student survey)	1.9	2.0	-0.1	0.98
Classroom Misbehavior/ Discipline Scale Score[b]	1,431 (teacher survey, student-specific portion)	0.5	0.6	0.0	1.00
General Behavior					
Bad Behaviors Scale Score[c]	2,286 (student survey)	0.3	0.4	-0.1	0.80
Attendance					
Percentage of Teachers that Reported that Their Students Were:[d]	1,425 (teacher survey, student-specific portion)				
Absent "two or more times" in the past 30 days		27.6	29.7	-2.1	1.00
Late for class "two or more times" in the past 30 days		15.3	20.5	-5.3	0.99
Percentage of Students in Attendance Each Day, on Average[e]	25 (administrative records)	94.2	93.3	0.9	0.55
Percentage of Students that Were Chronically Absent[f]	29 (administrative records)	15.9	17.4	-1.5	1.00
Discipline					
Percentage of Teachers that Reported that Their Students:[d]	1,427 (teacher survey, student-specific portion)				
Were suspended during this school year		0.1	0.1	0.0	1.00
Received a detention in the past 30 days		21.9	24.0	-2.1	1.00
Suspensions per 100 Students:[g]	29 (administrative records)	14.6	10.1	4.5	0.94
Expulsions per 100 Students:[h]	24 (administrative records)	0.1	0.3	-0.2	0.99
Number of Disciplinary Incidents:[i]	26 (discipline referral data)				
Overall		13.0	20.8	-7.8	0.50
At recess		2.4	2.4	0.0	1.00
In class		8.0	11.4	-3.4	0.86
In other location		2.7	5.2	-2.5	0.65
For fighting		2.0	4.3	-2.3	0.77

TABLE 11. (continued)
Impacts on Student Behavior

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Classroom Behavior					
For profanity		0.3	1.1	-0.8	0.41
For disrespect		3.1	4.3	-1.2	1.00
For harassment		0.5	1.2	-0.7	0.85
For disruption		2.9	3.9	-1.0	0.99
For another reason		1.7	2.6	-0.9	0.97
For multiple reasons		2.6	3.3	-0.7	1.00

Sources: Student and teacher surveys conducted in spring 2011 and spring 2012; discipline referral data collected in spring 2011 and spring 2012; administrative records.

Note: Random assignment block indicators were included as covariates in all impact models; indicators for teacher's race were included as covariates in the impact models for teacher survey outcomes. The *p*-values reported in this table account for clustering of students and teachers within schools and for multiple hypothesis testing (MHT) to control the probability of finding any falsely significant impacts (the family-wise error rate) at 5 percent. The adjustment for MHT is based on the multivariate t-distribution and takes into account correlations among test statistics. The adjustment accounts for the tests presented in this table, but not for tests presented in other tables. Sample sizes based on the same data source might be different due to missing responses. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding.

a The Disruptive Behavior in Class Scale averages student responses to five items from the student survey: (C8) "I sometimes annoy my teacher during class."; (C9) "I sometimes get into trouble with my teacher during class."; (C10) "I sometimes behave in a way during class that annoys my teacher."; (C11) "I sometimes don't follow my teacher's directions during class."; and (C12) "I sometimes disturb the lesson that is going on in class." Responses are coded on a 4-point scale ranging from 1 (not at all true) to 4 (very true). Higher values on the scale indicate disruptive behavior. The scale is coded as missing if responses were missing for two or more items.

b The Classroom Misbehavior/Discipline Scale averages teacher responses to three items from the student-specific portion of the teacher survey: (F9) "How many times in the past 30 days of school have you disciplined this child for misbehaving in class (e.g., asked this child to sit-out, miss recess)?"; (F10) "How many times in the past 30 days of school have you sent this child to the principal's office for misbehaving in class?"; and (F11) "How many times in the past 30 days of school have you contacted this child's parents regarding his/her behavior in class?" Responses are coded on a 4-point scale ranging from 0 (never) to 3 (four or more times). Higher values on the scale indicate greater misbehavior. The scale is coded as missing if responses were missing for one or more items. Teachers reported on up to five students. A scale was constructed for each student; for each teacher, we then constructed a weighted average of his or her students' scales. Impacts are estimated on the teacher-level weighted average. The reported sample size is the number of students about whom teachers reported.

c The Bad Behaviors Scale averages student responses to six items from the student survey: (C23) "During this school year, how many times were you sent to the principal's office for doing something wrong at recess?"; (C24) "During this school year, how many times did you get in trouble for something that happened during recess?"; (C25) "During this school year, how many times were you sent to the principal's office for bad behavior in the classroom?"; (C26) "During this school year, how many times did you have to sit out at recess for bad behavior?"; (C27) "During this school year, how many times were you given a detention (for example you had to stay after school or miss lunch for bad behavior)?"; and (C28) "During this school year, how many times were you suspended?" Responses are coded on a 4-point scale ranging from 0 (never) to 3 (five or more times). Higher values on the scale indicate greater engagement in bad behavior. The scale is coded as missing if responses were missing for two or more items.

d Teachers reported on up to five students. For each teacher, we constructed a weighted average of his or her students' outcomes. Impacts are estimated on the teacher-level weighted average. The reported sample size is the number of students about whom teachers reported.

e Percentage of Students in Attendance Each Day, on Average, comes from school administrative records and is typically calculated by dividing total days of attendance (the sum of days attended across all students) by the total number of student-days in the school year. The percentage is coded as missing if the definition was not known or was not consistent within schools in the random assignment block.

f Percentage of Students that Were Chronically Absent comes from school administrative records and is defined differently in different schools. Definitions include: (i) the percentage of students absent 10 or more days, (ii) the percentage of students absent 20 or more days, (iii) the percentage of students with 10 or more unexcused absences, (iv) the percentage of students absent for 10 or more percent of school days, (v) the percentage of students absent for 11 or more percent of school days, and (vi) the percentage of students with an unexcused

absence or tardy of more than 30 minutes on three or more days. Within random assignment block, all schools used the same definition of chronic absence.

g Suspensions per 100 Students comes from school administrative records. The number of suspensions is defined differently in different schools, sometimes including and sometimes excluding in-school suspensions. Within random assignment block, all schools used the same definition of suspensions.

h Expulsions per 100 Students comes from school administrative records. It is coded as missing if the definition of expulsions was not known or was not consistent within schools in the random assignment block.

i Schools were asked to report on discipline referrals that occurred over the course of one week during the spring semester. Each referral to the office during the week was treated as a separate incident. Discipline referral data were missing for one of the control schools; therefore, discipline referral data from this control school and the two treatment schools matched to that control school were not included in the impact analysis.

TABLE 12.
School Contextual Issues and Pre-Implementation Features in Treatment Schools

	Sample Size (data source)	
Percentage of Schools that:[a]	16 (principal interviews)	
Had recess in year before evaluation		87.5
Had a new principal in study year		37.5
Did not make AYP in prior year		43.8
Percentage of Schools in Which the Following Person Was Responsible for Bringing in the Program:	17 (principal interviews)	
Current principal		64.7
Another school staff member		11.8
School parent organization		11.8
Another person		11.8
Percentage of Schools in Which the Program Was Funded by:[b]	17 (principal interviews)	
School budget/Title I funds		58.8
School district		17.6
External grant		11.8
Parent organization		17.6
Percentage of Schools in Which the Key Reason the Program Was Desired Was to:[b]	17 (principal interviews)	
Organize/formalize recess games		52.9
Improve sense of community, teamwork, and school climate		35.3
Increase safety/decrease conflicts		23.5
Increase physical activity		23.5
Support the introduction of recess		17.6
Promote leadership		11.8

Source: Interviews with treatment school principals conducted in spring 2011 and spring 2012.

a One school was not in operation prior to the study year and therefore does not have prior year recess, principal, or AYP information.

b Percentages do not sum to 100 because data capture multiple responses per school.

AYP = adequate yearly progress.

TABLE 13.
Teacher Participation in and Perceptions of Playworks Class Game Time

	Sample Size (data source)	**Treatment**
Percentage of Teachers that Report that Their Class Participated in Playworks Class Game Time 2 or More Times in the Past 30 Days	174 (teacher survey)	70.4
Percentage of Teachers that "Agree" or "Strongly Agree" that Playworks Class Game Time:	174 (teacher survey)	
Helps their students learn new games		90.2
Helps them learn new games		85.3
Helps their students learn recess rules		81.9
Helps them learn recess rules		60.7
Is fun for their students		92.3
Is good exercise for their students		88.1
Takes away from students' academic learning		12.8
Allows them to play with their students		70.3
Is an important part of Playworks		85.5

Source: Teacher surveys conducted in spring 2011 and spring 2012.

Note: Sample sizes based on the same data source might be different due to missing responses.

TABLE 14.
Student Participation in and Perceptions of Playworks

	Sample Size (data source)	All Treatment Students	Treatment Students with Above-Average Participation[a]	Treatment Students with Below-Average Participation[a]
Percentage of Students that Report that They Did the Following "A Few Times" or "Many Times":	1,267 (student survey)			
Participated in activities organized by their Playworks coach during the past two weeks		72.9	92.1	49.5
Participated in games with their Playworks coach during class game time during the past two weeks		73.3	94.1	48.2
Participated in Playworks games led by their teacher in class during the past two weeks		48.0	71.6	19.3
Used ro-sham-bo to resolve conflicts during recess		68.3	88.7	43.5
Percentage of Students that Report that They Agree "A Little" or "A Lot" with the Following:	1,268 (student survey)			
I enjoy participating in class game time with my Playworks coach		90.3	96.3	83.0
I enjoy participating in Playworks games with my teacher		82.0	89.3	73.2
I enjoy recess activities with my Playworks coach		88.6	96.7	78.7
My Playworks coach does games with us that I like to play		86.9	93.8	78.4

Source: Student surveys conducted in spring 2011 and spring 2012.

Note: Sample sizes based on the same data source might be different due to missing responses.

a The treatment students with above-average participation in Playworks were the students with a value above the overall unweighted mean score for the Participation in Playworks Activities Scale. The treatment students with below-average participation were the students with a value below the overall unweighted mean score for the Participation in Playworks Activities Scale. The Participation in Playworks Activities Scale averages student responses to four items from the student survey: (I2) "During the last two weeks, how often have you participated in activities organized by your Playworks coach during recess?"; (I3) "During the last two weeks, how often has your class participated in games with your Playworks coach during class game time?"; (I4) "During the last two weeks, how often has your teacher led you in Playworks games during class?"; and (I5) "During recess, how often do you use 'ro-sham-bo' to resolve conflicts?" Responses are coded on a 4-point scale ranging from 0 (never) to 3 (many times). Higher values on the scale indicate greater participation in Playworks activities. The scale is coded as missing if responses were missing for two or more items.

Appendix 3
Additional Analyses

Differences in Impacts on Teacher- and Student-Reported Outcomes

This study found that Playworks had a statistically significant positive effect on several outcomes. The significant findings follow an interesting pattern—most of the positive impacts observed were found on outcomes reported by teachers. In particular, significant impacts were found on seven teacher-reported outcomes and one student-reported outcome.

There are many possible explanations for this pattern, including several that we were able to investigate:

- **Differing grade levels of teacher and student respondents.** The differences in impacts could be due to the fact that the teacher surveys were completed by teachers in grades one through five, whereas the student surveys were only completed by 4th and 5th grade students. To examine whether this was the case, we examined two sets of findings where the teacher-reported outcome was significant and the related student-reported outcome was not. These two areas were (1) the teacher- and student-reported scales on student safety and (2) the bullying/exclusion scale for teachers and the related aggressive behavior scale for students. To assess whether the difference in findings might have been due to different grades being used in the analyses, we re-estimated the impact of Playworks on the two teacher-reported scales by limiting the sample to only teachers of 4th- and 5th-grade classes. For this subsample—like the overall teacher sample—teachers in treatment schools reported significantly better safety and less bullying/exclusion than teachers in control schools. Therefore, the difference in grades between the teacher and student survey samples does not explain the difference in these findings.

- **Survey questions not appropriate to the study population.** Differences between teacher- and student-reported impacts could also arise if the survey questions were not appropriate for the study participants and did not correctly measure the constructs (e.g., if student survey items were not meant to be used with students in grades four and five or with students that are mostly black or Hispanic). However, the student scales used to measure school climate, conflict resolution/aggression and learning/academic performance—the three domains in which there were significant program impacts—were either developed and validated using populations similar to the sample in the current study or were created specifically for this study and pre-tested with students who had similar backgrounds. Most of these scales have also been used in previous impact studies, though not necessarily as part of studies that examined the impact of recess interventions.

- **Differences in the way scales were constructed.** Impacts on teacher and student outcomes could also differ because not all components of the scales used to measure outcomes are the same. For example, even though the items that make up the teacher- and student-reported safety at school and safety at recess scales are similar, the student scales both include four items, whereas the teacher scales include fewer items (Appendix 2, Table 3). When examining impacts on individual items from these scales, however, impacts on most teacher-reported individual items were significant and

impacts on all student-reported individual items were not significant, suggesting that this does not explain the difference in teacher- and student-reported findings for safety.

Though we can rule out these explanations, there are other possible reasons for the differences in impacts on outcomes reported by teachers and students that we cannot directly investigate to assess whether they explain the results:

- **Differences in framing or reference periods of survey questions.** Different questions were used to measure the teacher-reported student bullying/exclusion outcome (for which there were significant impacts) and the student-reported aggressive behavior outcome (for which impacts were not significant). Teachers were asked to report how often students had reported bullying or exclusionary behaviors; in contrast, students were asked how often they themselves perpetrated aggressive behavior. Though many of the items covered similar topics (for example, name-calling and hitting or pushing), the way that the items were framed and the reference period that was used (the "past 30 days" for the teachers' bullying/exclusion scale, and the "past two weeks" for the students' aggressive behavior scale) are different, which could result in differences between teacher and student reports.

- **Differences in teacher and student perceptions.** Teacher and student perceptions could simply be different due to the fact that they have different experiences at school; for example, students spend more time at recess than teachers, whereas teachers might observe recess informally or hear reports from students, coaches or playground monitors after recess. Teachers' perceptions are also likely based on their whole classroom, whereas students were asked to report on their own perceptions or experiences.

This study intentionally included both teacher- and student-reported outcomes because measuring impacts from both perspectives provides a more complete picture of the impacts of Playworks. Though there are differences between teacher and student reports, the pattern of findings suggests that Playworks led to improvements on some measures of school climate, conflict resolution/aggression and learning/academic performance, particularly on measures reported by teachers.

Cost Analysis

Administrators and policymakers interested in comparing Playworks to other programs might be interested in understanding how program impacts relate to costs. To provide evidence on the cost-effectiveness of Playworks, we estimated the costs of the program. To participate in Playworks, a school (or its district) contributes to the upfront costs of the program. For the schools in this study, the cost to participate in one year of Playworks ranged from $20,500 to $25,500 per school (depending on the school); the average amount contributed by a school or district for a school to participate in the program was $24,353. However, because of Playworks' efforts to secure donations and grants, the school or district contribution does not reflect the true upfront cost of the program. In particular, according to Playworks, the actual cost of providing Playworks to a single school was $61,200 in the 2010-2011 school year and $64,600 in the 2011-2012 school year (based on national estimates).

Furthermore, program-reported, upfront costs represent only part of the costs borne by schools or districts. Principal interviews in the 17 treatment schools gathered information on other types of costs, including:

- **Staff time.** The interviewers asked principals about time spent supervising recess, leading physical education, implementing the Playworks program, addressing behavioral/disciplinary events, staffing before- or after-school programs or carrying out other activities. For each category, principals were asked whether Playworks changed the number of staff positions or time that staff spent, whether the change was an increase or decrease (in paid staff or volunteer hours), the amount of time or positions added/cut and (if a paid position) the salary or hourly rate. Some principals reported staffing changes in response to Playworks; these included a reduction in volunteer time spent on recess (reported by one principal), a reduction in staff time devoted to physical education (reported by one principal), a reduction in staff time devoted to Playworks implementation (reported by one principal), a reduction in staff time spent addressing behavioral/disciplinary events (reported by two principals), a reduction in volunteer time spent on a before- or after-school program (reported by one principal), and a reduction in volunteer and staff time spent on special event and other activity planning (reported by one principal). Only one school's principal reported an increase in staff hours in response to Playworks: a four-hour per week increase in staff time on an after-school program. To include staffing changes in the cost estimates, we estimated the value of staff time using salary or hourly rate information.[1,2]
- **Space/equipment for play.** The interviewers asked principals about the purchase of playground equipment or materials, improvements to the playground or other spaces and other space or equipment changes. If a principal said that Playworks changed the school's decision in one of these areas, the principal was asked to describe the change and the cost (or cost savings) associated with that change.[3,4] Principals from two schools reported positive equipment costs, and one of them also reported positive space improvement costs. In a third school, the principal reported cost savings on both equipment and space improvements. Changes in equipment and space costs ranged from -$400 to $400.
- **Other student and staff programs.** Interviewers asked principals about changes to professional development and other programs funded by the school. If the principal said that Playworks changed the school's programming, the principal was asked to describe the change and the cost (or cost savings) associated with that change. All but one principal reported no changes to student and staff programs; in the remaining school, the principal reported that the school saved $200 because it no longer needed assemblies teaching sportsmanship, leadership and other qualities.

1 If information about salary or hourly rate was not available (or not applicable, in the case of volunteer hours), we converted changes in staff and volunteer time to dollar cost estimates using metropolitan area-level estimates of hourly mean earnings for full-time workers from the 2010 National Compensation Survey. In particular, teacher time was valued at the wage for elementary school teachers (except special education teachers), and volunteer time was valued at the wage for teacher assistants.

2 Some changes in staff time were reported at the weekly or monthly level. To convert those estimates to an annual value, we assumed a 36-week or 9-month school year.

3 In some cases, principals reported the purchase of equipment or other materials but did not provide an estimate of its value. Treating these cases as missing could lead us to understate program costs (or cost savings). Instead, we imputed the value of equipment as the average value across other schools reporting equipment purchases.

4 In a few cases, principals reported receiving donations of equipment or other supplies. The value of donated items is not included in these cost estimates.

As described above, principals reported few additional costs of Playworks. Of the 17 treatment schools, 11 principals reported no additional costs of Playworks in these categories. Of the others, 4 reported a cost savings and 2 reported an increase in costs. Because 2 schools reported substantial cost savings, the average additional cost of Playworks reported by principals was actually a cost *savings* of $6,492.

We constructed two estimates of average costs for the schools in this study. The first, the cost from the school or district's perspective, includes the upfront costs of the Playworks program borne by schools or districts (average of $24,353) as well as the average costs reported in the principal interview (a cost savings of $6,492), for a total cost per school of $17,861. We take the perspective of the policymaker in our second estimate of average costs. In that case, the cost is the total upfront cost (including costs covered donations or grants) of the Playworks program (average of $61,800) minus the cost savings reported in the principal interview (a cost savings of $6,492), for a total cost per school of $55,308. To compare Playworks to other programs, administrators or policymakers can compare the impact per dollar for different outcomes to the impact per dollar under other programs.

One potential limitation of our analysis is that our cost estimates are based on costs reported by school principals during their first year of Playworks implementation; the costs of Playworks may be lower (or higher) in subsequent years. Another potential limitation is that our analysis may not account for all possible costs (or benefits) of the Playworks program. To capture another dimension of the potential benefits of Playworks, we asked treatment group teachers responding to the teacher survey about changes in their time use as a result of having the Playworks program at their school for one school year.[5] On average, treatment group teachers reported that Playworks led to (1) small increases in the amount of time they spent supervising recess, leading physical education and participating in professional development and training related to Playworks and (2) a larger decrease in the amount of time they spent addressing behavioral or disciplinary issues (Exhibit 1).[6] The decrease in time teachers reported spending addressing behavioral problems is consistent with the impact finding that showed a reduction in the bullying/exclusion scale. Conversely, the increase in time teachers reported spending on supervision of recess seems somewhat counterintuitive; we know from our interviews with principals, however, that teachers in one treatment school were asked to serve as recess monitors, in order to give them direct experience with Playworks games and activities. When data from this school are excluded from the calculation, the average increase in time spent on supervision of recess is an increase of 0.9 hours per teacher over one school year. Together, these findings show that teachers report a modest net time savings, on average, associated with participating in Playworks.

5 In a few cases, a teacher reported a change in time spent but did not report the magnitude of the change. In these cases, we imputed the magnitude as the average among other teachers in the same school (or same city, if necessary) who reported a non-zero change.

6 The average change obscures substantial variation. Thirty-nine percent of teachers reported no change in time spent in any of the four categories; the median amount of time spent by category was zero hours. However, some teachers reported large (positive or negative) changes in the amount of time spent on these activities.

EXHIBIT 1.
Teacher-Reported Changes in Time Use Due to Playworks

	Average Change in Hours per Classroom Teacher, Annually	Average Change in Hours per Classroom Teacher, Annually
	17 Schools (full sample)	**16 Schools (excluding outlier)**
Changes (Due to Playworks) in Time Classroom Teachers Spent:		
Supervising recess	6.17	0.90
Leading physical education	0.96	1.05
Participating in professional development or training related to Playworks	0.98	0.96
Addressing behavioral or disciplinary issues	-9.37	-8.79

Source: Teacher surveys (n = 172) conducted in treatment schools in spring 2011 or spring 2012 (sample sizes may be smaller for some outcomes due to missing responses).

Note: Average change in time spent is the unweighted mean. The second column excludes one school in which teachers served as recess monitors to build experience with Playworks.

ABOUT THE ROBERT WOOD JOHNSON FOUNDATION

Robert Wood Johnson Foundation

The Robert Wood Johnson Foundation focuses on the pressing health and health care issues facing our country. As the nation's largest philanthropy devoted exclusively to health and health care, the Foundation works with a diverse group of organizations and individuals to identify solutions and achieve comprehensive, measurable, and timely change.

For 40 years, the Foundation has brought experience, commitment, and a rigorous, balanced approach to the problems that affect the health and health care of those it serves. When it comes to helping Americans lead healthier lives and get the care they need, the Foundation expects to make a difference in your lifetime.

For more information, visit *www.rwjf.org*. Follow the Foundation on Twitter at *www.rwjf.org/twitter* or on Facebook at *www.rwjf.org/facebook*.

Route 1 and College Road East
P.O. Box 2316
Princeton, NJ 08543-2316

ABOUT MATHEMATICA

MATHEMATICA
Policy Research

Mathematica Policy Research seeks to improve public well-being by conducting studies and assisting clients with program evaluation and policy research, survey design and data collection, research assessment and interpretation and program performance/data management. Its clients include foundations, federal and state governments and private-sector and international organizations. The employee-owned company, with offices in Princeton, NJ; Ann Arbor, MI; Cambridge, MA; Chicago, IL; Oakland, CA; and Washington, DC; has conducted some of the most important studies of education, health care, nutrition, international, disability, family support, employment and early childhood policies and programs.

P.O. Box 2393
Princeton, NJ 08543-2393
www.mathematica-mpr.com

ABOUT THE JOHN W. GARDNER CENTER FOR YOUTH AND THEIR COMMUNITIES

john w. gardner
center for youth and their communities

The John W. Gardner Center for Youth and Their Communities at Stanford University partners with communities to carry out three inter-related goals. *Develop Leadership:* Build relationships and capacity among community organizations to identify shared challenges; foster partnership between the university and community to engage in evidence-based inquiry and decision-making to find common solutions related to youth and communities. *Conduct Research:* Collect and analyze data to understand youth across contexts and across a range of developmental domains; engage in high-quality evaluation of youth-serving programs and services. *Effect Change:* Support community stakeholders to translate research findings into actionable knowledge, and to identify the most effective levers for programmatic and policy improvement.

Stanford University School of Education
365 Lasuen Street,
Stanford, CA 94305-3083
jgc.stanford.edu