Findings from an Experimental Evaluation of Playworks: Effects on Play, Physical Activity and Recess

MATHEMATICA Policy Research

john w. gardner center for youth and their communities

Robert Wood Johnson Foundation

EVALUATION REPORT

May 2013

Findings from an Experimental Evaluation of Playworks: Effects on Play, Physical Activity and Recess

TABLE OF CONTENTS

RESEARCHERS

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

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

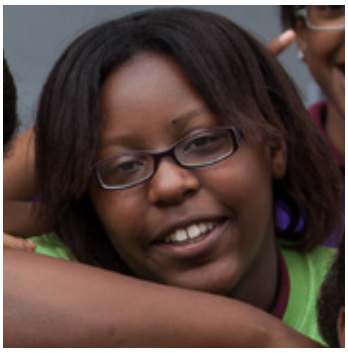
Recess periods often lack the structure needed to support physical activity and positive social development (Robert Wood Johnson Foundation 2010). The Playworks program places full-time coaches in low-income schools to provide opportunities for organized play during recess and throughout the school day. 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 program's impact on key outcomes related to play, physical activity and recess.

Additional analyses on the implementation of Playworks and its impact on school climate, academic learning, students' social skills and behavior are presented in a series of other study briefs: Bleeker et al. (2012), Fortson et al. (2013) and London et al. (2013).

Key Findings

The following significant impacts of Playworks on play, physical activity and recess were found:



- Accelerometer data showed that Playworks had a positive impact on students' physical activity during recess. Students in treatment schools engaged in physical activity during recess that was, on average, more intense than the physical activity engaged in by control students. Moreover, students in treatment schools spent significantly more time engaged in vigorous physical activity at recess than students in control schools.
- The main recess activity in which students were observed to be engaged was less likely to be a sedentary activity (such as sitting and talking) in treatment schools, compared with control schools.
- Playworks had an impact on the extent to which recess activities were organized by adults. The percentage of recess activities that were organized by school staff (or a Playworks coach, in the case of treatment schools) was significantly higher in treatment schools than in control schools. A significantly higher percentage of students in treatment schools, compared with control schools, also reported that adults helped them play games and sports "sometimes" or "a lot" during recess.
- Playworks had a positive impact on the availability of equipment at recess. Equipment was available at a significantly higher percentage of recess periods in treatment schools, compared with control schools. Some types of equipment (including jump ropes, cones and bases for ball games) were available at a higher percentage of recess periods in treatment schools, compared with control schools.
- Teachers in treatment schools reported better recess behavior and readiness for class than teachers in control schools.
- A higher percentage of teachers in treatment schools, compared with control schools, agreed or strongly agreed that their students enjoyed adult-organized activities at recess and felt ownership over their activities during recess.

…students in treatment schools spent significantly more time engaged in vigorous physical activity at recess than students in control schools.

The following key implementation findings related to play, physical activity and recess were observed:

- About three out of every four teachers interviewed at treatment schools commented that Playworks had provided students with increased opportunities for physical activity.
- Students in treatment schools were observed engaging primarily in organized games during recess, such as four-square, tetherball, or kickball, whereas students in control schools were observed engaging in games that were less organized, such as tag games without clear rules or boundaries.
- We observed students at treatment schools using the equipment as intended during recess. There was only one instance at one treatment school where equipment was being misused. In contrast, we observed students not using equipment as intended in 44 percent of control-school recesses. Moreover, we observed that misuse of equipment in control schools often led to unsafe behavior.

Impact and Implementation Findings from an Experimental Evaluation of Playworks: Effects on Play, Physical Activity and Recess

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).

In schools where recess is still offered, recess periods often lack the structure needed to support physical activity and positive social development.

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 (along with staff trainings), 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 and a sports league.

Past Research

The Playworks program attempts to engage students in more play and physical activity during recess periods by offering organized games and activities. Past research suggests that such an approach may foster a number of desirable student outcomes (Murray et al. 2013; CDC 2010; DHHS 2008; Ginsburg 2007). Recess has also been shown to play a critical role in students' academic and cognitive development (Robert Wood Johnson Foundation 2007; Ginsberg 2007; Pellegrini 2005; Jarrett et al. 1998; Pellegrini et al. 1995), social and emotional development (CDC 2000; Action for Healthy Kids 2011; Pellegrini et al. 2002) and physical development (Robert Wood Johnson Foundation 2007; Kahan 2008; Wechsler et al. 2000) by offering more opportunities for students to engage in physical activity during the school day. Moreover, increasing school-aged children's engagement in physical activity during recess has the potential to help improve health and reduce obesity (Ogden et al. 2006).

There is some debate about whether structured recess or unstructured recess is better for promoting physical activity, which, in turn, can help foster positive social and academic behaviors (Murray et al. 2013). Some believe recess periods should be structured like physical education classes to encourage students to be more active (Murray et al. 2013). A report by the Robert Wood Johnson Foundation (2007) cites some of the benefits associated with structured recess periods, such as improved behavior and attention in the classroom after recess (as reported by teachers) and development of interpersonal skills through coaching from adults. Others argue that the focus should be on offering a desirable play space for unstructured recess periods and not so much on making sure the recess periods are structured (Murray et al. 2013; Pellegrini and Bohn 2005; Holmes et al. 2006).

Playworks' activities are organized with common rules and separate play spaces, but they are unstructured in the sense that students are free to choose one of the Playworks offerings or another activity of their choosing (London et al. 2010). In other words, Playworks activities are intended to organize the playground and make playing easier and more available to students, but not to make recess structured like a physical education class. In this current brief, we investigate the impact of Playworks on outcomes related to play, physical activity and recess.

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 from analyses of both cohorts showed that Playworks had positive impacts on some measures of school climate, conflict resolution and aggression and learning and academic performance, and no negative impacts in any of the five domains that were examined (Fortson et al. 2013). Findings from analyses of cohort 1 alone also showed that Playworks had positive impacts on measures of recess experience (Bleeker et al. 2012).[1] Playworks implementation across the two cohorts was observed to be moderate or strong in 14 of 17 treatment schools, with stronger implementation occurring in schools that had recess in the past (London et al. 2013). Implementation findings also showed that students, teachers and principals had mostly positive perceptions of Playworks (London et al. 2013).

This current brief focuses on a different set of analyses—analyses that examine the impact of Playworks on measures of play, physical activity and recess. We set out to address the following research question related to the program's impact in this area:

1. What is the effect of Playworks on:

   (1) objective measures of students' physical activity at recess[2]

   (2) teachers' and students' reports of physical activity at recess

   (3) students' recess activities

   (4) availability and condition of recess equipment

   (5) student interactions at recess and

   (6) students' and teachers' perceptions of recess?

This brief also looks qualitatively at how Playworks is related to play, physical activity and recess.[3]

Study Design

Twenty-nine schools (17 treatment schools, 12 control schools) from six cities across the United States 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 schools' observed and unobserved characteristics and that the differences in outcomes between the two groups could be attributed solely to the effect of Playworks.[4] 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. 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. This brief focuses on outcomes related to play, physical activity and recess. The five other outcome domains (school climate, conflict resolution and aggression, learning and academic performance, youth development and student behavior) are the focus of three other study briefs (Bleeker et al. 2012; Fortson et al. 2013; London et al. 2013). For this study, 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 physical activity during recess, participation in sports and clubs outside of school, recess activities, interactions at recess and perceptions of recess.
- ***Teacher Survey.*** A total of 296 teachers from 29 study schools participated in a survey that included items related to student physical activity during recess, student interactions at recess and perceptions of student feelings about recess.

- ***Accelerometers.*** A total of 1,579 students from 98 4th- and 5th-grade classrooms in 28 study schools wore accelerometers during one or two school days for at least 10 minutes during their recess periods.[5] The accelerometers recorded intensity counts, which measure the intensity of students' movement. The accelerometers also record the number of steps taken during the period of time students wore the accelerometers.

- ***Structured Recess Observations.*** A total of six structured recess observations were conducted at each of the 29 study schools. These observations captured information on recess equipment, student physical activity levels, types of activities and student interactions. During each recess observation, students were observed in designated zones of the play area. The observations that took place within specific zones of the play area will be referred to as "scans" in this study brief. Sample sizes in the tables are reported either as the number of scans or the number of recess observations (six per school), depending on the outcome's unit of analysis.

- ***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.

- ***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.[6]

- ***Semi-Structured Recess Observations.*** Semi-structured observations of two days of recess (all or nearly all recess periods in each of the two days) were conducted at each of the 29 study schools. Unlike the structured recess observations described above, which included observations of students within zones of the play area, these observations focused on the play area as a whole, capturing information on organization of recess, engagement in games and play, conflict and conflict resolution and the inclusiveness of behavior by students and adults.

- ***Playworks Class Game-Time Observations***. Observations of Playworks class game time were conducted at the 17 treatment schools to examine the coaches' relationships with students in smaller groups and to examine teacher and coach interactions, discipline styles and students' physical activity levels.

Response rates and additional details about each of the study's data sources can be found in Appendix 1.

Key Impact and Implementation Findings

For this brief, we examined the effect of Playworks on play, physical activity and recess. Significant impacts were observed in the areas of physical activity at recess, recess activities, availability and condition of recess equipment, student interactions at recess and students' and teachers' perceptions of recess. (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 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.

Physical Activity at Recess. Playworks had a positive impact on physical activity at recess, as measured by accelerometers. In particular, 4th- and 5th-grade students in treatment schools had significantly higher average physical activity intensity levels during recess than 4th- and 5th-grade students in control schools (see Exhibit 1 and Appendix 2, Table 3).

We also used the accelerometer data to examine how much time students spent during recess in specific activity groups that are often used to categorize different levels of physical activity—namely, sedentary, light, moderate and vigorous physical activity. Students in treatment schools spent significantly more time than students in control schools engaging in vigorous physical activity (Exhibit 1 and Appendix 2, Table 4). There were no significant differences between treatment and control students on the other individual activity categories (sedentary, light or moderate) (Exhibit 1 and Appendix 2, Table 4).

To help interpret these findings, we will clarify what the accelerometer outcomes described above measure. Accelerometer intensity counts are used to objectively measure physical activity and to compare physical activity levels across individuals and periods of time. If, for example, two people wearing accelerometers are engaging in two different activities during a given period of time, one of which is more physically demanding than the other, the accelerometer worn by the person engaging in the more physically demanding activity will record more intensity counts (per unit of time) than the accelerometer worn by the person engaging in the less strenuous activity. The amount of time spent in sedentary, light, moderate or vigorous activity is defined based on the accelerometer intensity counts. Specifically, we used the accelerometer intensity "cut points" defined by Edwardson and

Gorely (2010) to create outcome variables for time spent in sedentary, light, moderate and vigorous activity during recess. Examples of vigorous activities are running or jogging, lifting weights and swimming laps. Activities like walking briskly or riding a bike are commonly thought of as moderately intense activities.

Although significant impacts of Playworks on physical activity were observed based on the accelerometer data, we did not find significant impacts of Playworks on physical activity based on the student surveys or teacher surveys. Similar percentages of students at treatment and control schools reported that they participate in activities at recess that make them sweat and breathe hard "sometimes" or "a lot" and similar percentages of teachers in both groups agreed or strongly agreed that their students participated in activities at recess that made them sweat and breathe hard (Appendix 2, Table 5). When considering the structured recess observations, similar percentages of students at treatment and control schools were observed engaging in moderate or vigorous physical activity during recess (Appendix 2, Table 5).

When we focused on the rest of the school day (excluding recess periods), we found no significant impacts on physical activity based on the accelerometer data. Students in treatment and control schools who wore an accelerometer for four or more hours during the school day had similar levels of physical activity intensity (and spent similar amounts of time engaging in vigorously intense activity) during non-recess time.

Reports from teacher interviews at treatment schools were consistent with the accelerometer impact findings. Thirty-seven out of the 51 teachers interviewed at treatment schools for the implementation component of the study commented that Playworks provided students with more physical activity opportunities.

The semi-structured recess observations provided information about situations in which physical activity at both treatment and control schools was hindered. Physical activity levels were lower when bad weather kept students indoors for recess and class game time; when there were more students involved in a game; when the space available for play was small; and when students spent more time arguing than actually playing the game as intended.

EXHIBIT 1.
Impacts on Students' Physical Activity at Recess As Measured by Accelerometers

Outcome	Treatment	Control	Difference
Mean Number of Accelerometer Intensity Counts Recorded Per Minute During Recess	1312.9	1013.0	299.9*
Mean Number of Steps Taken Per Minute During Recess	29.7	25.4	4.3
Mean Percentage of Accelerometer Wear Time During Recess Spent in:			
Sedentary Activity	44.1	50.6	–6.5
Light Activity	19.6	18.6	1.0
Moderate Activity	22.7	21.2	1.5
Vigorous Activity	13.7	9.6	4.1*

Source: Accelerometer data from students (n = 1,579) collected in spring 2011 and spring 2012.

Note: See full tables in Appendix 2, Tables 3 and 4.

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

Recess Activities. During the structured recess observations (where scans were made within zones of the play area), we recorded the main activity in which students were engaged within each zone. We then coded these main activities into seven distinct categories, such as "sitting, talking or other sedentary activity" or "jumping, running, tag or chase games." The main recess activity in which students were observed to be engaged was less likely to be sedentary in treatment schools than control schools. (Exhibit 2 and Appendix 2, Table 6). However, no significant impacts were found for other coded main recess activities, such as standing or walking; jumping, running, tag or chase games; climbing, sliding or swinging on a play structure; or team sports like soccer and baseball (Exhibit 2 and Appendix 2, Table 6).

Playworks had an impact on the extent to which games and activities were organized by an adult during recess. During the structured recess observations, a significantly higher percentage of the recess activities at treatment than control schools were organized by adults—either school staff or, in the case of treatment schools, Playworks coaches (Exhibit 2 and Appendix 2, Table 6). Similarly, a higher percentage of students in treatment schools, compared with control schools, reported that adults helped them play games and sports during recess "sometimes" or "a lot" (Exhibit 2 and Appendix 2, Table 6).

These findings are consistent with what was observed during the semi-structured recess observations. Playworks coaches at treatment schools were observed setting up a variety of games and associated equipment and helping sustain games by moving around the playground to support games either by participating or by helping students follow the rules or resolve conflicts.

The semi-structured recess observations found that Playworks also had an impact on the number of games that were organized (either with or without adult assistance) during recess. The average number of organized games observed during the semi-structured recess observations was significantly larger at treatment schools than at control schools (Exhibit 2 and Appendix 2, Table 6).

EXHIBIT 2.
Impacts on Students' Recess Activities

Outcome (mean unless otherwise noted below)	**Treatment**	**Control**	**Difference**
Observed Recess Activities			
Percentage of Scans in Which the Following Activity Was the Main Activity Observed (Structured Recess Observations):			
Sitting, talking or other sedentary activity	6.7	16.8	–10.0**
Standing or walking	15.3	20.9	–5.6
Jumping, running, tag or chase games	19.7	13.0	6.7
Climbing, sliding or swinging on a play structure	9.9	12.3	–2.4
Soccer, football or hockey	5.1	6.3	–1.2
Baseball, kickball, basketball or volleyball	20.8	14.6	6.2
Playground games (e.g., four-square, tetherball, etc.)	17.8	9.3	8.4
Percentage of Scans in Which the Main Observed Activity Was Organized by School Staff or a Playworks Coach (Structured Recess Observations)	38.6	17.2	21.3*
Average Number of Organized Games Observed During Recess (Semi-Structured Recess Observations)	4.8	2.7	2.1***
Percentage of Recesses in Which Students Were Engaged in Chase or Non-Playworks Tag Games (Semi-Structured Recess Observations)	40.6	71.0	30.4
Student Reports on Recess Activities			
Percentage of Students Who Report That the Following Happens "Sometimes" or "A Lot":			
Adults help them play games and sports during recess	61.5	45.4	16.1*
Their teacher helps them play games and sports during recess	39.9	30.7	9.2
They talk with friends during recess	90.0	89.9	0.1

Sources: Structured recess observation scans (n = 1,668) and semi-structured recess observations (n = 140) conducted in spring 2011 and spring 2012 (sample sizes may be smaller for some outcomes due to missing responses).

Note: See full table in Appendix 2, Table 6.

* 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.

...Equipment was more likely to be available to students in treatment schools than control schools.

Recess Equipment. Playworks had a positive impact on availability of recess equipment. The scans conducted as part of the structured recess observations showed that equipment was more likely to be available to students in treatment than control schools (Exhibit 3 and Appendix 2, Table 7). Moreover, some specific types of equipment, including bases (for ball games), cones and jump ropes, were more likely to be available in treatment than control schools (Exhibit 3 and Appendix 2, Table 7).

During the semi-structured recess observations, we observed students at treatment schools using the equipment as intended during recess. There was only one instance at one treatment school where equipment was being misused. In contrast, we observed equipment being misused at 9 out of 12 control schools (and 44 percent of control-school recess periods), which we observed leading to unsafe behavior at 8 control schools (and 30 percent of control-school recess periods). Examples of equipment misuse included taking equipment

from other students, holding on to equipment instead of using it to play, throwing balls at other students and using jump ropes as leashes or pull toys instead of for jumping.

EXHIBIT 3.
Impacts on Availability and Condition of Recess Equipment

Outcome	Treatment	Control	Difference
Availability of Equipment			
Percentage of Scans in Which Equipment Was Available for Students to Use	84.4	55.2	29.2***
Percentage of Recess Observations in Which the Following Were Available for Use:			
Balls	92.2	89.6	2.6
Bases (for ball games)	40.6	1.4	39.2***
Cones	84.8	29.8	55.0***
Hula hoops	47.8	40.8	7.0
Jump ropes	86.0	50.7	35.2***
Frisbees	6.5	12.6	–6.1
Goals (e.g., for soccer, football)	23.0	35.1	–12.2
Lines on ground (for four-square, etc.)	90.9	85.9	5.0

Source: Structured recess observations (n = 1,668 and n = 170) conducted in spring 2011 and spring 2012 (the first sample size refers to the number of scans and the second sample size refers to the number of recess observations; multiple scans were taken during each recess observation).

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

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

Student Interactions at Recess. To assess the nature of student interactions at recess, we estimated impacts using data from the structured recess observations and the teacher and student surveys. There was one significant impact on 11 measures of student interactions at recess. Teachers in treatment schools reported significantly better recess behavior and readiness for class than teachers in control schools (Appendix 2, Table 8). This finding was also observed based on the analyses of cohort 1 only (Bleeker et al. 2012). In contrast, we found no significant difference between the treatment and control groups in terms of positive and negative behaviors observed during the structured recess observations. We also found no impact on student reports of their own negative behavior during recess. One potential explanation for this pattern of findings could be that student conflict still occurs with Playworks (so no impacts on student interactions were found based on student survey or recess observation data), but that conflict may be resolved more quickly while students are still at recess so that teachers do not see conflicts returning to their classrooms after recess.

Perceptions of Recess. Perceptions of recess were measured using outcomes constructed from the student and teacher surveys. There were significant impacts on two of four measures of teachers' perceptions of recess. A higher percentage of teachers in treatment schools reported that their students enjoyed adult-organized activities at recess and felt ownership over their activities during recess (Appendix 2, Table 9). There were no significant impacts on seven measures of students' perceptions of recess; that is, the levels of enjoyment at recess reported by students in treatment and control schools did not differ significantly.

Discussion of Key Findings. Playworks had a statistically significant positive effect on several outcomes related to play, physical activity and recess at school and the significant impacts were based on data from a range of sources—accelerometers, structured and semi-structured recess observations, teacher surveys and students surveys.

One key finding was that Playworks had a significant positive impact on students' physical activity during recess. This finding is particularly noteworthy because it is based on accelerometers, which are thought to offer one of the most objective measurements of physical activity in free-living settings outside of the laboratory environment. However, there was no significant difference in physical activity, based on accelerometer data, when we focused on students' activity during periods of the school day other than recess. This is, perhaps, not surprising given that one of the main focuses (if not *the* main focus) of Playworks is schools' recess periods. Therefore, if an impact on physical activity was going to be observed, it would most likely be observed during recess and not during other times of the school day.

We found significant impacts on outcomes related to physical activity from multiple data sources. For example, we found a significant impact on time spent in vigorous activity at recess based on the accelerometer data and we found a significant impact on the percentage of scans in which the main activity students were observed to be engaged was sedentary. These findings suggest, in different ways, that students at treatment schools were engaged in higher levels of physical activity, compared with students at control schools. There are clear differences between the data sources for these two findings, which could explain why they are different. In particular, the accelerometer data were collected from 4th- and 5th-grade students only, whereas the recess observations spanned all grades. Nevertheless, the pattern of findings is consistent with the notion that Playworks led to increased physical activity at recess, at least for some students.

Cost Analysis

As described above, Playworks had a positive impact on outcomes related to play, physical activity and recess. 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.

Summary of Findings

The current evaluation found positive impacts of the Playworks program on multiple measures associated with play, physical activity and recess.[7] In particular, our impact analyses showed the following:

Teachers in treatment schools reported better recess behavior and readiness for class than teachers in control schools.

- Fourth- and 5th-grade students in treatment schools engaged in physical activity during recess that was, on average, more intense than the physical activity engaged in by control students. Students in treatment schools spent significantly more time engaged in vigorous physical activity at recess than students in control schools. These findings are based on data recorded by accelerometers, which offer one of the most objective measures of physical activity in non-laboratory settings.
- The main recess activity in which students were observed to be engaged was less likely to be sedentary in treatment schools, compared with control schools.
- The percentage of recess activities that were organized by school staff (or a Playworks coach in treatment schools) was significantly higher at treatment schools, compared with control schools. A significantly higher percentage of students in treatment schools also reported that adults helped them play games and sports "sometimes" or "a lot" during recess.
- Equipment was made available at a significantly higher percentage of recess periods in treatment schools, compared with control schools. Some types of equipment (jump ropes, cones and bases for ball games) were available at a higher percentage of recess periods in treatment schools.
- Teachers in treatment schools reported better recess behavior and readiness for class than teachers in control schools.
- A higher percentage of teachers in treatment schools, compared with control schools, agreed or strongly agreed that their students enjoyed adult-organized activities at recess and felt ownership over their activities during recess.

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. When it came to play, physical activity and recess, these were the key implementation findings that were observed:

- About three out of every four teachers interviewed at treatment schools commented that Playworks had provided students with increased opportunities for physical activity.
- Students in treatment schools were observed engaging primarily in organized games during recess, such as four-square and tetherball, whereas students in control schools were observed engaging in games that were less organized, such as tag games without clear rules or boundaries.
- Students in treatment schools were observed using the equipment as intended during recess, but students in control schools were observed not using the equipment as intended during 44 percent of control-school recess.

References

ActiLife 5 – User's Manual. Prepared by ActiGraph R&D and Software Departments, 2011. Available at *http://www.actigraphcorp.com/support/*. Accessed September 15, 2011.

Action for Healthy Kids. "Action for Healthy Kids Commitment to Change." Chicago, IL: Action for Healthy Kids, 2008. Available at *http://www.actionforhealthykids.org/resources/files/commitmenttochange.pdf*. Accessed September 13, 2011.

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.

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.

Centers for Disease Control and Prevention. "Promoting Better Health for Young People Through Physical Activity and Sports." Atlanta, GA: U.S. Department of Health and Human Services, 2000.

Edwardson CL and Gorely T. "Epoch Length and Its Effect on Physical Activity Intensity." *Medicine and Science in Sports and Exercise*, 42(5): 928–934, 2010.

Fortson J, James-Burdumy S, Bleeker M, et al. "Impact and Implementation Findings from an Experimental Evaluation of Playworks: Effects on School Climate, Academic Learning, Student Social Skills and Behavior." Report submitted to the Robert Wood Johnson Foundation. Princeton, NJ: Mathematica Policy Research, February 2013.

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.

Holmes RM, Pellegrini AD and Schmidt SL. "The Effects of Different Recess Timing Regimens on Preschoolers' Classroom Attention." *Early Child Development and Care*, 176(7): 735–743, 2006.

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.

Kahan D. "Recess, Extracurricular Activities and Active Classrooms: Means for Increasing Elementary School Students' Physical Activity." *Journal of Physical Education, Recreation and Dance*, 79(2): 26–39, 2008.

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, Castrechini S, Stokes-Guinan K, et al. "Playworks Implementation in 17 Schools Nationwide." Report submitted to the Robert Wood Johnson Foundation. Princeton, NJ: Mathematica Policy Research, February 2013.

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.

McKenzie TL, Marshall SJ, Sallis JF, et al. "Leisure-Time Physical Activity in School Environments: an Observational Study Using SOPLAY." *Preventive Medicine*, 30(1): 70–77, 2000.

Murray R, Ramstetter C, Devore C, et al. "The Crucial Role of Recess in School." *Pediatrics*, 131(1): 183–188, 2013.

Ogden CL, Carroll MD, Curtin LR, et al. "Prevalence of Overweight and Obesity in the United States, 1999–2004." *Journal of the American Medical Association*, 295(13): 1549–1555, 2006.

Pellegrini AD. *Recess: Its Role in Education and Development*. Mahwah, NJ: Erlbaum, 2005.

Pellegrini AD and Bohn C. "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. "Recess Rules: Why the Undervalued Playtime May Be America's Best Investment for Healthy Kids and Healthy Schools Report." Princeton, NJ: RWJF, 2007. Available at *http://www.rwjf.org/files/research/sports4kidsrecessreport.pdf*. Accessed September 13, 2011.

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.

Wechsler H, Devereaux AB, Davis M, et al. "Using the School Environment to Promote Physical Activity and Healthy Eating." *Preventive Medicine*, 31: S121–S137, 2000.

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. Findings on measures of recess experience based on both cohorts are included in this brief.
2. In this study, physical activity at recess is objectively measured using accelerometers. Accelerometers are devices worn on the body (usually around the waist) that record intensity counts based on movement. More intensity counts for a given unit of time indicate more intense physical activity during that unit of time. The accelerometers used for this study also measure the number of steps taken, which, like intensity counts, can be used to assess levels of physical activity engagement.
3. Unlike the previous two impact briefs (Fortson et al. 2013; Bleeker et al. 2012), which presented impact findings and implementation findings separately, in this brief we integrate our discussion of the impact and implementation findings.
4. Impacts on schools potentially combine the effects of Playworks on the behavior of individual students with impacts on the composition of the student body. That is, if implementing Playworks changes which students attend a school (admittedly an unlikely scenario), then the impacts reported in this brief reflect that effect. In Appendix 2, Table 2, we compare the characteristics of students in treatment schools and control schools and find no compositional differences. We therefore conclude that the impacts reported in this study most likely do not reflect compositional effects, but rather only student-level behavior effects (though we acknowledge that impacts including compositional effects can also be of interest).
5. Students wore accelerometers during the entire school day. For this brief, we focus on the accelerometer data recorded during scheduled recess periods since the Playworks intervention focuses primarily on recess.
6. Focus groups were not organized at one of the 17 treatment schools because junior coaches in that school did not get their permission slips signed.
7. We might expect impacts to be larger in schools with stronger implementation. (A school exhibits “strong” implementation of Playworks if recesses are well organized, school staff are knowledgeable about Playworks, principals are willing to schedule regular class game times and school policies support Playworks [London et al. 2013].) 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 13 outcomes with significant findings in Appendix 2, Tables 3–9. Implementation strength was not significantly related to the magnitude of impacts for 12 of the 13 outcomes. For one outcome—average number of organized games observed during recess (Appendix 2, Table 6)—we found evidence suggesting that schools with strong implementation experienced larger positive impacts of Playworks on the number of organized games.

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 2 blocks of matched schools, one had 4 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, 1 of which was a foursome, 3 of which were trios and 8 of which were pairs. For a block of paired schools, 1 school was randomly assigned to the treatment group and 1 school to the control group. For blocks of 3 schools, 2 were randomly assigned to the treatment group and 1 to the control group. For the block of 4 schools, 3 were randomly assigned to the treatment group and 1 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), recess observations 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 the structured recess observations and semi-structured recess observations, we used the same model structure as the one described above for teacher-level data, except we did not include the X_{js} and β terms in the recess-data models. Thus, we estimated the following model:

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

where Y_{js} is the outcome for recess observation 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_{js} is a recess observation-level random error term and γ is the parameter to be estimated.

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). For outcomes based on the student-level accelerometer data, we used the above model but took a weighted averaged of the outcomes for students with two days of accelerometer wear.

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 generalized estimating equations. 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 with some exceptions.[1] For the adjustments, we calculated statistical significance tests based on critical

1 The impacts in Appendix 2, Table 4 are not based on MHT adjustments (the *p*-values shown are unadjusted *p*-values) because each of the outcomes is based on the same piece of information (accelerometer intensity counts). The findings in Exhibit 1 also reflect these considerations for MHT adjustments.

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 to estimate the impacts of teacher- and student-level outcomes and outcomes based on recess observations. For the accelerometer data, sampling weights also accounted for the selection of some students for two days of accelerometer wear. The student and teacher 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. Sampling weights for the structured and semi-structured recess observations were constructed so that equal weight was given to each school. This approach ensures that schools with smaller play areas get the same weight as those with larger play areas, even if there were more scans made during recess observations at the schools with larger play areas (in the case of the structured recess observations).

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, collection of physical activity data via accelerometers, structured observations of recess periods and collection of administrative records. The implementation study included interviews with principals, teachers and Playworks coaches; focus groups with Playworks junior coaches; semi-structured observations of recess; and observations of Playworks class game time. 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.[2]

- **Accelerometer Data Collection.** A total of 1,579 students from 98 4th- and 5th-grade classrooms in 27 study schools wore ActiGraph GT3X accelerometers for 10 minutes or more during their scheduled recess periods[3] and were included in the main accelerometer analysis sample. Accelerometers are recognized as one of the most effective ways to objectively monitor intensity and duration of physical activity outside of research laboratories because of their cost and minimal burden to study participants. We processed the accelerometer data using the ActiLife 5 software package (ActifLife 5 – User's Manual, 2011). We used this package to create measures of vertical movement and the number of steps taken (these measures were used to construct all outcomes based on accelerometer data in Appendix 2, Table 3). In schools with four or fewer 4th- and 5th-grade classrooms, all 4th- and 5th-grade classrooms were asked to participate in the accelerometer data collection. In schools with more than four classrooms, we selected a random sample of four classrooms, balanced across the 4th and 5th grades. Students from one study school did not participate in the accelerometer data collection 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 accelerometer data analysis, leaving 27 schools. From the four (or fewer) classrooms selected, we also randomly selected an additional classroom; students in that classroom wore the accelerometers for an additional school day (two days in total). Parents were required to sign and return

2 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.

3 Each school provided information about the start and end time for recess periods for each grade. We used accelerometer data measured during these scheduled recess periods for analysis.

consent forms before their children were allowed to participate in the accelerometer data collection; 30 percent of parents did not provide parental consent (treatment schools: 29 percent; control schools: 30 percent). The overall response rate for the accelerometer data collection was 66 percent (treatment schools: 66 percent; control schools: 66 percent).

- **Structured Recess Observations.** Mathematica staff conducted six structured recess observations at all 29 study schools during the study year. These observations captured information on recess equipment, student physical activity levels, types of activities and student interactions. During each of the six observations, students were observed in designated zones of the play area. The SOPLAY observation tool (McKenzie et al. 2000) was used to collect information on types of activities and student physical activity levels within each zone. A separate observation tool was used to collect information on student interactions within each zone, indicating whether positive or negative interactions occurred or if conflict-resolution strategies were used. Each observation taken within a zone is referred to as a "scan" in this study brief. Observers were provided with a three-day training that covered operational definitions of recess activities and student behaviors, mapping of the physical recess space and coding conventions. Observers watched the SOPLAY training video, as well as a series of short video clips selected to represent positive and negative behaviors, and took an assessment. Observers also practiced live mapping and coding with children on a playground.

  To assess inter-rater reliability in the field, two certified observers independently conducted at least 33 percent of all recess observations at each site simultaneously. SOPLAY reliability was then assessed between the two observers by aggregating counts of levels of physical activity intensity (sedentary, walking and vigorous) across boys and girls and across all intervals within an observation. The Pearson correlation coefficient between observers was 0.98, 0.98 and 0.97 for sedentary, walking and vigorous activity counts, respectively. Reliability of the student interaction observations was assessed by collapsing the counts into positive and negative behaviors. Counts of negative, positive and conflict-resolution behaviors were aggregated across boys and girls, across scans within an interval and then across intervals within an observation. The Pearson correlation coefficient between observers was 0.70, 0.94 and 0.95 for negative, positive and conflict-resolving counts, respectively.

- **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.

- **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 three 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.

- **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.[4] 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. Focus groups were organized in only 16 out of the 17 treatment schools because junior coaches in one of the treatment schools did not get their permission slips signed.

- **Semi-Structured Recess Observations.** JGC staff conducted semi-structured observations of two days of recesses (all or nearly all recess periods on each day) at each of the 29 treatment and control schools to examine 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. These observations are different from the structured observations that focused on specific zones within the play area and were conducted independently from the structured observations (i.e., a separate team conducted the semi-structured recess observations).

- **Playworks Class Game-Time Observations.** JGC staff observed Playworks class game time in the 17 treatment schools to assess the coaches' relationships with students in smaller groups and to examine teacher and coach interactions, discipline styles and levels of physical activity.

Additional Study Briefs

Three separate study briefs reported findings related to Playworks implementation (London et al. 2013) and the impact of the program on school climate, conflict resolution and aggression, learning and academic performance, youth development and student behavior (Bleeker et al. 2012; Fortson et al. 2013). The implementation brief (London et al. 2013) describes Playworks implementation in the 17 treatment schools, focusing on the

4 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.

implementation of recess, the junior coach program and class game time. London et al. (2013) also describes the contextual factors that influence recess functioning and program implementation quality and reports on school staff and students' perceptions of Playworks.

The other two briefs (Fortson et al. 2013; Bleeker et al. 2012) present impacts of Playworks on school climate, conflict resolution and aggression, learning and academic performance, youth development and student behavior. Bleeker et al. (2012) reports findings from cohort 1 and Fortson et al. (2013) reports findings from cohorts 1 and 2.

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 Who 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 of the Student Survey Sample[a]					
Percentage of Students Who Are Female	2,305 (student survey)	52.5	50.5	2.0	0.75
Percentage of Students Who 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 Who Are the Following Race/ Ethnicity:[b]	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
Student Characteristics of the Accelerometer Sample[c]					
Percentage of Students Who Are Female[d]	1,510 (student survey)	52.4	51.9	0.6	1.00
Percentage of Students Who Are in the Following Grades:	1,579 (administrative records)				
4th		54.5	54.8	–0.3	1.00
5th		45.5	45.2	0.3	1.00
Percentage of Students Who Are the Following Race/ Ethnicity:[e]	1,488 (student survey)				
Black or African American		30.9	29.5	1.4	1.00
Hispanic		35.7	47.0	–11.2	0.82
White		26.3	21.9	4.4	0.62
Asian, Native Hawaiian, or Other Pacific Islander		24.0	14.8	9.3	0.68
American Indian or Alaskan Native		8.4	5.8	2.6	0.74

TABLE 2. (continued)
Characteristics of Students and Teachers in the Study

Characteristic (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Teacher Characteristics					
Percentage of Teachers Who Are Female	297 (teacher survey)	90.1	88.0	2.1	1.00
Percentage of Teachers Who Are Hispanic or Latino	293 (teacher survey)	8.4	14.6	–6.2	0.61
Percentage of Teachers Who Are the Following Race:[f]	273 (teacher survey)				
White		86.2	73.1	13.0**	0.02
African American		8.1	18.6	–10.5**	0.02
Other race[g]		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 are the baseline characteristics for the sample of students who completed the student surveys.

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

c These are the baseline characteristics for the sample of students who wore the accelerometers; the students selected to wear the accelerometers were a subsample of the students selected to complete the student surveys.

d Some students from the accelerometer sample are not included in these results for gender because they did not complete the student survey.

e These percentages can sum to more than 100 because students could report more than one race or ethnicity. Some students from the accelerometer sample are not included in these results for race/ethnicity because they did not complete the student survey.

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

g 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 Students' Physical Activity at Recess (Intensity Counts and Steps)[a]

Outcome	Sample Size (data source)	Treatment	Control	Difference	*p*-Value
Mean Number of Accelerometer Intensity Counts Recorded Per Minute During Recess[b]	1579 (accelerometers)	1312.9	1013.0	299.9*	0.10
Mean Number of Steps Taken Per Minute During Recess[b]	1579 (accelerometers)	29.7	25.4	4.3	0.20

Source: Accelerometer data collected in spring 2011 and spring 2012.

Note: Random assignment block indicators were included as covariates in all impact models. The *p*-values reported in this table account for clustering of students 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 The results in this table are based on accelerometer data. Students wore accelerometers during one or two full school days; the accelerometer data recorded during scheduled recess periods were included for estimating impacts. Students with 10 or more minutes of usable accelerometer data during their recess periods are included in the analysis sample. Any data from time periods during which there was a device malfunction or students were not wearing the accelerometer were considered unusable and not included in the analyses. The accelerometer sample of students includes a subsample of students who wore an accelerometer on two different school days. The accelerometer sample sizes in the table reflect the number of students in the accelerometer sample.

b Accelerometers produce intensity counts and number of steps taken for each second of accelerometer wear time. More intensity counts and more steps taken both suggest that the student is engaged in more intense physical activity.

* 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.
Impacts on Students' Physical Activity at Recess (Time Spent in Activity Groups)[a]

Outcome	Sample Size (data source)	Treatment	Control	Difference	*p*-Value
Mean Percentage of Accelerometer Wear Time During Recess Spent in:[b]	1,579 (accelerometers)				
Sedentary Activity		44.1	50.6	–6.5	0.16
Light Activity		19.6	18.6	1.0	0.25
Moderately Intense Activity		22.7	21.2	1.5	0.46
Vigorously Intense Activity		13.7	9.6	4.1*	0.06

Source: Accelerometer data collected in spring 2011 and spring 2012.

Note: Random assignment block indicators were included as covariates in all impact models. The *p*-values reported in this table account for clustering of students within schools. Multiple hypothesis testing (MHT) adjustments were not used for calculating *p*-values in this table because the outcomes are all constructed from the same accelerometer intensity counts. The treatment mean minus the control mean does not always equal the number shown in the difference column due to rounding.

a The results in this table are based on accelerometer data. Students wore accelerometers during one or two full school days; the accelerometer data recorded during scheduled recess periods were included for estimating impacts. Students with 10 or more minutes of usable accelerometer data during their recess periods are included in the analysis sample. Any data from time periods during which there was a device malfunction or students were not wearing the accelerometer were considered unusable and not included in the analyses. The accelerometer sample of students includes a subsample of students who wore an accelerometer on two different school days. The accelerometer sample sizes in this table reflect the number of students in the accelerometer sample.

b Time spent in the sedentary, light, moderate and vigorous activity groups are determined using the accelerometer intensity counts. Each five-second interval during accelerometer wear time is categorized as sedentary, light, moderate or vigorous if the intensity counts are between 0 and 8, 9 and 74, 75 and 288 or 289 or greater, respectively, for the five-second interval. See Edwardson and Gorely (2010) for more information on converting intensity counts into time-based metrics of physical activity.

* 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 5.
Impacts on Students' Physical Activity at Recess and Participation in Sports and Clubs

Outcome	Sample Size (data source)	Treatment	Control	Difference	*p*-Value
Reports on Physical Activity During Recess					
Percentage of Students Who Participate in Recess Activities That Make Them Sweat and Breathe Hard "Sometimes" or "A Lot"	2,278 (student survey)	76.8	77.1	–0.3	1.00
Percentage of Teachers Who "Agree" or "Strongly Agree" That Their Students Participate in Physical Activities That Make Them Sweat and Breathe Hard During Recess	293 (teacher survey)	76.9	57.9	19.0	0.14
Mean Percentage of Students Who Were Observed Engaging in Moderate or Vigorous Activity During Recess[a]	1,668 (structured recess observation scans)	63.0	67.2	–4.2	0.50
Participation in Sports and Clubs					
Percentage of Students Who Participated in a Sports Team After School or on the Weekend During the School Year	2,272 (student survey)	58.8	51.6	7.2	0.16
Percentage of Students Who Participated in Dance, Gymnastics or Martial Arts Lessons After School or on the Weekend During the School Year	2,274 (student survey)	46.0	41.1	4.9	0.47

Sources: Student and teacher surveys and structured recess observations 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, teachers and recess observations 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 study team used the SOPLAY recess observation instrument to systematically scan up to eight zones of the recess space, during each structured recess observation, to determine physical activity levels of individuals. Within each zone, separate scans were made for girls and boys. Each individual student was coded as engaging in sedentary, moderate or vigorous activity.

* 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 6.
Impacts on Students' Recess Activities

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Observed Recess Activities					
Percentage of Scans in Which the Following Activity Was the Main Activity Observed:[a]	1,668 (structured recess observation scans)				
Sitting, talking or other sedentary activity		6.7	16.8	–10.0**	0.04
Standing or walking		15.3	20.9	–5.6	0.98
Jumping, running, tag or chase games		19.7	13.0	6.7	0.49
Climbing, sliding or swinging on a play structure		9.9	12.3	–2.4	1.00
Soccer, football or hockey		5.1	6.3	–1.2	1.00
Baseball, kickball, basketball or volleyball		20.8	14.6	6.2	0.95
Playground games (e.g., four-square, tetherball, dodge ball, wall ball, Simon says, tug of war)		17.8	9.3	8.4	0.18
Percentage of Scans in Which the Main Observed Activity Was Organized by School Staff or a Playworks Coach[b]	1,668 (structured recess observation scans)	38.6	17.2	21.3*	0.05
Average Number of Organized Games Observed During Recess[c]	140 (semi-structured recess observations)	4.8	2.7	2.1***	0.00
Percentage of Recesses in Which Students Were Engaged in Chase or Non-Playworks Tag Games[c]	118 (semi-structured recess observations)	40.6	71.0	–30.4	0.42
Student Reports on Recess Activities					
Participation in Individual Activities During Recess Scale Score[d]	2,289 (student survey)	0.7	0.8	–0.1	0.74
Participation in Games During Recess Scale Score[e]	2,280 (student survey)	2.1	1.9	0.1	0.11
Percentage of Students Who Report That the Following Happens "Sometimes" or "A Lot":	2,282 (student survey)				
Adults help them play games and sports during recess		61.5	45.4	16.1*	0.05
Their teacher helps them play games and sports during recess		39.9	30.7	9.2	0.25

TABLE 6. (continued)
Impacts on Students' Recess Activities

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
They talk with friends during recess		90.0	89.9	0.1	1.00

Sources: Student surveys and structured and semi-structured recess observations conducted in spring 2011 and spring 2012.

Note: Random assignment block indicators were included as covariates in all impact models. The *p*-values reported in this table account for clustering of students and recess observations 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 study team used the SOPLAY recess observation instrument to systematically scan up to eight zones of the recess space, during each structured recess observation. Observers made two separate scans of each recess zone they were observing—one scan for boys and one scan for girls. During each scan, they recorded the main activity in which students were engaged.

b The study team used the SOPLAY recess observation instrument to systematically scan up to eight zones of the recess space, during each recess structured observation. Observers made two separate scans of each recess zone they were observing—one scan for boys and one scan for girls. During each scan, they recorded whether the main activity in which students were engaged was organized by school staff or, in the case of the treatment schools, school staff or a Playworks coach.

c These outcomes are based on the semi-structured recess observations that were conducted separately from the SOPLAY recess observations, and focused on the overall playground instead of individual zones of the playground.

d The Participation in Individual Activities During Recess Scale averages student responses to four items from the student survey: (D12) "How often do you stand or sit someplace during recess?"; (D14) "How often do you read during recess?"; (D15) "How often do you do schoolwork during recess?"; and (D16) "How often do you play or hang out by yourself during recess?" Responses are coded on a 4-point scale, ranging from 0 (never) to 3 (a lot) for all items. Higher values on the scale indicate greater participation in solo activities during recess. The scale is coded as missing if responses were missing for two or more items.

e The Participation in Games During Recess Scale averages student responses to six items from the student survey: (D17) "How often do you play a game or sport with another student or students during recess?"; (D18) "How often do you stay involved in games during recess?"; (D19) "How often do you feel confident trying a new game during recess?"; (D20) "How often do you invite another student to play a game during recess?"; (D21) "How often do you play a game or sport that an adult has started during recess?"; and (D22) "How often do you play a game or sport that you or another student has started during recess?" Responses are coded on a 4-point scale, ranging from 0 (never) to 3 (a lot) for all items. Higher values on the scale indicate greater participation in organized games during recess. The scale is coded as missing if responses were missing for two 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.

TABLE 7.
Impacts on Availability and Condition of Recess Equipment

Outcome	Sample Size (data source)	Treatment	Control	Difference	*p*-Value
Availability of Equipment					
Percentage of Scans in Which Equipment Was Available for Students to Use[a]	1,668 (structured recess observation scans)	84.4	55.2	29.2***	0.00
Percentage of Recess Observations in Which the Following Were Available for Students to Use:[b]	170 (structured recess observations)				
Balls		92.2	89.6	2.6	1.00
Bases (for ball games)		40.6	1.4	39.2***	0.00
Cones		84.8	29.8	55.0***	0.00
Hula hoops		47.8	40.8	7.0	1.00
Jump ropes		86.0	50.7	35.2***	0.01
Frisbees		6.5	12.6	–6.1	0.99
Goals (e.g., for soccer, football)		23.0	35.1	–12.2	0.72
Lines on ground (for four-square, etc.)		90.9	85.9	5.0	1.00
Condition of Equipment and Recess Space					
Percentage of Schools in Which the Overall Condition of the Structures Available Was "Excellent/Like New"[b]	29 (structured recess observations)	14.8	22.0	–7.3	1.00
Percentage of Recess Observations in Which There Was Little or No Litter in the Recess Space[b]	171 (structured recess observations)	88.5	94.9	–6.4	0.93

Source: Structured recess observations conducted in spring 2011 and spring 2012.

Note: Random assignment block indicators were included as covariates in all impact models. The *p*-values reported in this table account for clustering of recess observations 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 study team used the SOPLAY recess observation instrument to systematically scan up to eight zones of the recess space, during each structured recess observation. Observers made two separate scans of each recess zone they were observing—one scan for boys and one scan for girls. During each scan, they recorded whether equipment was available for students to use.

b As part of structured SOPLAY recess observations, observers assessed the overall recess space in terms of the equipment that was provided, the condition of the climbing and play structures and the amount of litter that was present. Equipment and litter were assessed during each structured recess observation. The condition of the climbing and play structures was assessed once at each school.

TABLE 8.
Impacts on Student Interactions at Recess

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Student and Teacher Reports of Conflict and Behavior During Recess					
Recess Behavior and Readiness for Class Scale Score[a]	294 (teacher survey)	3.9	3.3	0.5***	0.01
Recess Misbehavior/Discipline Scale Score[b]	2,349 (teacher survey, student-specific portion)	0.1	0.2	0.0	0.92
Percentage of Students Who Report That They Do the Following "Sometimes" or "A Lot":	2,282 (student survey)				
Ask an adult to help them solve a conflict during recess		36.8	37.2	–0.4	1.00
Get teased about not being good at games or sports during recess		27.4	28.8	–1.5	1.00
Get into an argument with other students during recess		31.9	33.9	–1.9	1.00
Fight or hit other students during recess		7.9	11.0	–3.0	0.81
Work out problems with other students during recess without fighting		70.2	64.9	5.3	0.57
Say encouraging things to other students during recess		84.3	79.5	4.9	0.29
Observed Behaviors During Recess					
Percentage of Zones in Which One or More Negative Behaviors Were Observed[c]	877 (structured recess observation zones)	17.3	20.5	–3.2	0.98
Percentage of Zones in Which One or More Positive Behaviors Were Observed[c]	874 (structured recess observation zones)	26.1	17.6	8.5	0.54
Percentage of Zones in Which One or More Conflict-Resolution Strategies Were Observed[c, d]	874 (structured recess observation zones)	2.7	0.7	2.0	0.60

Sources: Student and teacher surveys and structured recess observations 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, teachers and recess observations 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 Recess Behavior and Readiness for Class Scale averages teacher responses to five items from the teacher survey: (A5) "My students come back to class ready for learning after recess."; (A6) "My students come back from recess with a good report from the recess supervisor."; (A7) "My students need to be spoken to by the school principal after recess."; (A8) "My students need me to speak with them about their recess behavior."; and (A9) "My students get along with one another and the other classes at recess." Responses are coded on a 5-point scale, ranging from 1 (never) to 5 (always or almost always) for A5, A6 and A9 and from 1 (always or almost always) to 5 (never) for the other two items. Higher values on the scale indicate better recess behavior and readiness for class. The scale is coded as missing if responses were missing for two or more items.

b The Recess Misbehavior/Discipline Scale averages teacher responses to three items from the student-specific portion of the teacher survey: (F13) “How many times in the past 30 days of school have you disciplined this child for misbehaving at recess (e.g., asked this child to ‘sit out,’ miss recess)?”; (F14) “How many times in the past 30 days of school have you sent this child to the principal’s office for misbehaving at recess?”; and (F15) “How many times in the past 30 days of school have you contacted this child’s parents regarding his/her behavior at recess?” 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 study team scanned up to eight zones of the recess space, during each structured recess observation. Within each zone, observers were instructed to count all observed negative or positive student behaviors during two separate scans. Observers also recorded any conflict-resolution strategies that were used by students in the zone. Counts based on the two separate scans within each zone were summed for this analysis.

d When percentages are close to 0 or 100, as they are in this row, it may become more difficult to reject the null hypothesis because the assumption of normality is no longer valid for hypothesis testing. Readers should use caution when interpreting the results provided in this row, given these small percentages.

* 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 9.
Impacts on Students' and Teachers' Perceptions of Recess

Outcome (mean unless otherwise noted)	**Sample Size** (data source)	**Treatment**	**Control**	**Difference**	***p*-Value**
Student Perceptions of Recess, Sports and Games					
Enjoyment of Recess Scale Score[a]	2,284 (student survey)	3.6	3.6	0.0	1.00
Percentage of Students Who Agree "A Little" or "A Lot" That They:	2,287 (student survey)				
Like to play games and sports at recess		90.2	88.6	1.5	0.94
Feel like they can join other kids in a game on the playground		80.8	80.0	0.8	1.00
Feel left out at recess		20.8	23.9	–3.1	0.75
Percentage of Students Who Report That They Do the Following "Sometimes" or "A Lot":	2,285 (student survey)				
Get to play a game that they want during recess		78.5	75.7	2.8	0.94
Have to play a game that adults want them to play during recess		41.3	38.7	2.6	1.00
Games and Sports Self-Concept Scale Score[b]	2,277 (student survey)	1.8	1.7	0.0	0.13
Teacher Perceptions of Student Feelings About Recess					
Percentage of Teachers Who "Agree" or "Strongly Agree" That Their Students:	296 (teacher survey)				
Look forward to recess		97.9	95.5	2.4	0.97
Enjoy adult-organized activities at recess		94.4	69.8	24.6***	0.00
Would be upset about missing recess		95.5	90.0	5.5	0.55
Feel ownership over their activities during recess		76.9	50.5	26.4**	0.03

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 Enjoyment of Recess Scale averages student responses to seven items from the student survey: (D26) "How much do you look forward to recess?"; (D27) "How much do you like recess?"; (D28) "How much would you like to sit out at recess?"; (D29) "How happy do you usually feel at recess?"; (D30) "How would you feel if you had to miss recess?"; (D31) "How would you feel if recess was made longer?"; and (D32) "How would you feel if your school stopped having recess?" Responses are coded on a 4-point scale, ranging from 1 (not at all) to 4 (very much) for D26 and D27, from 1 (very much) to 4 (not at all) for D28, from 1 (not at all happy) to 4 (very happy) for D29 and D31, and from 1 (very happy) to 4 (not at all happy) for D30 and D32. Higher values on the scale indicate more positive views of recess. The scale is coded as missing if responses were missing for two or more items.

b The Games and Sports Self-Concept Scale averages student responses to 11 items from the student survey: (E1) "I am good enough at sports"; (E2) "I do very well at sports"; (E3) "I am better than most other kids my age at sports"; (E4) I like to play sports rather than watch"; (E5) "I would do well at a new sport I haven't tried before"; (E6) "I am less clumsy at sports than most other kids my age"; (E7) "I do very well at playground games"; (E8) "It

does not take me long to learn new playground games"; (E9) "I usually play games on the playground rather than just watch"; (E10) "I am better than most other kids my age at playground games"; and (E11) "I am less clumsy at playground games than most other kids my age." Responses are coded on a 2-point scale. Higher values on the scale indicate a more positive self-concept of ability in games and sports. 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.

Appendix 3
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

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.

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.

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, four reported a cost savings and two 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)

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.

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.

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.

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.

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.

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