What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

April 2013

WWC Review of the Report "Findings from a Randomized Experiment of Playworks: Selected Results from Cohort 1"[1,2]

The findings from this review do not reflect the full body of research evidence on the *Playworks* program.

What is this study about?

The study examined whether *Playworks*, a program that utilizes full-time coaches to provide structured play opportunities during recess and class time, reduces the number of disciplinary referrals in low-income elementary schools.[3]

Study authors randomly assigned 25 schools in five US cities to either a *Playworks* group during the 2010–11 school year or to a comparison group that did not use the *Playworks* program but was eligible to implement the program during the following year.

To estimate the effect of *Playworks* on disciplinary referrals, study authors compared the number of disciplinary referrals reported by school principals over the course of one week during the spring of 2011 in 12 intervention schools to the number of reported referrals in 10 comparison schools.

What did the study find?

The study did not find any statistically significant differences between the *Playworks* schools and comparison schools in the number of disciplinary referrals.

Features of the *Playworks* Program

The *Playworks* program places full-time coaches in low-income elementary schools to provide opportunities for organized play during recess and throughout the school day. As implemented in this study, the program includes three main components:

- *Structured recess activities:* Coaches teach and foster positive play and improved conflict resolution techniques.
- *Class game time:* In coordination with teachers, coaches engage in classroom games and model positive language.
- *Junior coach program:* Fourth- and fifth-grade students, trained in leadership and conflict resolution, act as role models and facilitators during recess.

WWC Rating

The research described in this report meets WWC evidence standards without reservations

Strengths: This study is a well-implemented randomized controlled trial with low attrition at the school level.

Appendix A: Study details

Bleeker, M., James-Burdumy, S., Beyler, N., Dodd, A. H., London, R. A., Westrich, L., Stokes-Guinan, K., & Castrechini, S. (2012). *Findings from a randomized experiment of Playworks: Selected results from cohort 1*. Report prepared for the Robert Wood Johnson Foundation by Mathematica Policy Research and Stanford University's John W. Gardner Center for Youth and Their Communities.

Setting

The study was conducted in 25 schools in five cities across the United States.

Study sample

During the fall of 2010, study authors formed matched pairs or trios of schools within cities on the basis of highest school grade, school size, percentage of major demographic groups, and percentage of students eligible for free or reduced-price lunch. For each pair, one school was randomly assigned to the *Playworks* group and one to the comparison group. For trios, two schools were randomly assigned to the *Playworks* group and one to the comparison group, resulting in 14 *Playworks* schools and 11 comparison schools. One comparison school did not provide discipline referral data; this school and the two *Playworks* schools in its randomization trio were dropped from the analysis, leaving 12 *Playworks* schools and 10 comparison schools in the analysis sample.[4]

Eighty-five percent of the original 14 *Playworks* schools and 82% of the original 11 comparison schools had Title 1 Schoolwide programs. More than 81% of students in all study schools were eligible for free or reduced-price lunch. Among the *Playworks* schools, 44% of students were African American, 22% were Hispanic, 16% were White, 15% were Asian, and less than 1% were Native American. In comparison group schools, 42% of students were African American, 28% were Hispanic, 12% were White, 8% were Asian, and less than 1% were Native American.

Intervention group

At intervention schools, full-time *Playworks* coaches worked with classes during recess and class game time. During recess, the coaches promoted positive behavior by using positive messaging, promoting inclusive behavior, supervising games, and managing conflicts. During classroom game time, the coaches provided similar behavioral supports and also acted as role models for teachers. The program also included "junior coaches," fourth- and fifth-grade students who provided positive support to lower-grade students during recess.

Comparison group

Comparison schools did not receive *Playworks* during the 2010–11 school year, but were eligible to participate in *Playworks* during the following school year.

Outcomes and measurement

Study authors examined student behavior in schools by comparing the number of principal referrals for disciplinary incidents over the course of one week in the spring of 2011 in *Playworks* and comparison schools. In addition to the number of overall disciplinary incidents, the schools reported disciplinary incidents at recess, in class, in another location, for fighting, for profanity, for disrespect, for harassment, for disruption, for another reason, and for multiple reasons. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Playworks coaches were in the schools full-time, supporting students and acting as role models for teachers. Information was not included in the report about whether training was provided for the fourth- and fifth-grade junior coaches.

Reason for review

This study was identified for review by the WWC by receiving significant media attention.

Appendix B: Outcome measures for the external behavior domain

External behavior	
Behavior problems—overall	Number of disciplinary referrals over a one-week period (spring 2011) for any reason, as reported by the principal.
Behavior problems—at recess	Number of disciplinary referrals at recess over a one-week period (spring 2011), as reported by the principal.
Behavior problems—in class	Number of disciplinary referrals in class over a one-week period (spring 2011), as reported by the principal.
Behavior problems—in another location	Number of disciplinary referrals in another location (i.e., not at recess or in class) over a one-week period (spring 2011), as reported by the principal.
Behavior problems—fighting	Number of disciplinary referrals for fighting over a one-week period (spring 2011), as reported by the principal.
Behavior problems—profanity	Number of disciplinary referrals for profanity over a one-week period (spring 2011), as reported by the principal.
Behavior problems—disrespect	Number of disciplinary referrals for disrespect over a one-week period (spring 2011), as reported by the principal.
Behavior problems—harassment	Number of disciplinary referrals for harassment over a one-week period (spring 2011), as reported by the principal.
Behavior problems—disruption	Number of disciplinary referrals for disruption over a one-week period (spring 2011), as reported by the principal.
Behavior problems—another reason	Number of disciplinary referrals for another reason (i.e., not for any previously listed) over a one-week period (spring 2011), as reported by the principal.
Behavior problems—multiple reasons	Number of disciplinary referrals for multiple reasons over a one-week period (spring 2011), as reported by the principal.

Appendix C: Study findings for each domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
External behavior								
Behavior problems—overall	Spring 2011	22 schools	15.10 (12.74)	21.50 (14.30)	6.40	0.46	+18	1.00
Domain average for external behavior						**na**	**na**	**Not statistically Significant**

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. Standard deviations were provided by the authors after an inquiry by the WWC. The effect size is a standardized measure of the effect of an intervention on school-level outcomes, representing the change (measured in standard deviations) in an average school's outcome that can be expected if the school provides the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average school's percentile rank that can be expected if the school provides the intervention. Because these effect sizes were computed at the school level, they, and their accompanying improvement indices, are not comparable to student-level effect sizes, which are typically the focus of WWC reviews; therefore, the WWC did not compute average effect sizes for this study. The WWC currently does not have an approved standard for determining whether cluster-level effect sizes are substantively important. The study is characterized as having an indeterminate effect on school-level disciplinary referrals because the estimated impact was not statistically significant. na = not applicable.

Study Notes: No corrections for clustering or multiple comparisons were needed. The *p*-value presented here was reported in the original study.

Appendix D: Supplemental findings by domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Behavior problems by location								
Behavior problems—at recess	Spring 2011	22 schools	1.90 (2.42)	1.40 (2.27)	–0.50	–0.20	–8	1.00
Behavior problems—in class	Spring 2011	22 schools	10.00 (9.29)	12.50 (10.59)	2.50	0.24	+10	1.00
Behavior problems—in another location	Spring 2011	22 schools	3.20 (3.27)	5.60 (5.42)	2.40	0.53	+20	0.99
Behavior problems by reason								
Behavior problems—fighting	Spring 2011	22 schools	2.40 (2.31)	4.40 (4.77)	2.00	0.53	+20	1.00
Behavior problems—profanity	Spring 2011	22 schools	0.30 (0.65)	1.00 (0.82)	0.70	0.92	+32	0.85
Behavior problems—disrespect	Spring 2011	22 schools	2.80 (2.59)	4.60 (5.34)	1.80	0.43	+16	1.00
Behavior problems—harassment	Spring 2011	22 schools	0.70 (0.98)	1.10 (1.10)	0.40	0.37	+14	1.00
Behavior problems—disruption	Spring 2011	22 schools	3.00 (3.79)	4.00 (4.52)	1.00	0.23	+9	1.00
Behavior problems—another reason	Spring 2011	22 schools	1.50 (2.65)	2.50 (2.32)	1.00	0.38	+15	1.00
Behavior problems—multiple reasons	Spring 2011	22 schools	4.30 (5.63)	3.80 (8.98)	–0.50	–0.07	–3	1.00

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. Standard deviations were provided by the authors after an inquiry by the WWC. The effect size is a standardized measure of the effect of an intervention on school-level outcomes, representing the change (measured in standard deviations) in an average school's outcome that can be expected if the school provides the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average school's percentile rank that can be expected if the school provides the intervention. Because these effect sizes were computed at the school level, they, and their accompanying improvement indices, are not comparable to student-level effect sizes, which are typically the focus of WWC reviews. The WWC currently does not have an approved standard for determining whether cluster-level effect sizes are substantively important. The sums of disciplinary referrals by location and by reason, as published in the report, do not always equal the overall number of referrals.

Study Notes: No corrections for clustering or multiple comparisons were needed. The *p*-values presented here were reported in the original study.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author[s]) to assess whether the study design meets WWC evidence standards. The review reports the WWC's assessment of whether the study meets WWC evidence standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the single study review protocol, version 2.0. A quick review of this study was released in June 2012, and this report is the follow-up review that replaces that initial assessment.

[2] Absence of conflict of interest: This study was conducted by staff from Mathematica Policy Research. Because Mathematica operates the WWC, this study was reviewed by staff from subcontractor organizations.

[3] This WWC review does not include impacts of *Playworks* on student-level outcomes. The study authors examined student-level impacts using a model in which only students who were present at schools in spring of the year of implementation were included; the WWC is considering how to evaluate analyses that use this design. An updated single study review will cover the student-level analyses once this process is complete.

[4] Four additional schools were randomly assigned to the *Playworks* group in the following school year and are not included in this analysis.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, April). *WWC review of the report: Findings from a randomized experiment of Playworks: Selected results from cohort 1.* Retrieved from http://whatworks.ed.gov.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample are spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.