

# Examining the Relationship between Physical Activity and Student Achievement and Behavior: Results from a Three-Year Intervention

Dick M. Carpenter II<sup>1,\*</sup>, Cheryl Kelly<sup>2</sup>, Timothy K. Behrens<sup>3</sup>, Julaine Field<sup>1</sup>,  
Carmen Luna<sup>2</sup>, Elizabeth Tucker<sup>1</sup>

<sup>1</sup>University of Colorado Colorado Springs

<sup>2</sup>Institute for Health Research, Kaiser Permanente Colorado

<sup>3</sup>Northern Arizona University

\*Corresponding author: [dcarpent@uccs.edu](mailto:dcarpent@uccs.edu)

Received September 10, 2018; Revised October 30, 2018; Accepted November 29, 2018

**Abstract** This study examines the relationship between physical activity (PA) levels and academic achievement and student behavior in schools that completed a three-year initiative to increase and improve PA for students. Using difference-in-difference and discontinuity analyses, we find schools that participated in a PA intervention saw greater academic achievement during the three years of PA programming as compared to non-participating schools. The significant effects of the intervention were realized during the first year of programming; subsequent years saw stable achievement but not significant growth. There did not, however, appear to be differences in maladaptive student behavior between schools based on participation in the PA intervention.

**Keywords:** *physical activity, student achievement, student discipline*

**Cite This Article:** Dick M. Carpenter II, Cheryl Kelly, Timothy K. Behrens, Julaine Field, Carmen Luna, and Elizabeth Tucker, "Examining the Relationship between Physical Activity and Student Achievement and Behavior: Results from a Three-Year Intervention." *Journal of Physical Activity Research*, vol. 3, no. 2 (2018): 109-117. doi: 10.12691/jpar-3-2-8.

## 1. Introduction

This study examines the relationship between physical activity (PA) levels and academic achievement and student behavior in U.S. schools that completed a three-year initiative to increase and improve PA opportunities for students. In recent years, school leaders and policymakers in the United States have paid increasing interest to children's levels of PA in school. Although much of the interest is in the relationship between PA and health, the link between PA and other outcomes has also seen growing attention. This recent interest is in contrast to a trend in less school-based PA that began almost 20 years ago.

With the passage of No Child Left Behind (NCLB) legislation in 2002 (P.L. No. 107-110, Section 115) by the U.S. Congress, schools increasingly focused on academic achievement and standardized test scores and placed less value on, among other things, PA through physical education or recess [1,2,3]. Yet, evidence suggests devaluing physical education may be counter-productive with respect to increasing academic achievement, as an increasing number of studies find PA can contribute to improved academic performance. In 2010, the Centers for Disease Control and Prevention (CDC) reviewed the related literature and

found half of the associations between PA and academic achievement to be positive and almost half of the other associations reporting null relationships [4]. Only a small portion of the studies found a negative association between PA and academic achievement. Moreover, a meta-analysis published in 2017 concluded PA improves mathematics, reading, and general composite scores in youth [5].

While evidence suggests providing PA breaks during the school day can improve academic performance, there is less research to show PA decreases maladaptive behavior of students (i.e., behaviors that lead to detentions, suspensions, expulsions, etc.). Several studies [6,7] have found PA breaks during the school day improve on-task classroom behavior (e.g., working, fidgeting). However, there is no research of which we are aware that has assessed the impact of PA during the school day on the rate of more extreme behavior problems. This is a striking omission given the clear need for interventions to reduce aggression, bullying, violence, and other behaviors that disrupt the learning environment and lead to serious disciplinary actions such as suspensions and expulsions. This illustrates yet another potential negative implication from a move away from opportunities for PA in schools.

The post-NCLB trend away from school-based PA is also in contrast to state-level policies about physical education. All states in the United States have policies

that require schools to provide some form of physical education, although policies vary by state in the number of minutes and the type of activity that must be provided. In Colorado-the state in which the present study occurred-House Bill 1069 established a minimum requirement for the amount of time each elementary school must provide students in the form of opportunities to be physically active during the school day (150 minutes a week). Schools were required to comply with this mandate starting in the 2011-12 academic year.

To help schools achieve this, Kaiser Permanente Colorado-a large nonprofit health insurance provider in the United States-funded school districts through their Thriving Schools (TS) initiative to increase opportunities for PA in schools. Spanning school years 2015-2017, the goal of this initiative was for schools to provide at least 30 minutes of PA a day (or 150 minutes a week) through before, during, and after school programming. Funded districts selected priority schools and used the funding to provide professional development for physical education teachers, recess monitors, and classroom teachers and staff about how to increase opportunities for students and staff to be active during school hours (e.g., quality physical education, in-class activity breaks, or movement-based brain breaks) and to be active before or after school (e.g., walking school buses, intramural, or PA club sports). For more information about the intervention, see [8]. The purpose of this study was to assess if schools in the TS initiative demonstrated improved academic performance and reduced maladaptive student behavior compared to schools not participating in the initiative.

The unique contribution of our study is four-fold. First, we examine the results of a PA initiative that was widely implemented across many schools and districts in one state. Many prior studies examine the effects of PA in more limited settings. Second, the analysis considered results from a PA initiative that was implemented over a three-year period; many prior studies considered short program implementations. Third, we examine a unique dependent measure-maladaptive student behavior-that appears to be essentially unexamined. Finally, our design and analyses are unique-a pre-post design with three years of pre-program data and three years of in-program data comparing intervention to non-intervention schools using difference-in-difference and discontinuity analyses.

## 2. Methods

The primary question guiding our research is: Is there a significant difference in student achievement and behavior between schools that implemented a structured PA program and those that did not? Or more specifically, did schools in the Thriving Schools initiative demonstrate improved academic performance and fewer disciplinary actions compared to schools not participating in the PA initiative?

### 2.1. Sample

The unit of analysis for this study is pre K-12 public schools. The sample includes all non-alternative public schools in Colorado for the school years 2012 through

2017. Colorado is a mountain-west state with a population of approximately 5.6 million residents, about 850,000 of which are school-aged children. The state educates its pre Kindergarten through 12<sup>th</sup> grade children through 178 public school districts operating approximately 1,800 schools.

The treatment schools are those that participated in the aforementioned TS program, while control schools did not. We excluded from the sample any alternative public schools, such as correctional or psychiatric facilities, etc. The final sample included 1,726 schools for the achievement data and 1,699 for the discipline data-100 TS schools and 1,626 non-TS schools in achievement and 1,599 non-TS schools in discipline. According to the achievement data, the total number of students served by these schools ranged from 50,316 in 2012 to 52,370 in 2017 for TS schools and 737,305 in 2012 to 818,969 in 2017 for non-TS schools.

Table 1 includes descriptive statistics for the sample using the achievement data, disaggregated by treatment and control schools. Means and standard deviations are derived across years for each group. For example, the mean size of TS schools for the years 2012 through 2017 was almost 509 students. As indicated, the mean enrollments for all schools across all years were similar between groups-approximately 500 students. The percentage of students qualifying for the free and reduced lunch program was even more similar at about 23%. The percentage of minority students served was somewhat greater for non-TS schools, at 43%, although the difference on this indicator was significant ( $p=.03$ ). Thus, at least on the demographic variables used in the analyses below, the schools were quite similar.

Table 1. Sample Descriptive Statistics

	Non-TS		TS		T-test
	mean	sd	mean	sd	p
Enrollment	491.28	409.32	508.50	310.60	.20
Percent FRL	0.23	0.30	0.23	0.29	.96
Percent Minority	0.43	0.27	0.41	0.24	.03

On one of the study's most important variables-mean minutes per day spent in PA-we do not know definitively how much the TS schools differed from non-TS schools. We know TS schools offered a mean of 55.5 minutes of PA per day (SD = 3.0), but because non-TS schools were not part of the study, data collection related to PA was not possible in those schools. Yet, because of the ubiquity of recess and PE classes, we believe it is safe to assume non-TS schools offered some amount of PA to students. And given Colorado law requiring an average of at least 30 minutes per day, a reasonable assumption could be that during the initiative period (2015-2017), TS schools offered, on average, about 25 minutes more PA per day than non-TS schools.

### 2.2. Data and Variables

Almost all of the study data were drawn from statistics reported by the Colorado Department of Education (CDE) on its website. The school discipline data were provided by CDE upon request. The variables used in

the analyses are listed in Table 2, with respective scales of measurement. Details for each are self-evident, save for ELA, Math, and the discipline variables. Like all other states, Colorado tests students annually in English Language Arts (ELA) and Math, and mean scale scores are reported by the state for every school. These scale scores could not be used in their original forms, however, because Colorado began using a new assessment system in 2015, and scores were not scaled for comparisons between the old and new assessment. Therefore, we transformed school scale scores into standard scores (z-scores) for analyses. As for the discipline variables, these represent the percentage of the student body receiving said disciplinary actions.

**Table 2. Analytical Variables, Coding, and Scales of Measurement**

Variable	Scale of Measurement	Coding
Enrollment	Continuous	1...n
Percent FRL	Continuous	0...1
Percent minority	Continuous	0...1
ELA	Continuous	z-scores
Math	Continuous	z-scores
TS school	Nominal	0=non-TS, 1=TS
Year	Continuous	2012 through 2017
Pre/post	Nominal	0=pre TS, 1=TS
Time	Continuous	0...5
Intercept	Nominal	0=pre TS, 1=TS
Slope	Continuous	0=2012 to 2014, 1=2015, 2=2016, 3=2017
Removed from class	Continuous	Percentage
In school suspension	Continuous	Percentage
Out of school suspension	Continuous	Percentage
Expulsions	Continuous	Percentage
Law enforcement referral	Continuous	Percentage
Other actions	Continuous	Percentage
Unduplicated count	Continuous	Percentage

We included enrollment, percent free and reduced lunch (FRL)-a measure of family poverty-and percent racial/ethnic minority because of previous research demonstrating significant relationships between these variables and student/school achievement or student behavior [9-14]. Therefore, these variables may confound the relationship between TS participation and achievement or behavior. Moreover, as discussed above, the TS intervention was not randomly assigned, making our analyses quasi-experimental. Even though we use strong analytical methods that help overcome omitted variable bias from time invariant variables, these three variables are time varying and can be included in the models below to control for their effects and improve model estimates.

## 2.3. Analysis

Our study used two primary analyses to answer the research question-difference-in-difference analysis and discontinuity analysis.

### 2.3.1. Difference-in-Difference

Using three years of data pre TS and three years during TS enabled us to examine if the TS schools reported

a significant difference in pre to post achievement and disciplinary actions between periods as compared to the non-TS schools. As such the general model took the form:

$$Y_{it} = \left[ \begin{array}{l} \alpha_i + \beta_1(enroll_{it}) + \beta_2(percFRL_{it}) \\ + \beta_3(percminority_{it}) + \beta_4(TSschool_{it}) \\ + \beta_5(pre/post_{it}) \\ + \beta_6(TSschool * pre/post_{it}) + \delta + \varepsilon_{it} \end{array} \right]$$

where  $\delta$  = year fixed effects, Y = ELA, Math, Removed from class, In school suspension, Out of school suspension, Expulsion, Law enforcement referral, Other actions, or Unduplicated count and all other variables are defined as above.

We ran this model with and without school fixed effects; results from both iterations are reported below. Note that this model estimates average effects when comparing the pre to post periods. That is, the outcome measures are averaged in the pre and post periods separately and then compared. This model does not measure change within the pre or post periods.

### 2.3.2. Discontinuity Analysis

Unlike the difference-in-difference analysis, the discontinuity analysis enables us to measure if the slopes in the pre and post periods differ between TS and non-TS schools. This model takes the form:

$$Y_{it} = \left[ \begin{array}{l} \alpha_i + \beta_1(enroll_{it}) + \beta_2(percFRL_{it}) \\ + \beta_3(percminority_{it}) + \beta_4(TSschool_{it}) \\ + \beta_5(time_{it}) + \beta_6(intercept_{it}) \\ + \beta_7(slope_{it}) + \beta_8(TSschool * time_{it}) \\ + \beta_9(TSschool * intercept_{it}) \\ + \beta_{10}(TSschool * slope_{it}) + \delta + \varepsilon_{it} \end{array} \right]$$

Given the coding described in Table 2, and consistent with Singer and Willett [15], this model examines the effect of the intervention at both the time of its introduction (the TSschool\*intercept term) and also the effect of the intervention in the years after the intervention (the TSschool\*slope term). If, for the TS schools compared to the non-TS schools, the intervention results in an immediate change in test scores or disciplinary actions or a change over time after the introduction of the PA program, the respective regression terms- $\beta_9$  or  $\beta_{10}$ -will be significant. As with the difference-in-difference model, we ran this model with and without school fixed effects; results from both iterations are reported below.

## 3. Results

### 3.1. Student Achievement

We begin by presenting the achievement trends disaggregated by TS and non-TS schools. As illustrated in Figure 1 and Figure 2, trends were quite similar in ELA

and math-TS schools showed achievement levels less than those of non-TS school in the pre-intervention period. In the first year of implementation TS schools saw sharp increases in achievement, followed by stable achievement levels in subsequent years. The trends of non-TS schools showed generally stable achievement levels across all years of the study, with no sharp increases or decreases in achievement. Achievement gaps between TS and non-TS schools did not close entirely during the PA implementation period, but they did significantly narrow during that time.

Statistical analyses largely confirm what is displayed in Figure 1 and Figure 2. First, we report results from the difference-in-difference analysis. Table 3 includes findings from ELA and math analyses. For each test type, Model 1 reports results without school fixed effects, and Model 2 includes school fixed effects. As indicated, the differences between models are trivial.

For both ELA and math, the variable of primary interest is the TSXprepost interaction. This indicates if the pre-post differences evident in Figure 1 and Figure 2 were greater for TS schools as compared to non-TS schools. Results for both ELA and math show there are significant differences between TS and non-TS schools. Moreover, the positive coefficients indicate the pre-post difference for TS schools were greater than those in non-TS schools, confirming what is plainly evident in the figures. Thus, TS schools saw, on average, significantly greater growth in ELA and math scores between the three years before the implementation of the TS PA programs and the three years of TS implementation.

Turning to the discontinuity analysis, Table 4 includes results for ELA and math. As in Table 3, two models are reported in Table 4-Model 1 without school fixed effects and Model 2 with school fixed effects. And as in Table 3, the differences in results based in fixed effects are trivial.

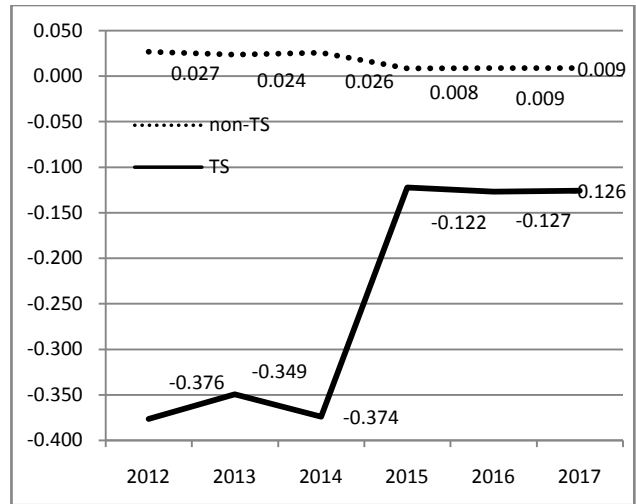


Figure 1. ELA Trends Pre (2012-2014) and Post (2015-2017) TS

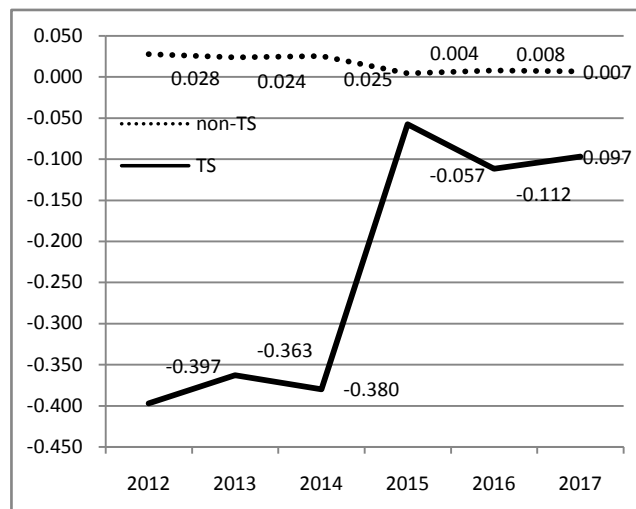


Figure 2. Math Trends Pre (2012-2014) and Post (2015-2017) TS

Table 3. Difference-in-Difference Results, Student Achievement

	ELA						
	Model 1			Model 2			
	Coef.	se	p	Coef.	se	p	
enroll	0.000	0.000	0.000	0.000	0.000	0.013	
firl	0.001	0.001	0.097	0.004	0.001	0.000	
pctminor	-2.225	0.067	0.000	-2.146	0.346	0.000	
TSparticipant	-0.455	0.068	0.000				
prepost	0.090	0.046	0.051				
TSXprepost	0.257	0.096	0.007	0.251	0.096	0.009	
_cons	0.715	0.044	0.000	0.987	0.178	0.000	
School fixed effects				x	x	x	
	Math						
	Model 1			Model 2			
	Coef.	se	p	Coef.	se	p	
enroll	0.001	0.000	0.000	0.000	0.000	0.013	
firl	0.001	0.001	0.070	0.004	0.001	0.000	
pctminor	-2.155	0.067	0.000	-1.870	0.368	0.000	
TSparticipant	-0.468	0.067	0.000				
prepost	0.105	0.047	0.027	0.377	0.049	0.000	
TSXprepost	0.288	0.087	0.001	0.271	0.085	0.001	
_cons	0.616	0.044	0.000	0.800	0.184	0.000	
School fixed effects				x	x	x	

**Table 4. Discontinuity Results, Student Achievement**

	ELA					
	Model 1			Model 2		
	Coef.	se	p	Coef.	se	p
enroll	0.000	0.000	0.000	0.000	0.000	0.013
frl	0.001	0.001	0.098	0.004	0.001	0.000
pctminor	-2.225	0.067	0.000	-2.146	0.346	0.000
TSparticipant	-0.451	0.068	0.000			
newintercept	0.064	0.045	0.157	0.163	0.046	0.000
slope	-0.012	0.008	0.156	-0.006	0.008	0.459
time	0.013	0.004	0.003	0.015	0.005	0.001
TSXintercept	0.266	0.097	0.006	0.267	0.095	0.005
TSXslope	0.001	0.025	0.957	0.004	0.024	0.875
TSXtime	-0.004	0.013	0.777	-0.008	0.014	0.560
_cons	0.714	0.044	0.000	0.985	0.178	0.000
School fixed effects				x	x	x

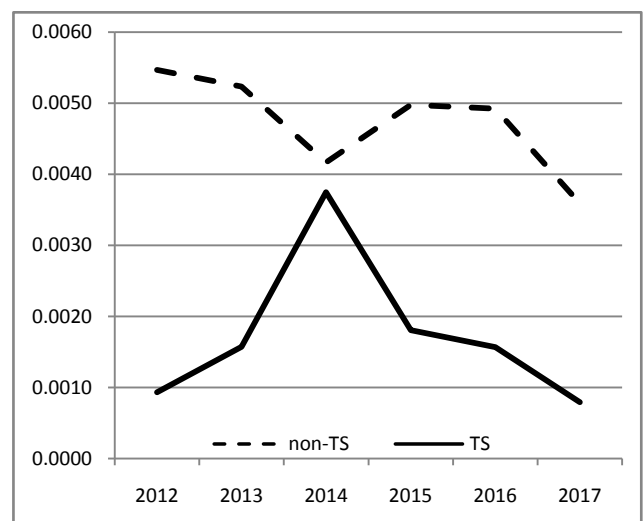
	Math					
	Model 1			Model 2		
	Coef.	se	p	Coef.	se	p
enroll	0.001	0.000	0.000	0.000	0.000	0.015
frl	0.001	0.001	0.074	0.004	0.001	0.000
pctminor	-2.154	0.067	0.000	-1.862	0.369	0.000
TSparticipant	-0.474	0.067	0.000			
newintercept	0.223	0.050	0.000	0.384	0.051	0.000
slope	-0.061	0.010	0.000	-0.063	0.010	0.000
time	0.010	0.004	0.018	0.010	0.004	0.025
TSXintercept	0.316	0.089	0.000	0.301	0.087	0.001
TSXslope	-0.023	0.030	0.456	-0.019	0.031	0.541
TSXtime	0.006	0.013	0.633	0.003	0.013	0.804
_cons	0.616	0.044	0.000	0.790	0.184	0.000
School fixed effects				x	x	x

In this analysis, there are two variables of interest-TSXintercept and TSXslope. For both ELA and math, results for TSXintercept are statistically significant, and the coefficients are positive. This means at the end of the first year of the PA intervention (2015), TS schools saw an immediate and significant increase in ELA and math achievement compared to non-TS schools, which is demonstrated in the figures. The TS schools did not, however, see a comparatively greater increase in achievement in the years following the implementation of the PA programs. In addition to the trend lines in the figures, this is also indicated by the non-significant TSXslope interaction variables for both ELA and math. It is also important to note that while TS schools did not see comparatively greater achievement growth after the implementation of PA, they also did not see a decrease in achievement. Thus, PA programs appeared to produce immediate effects, which then remained stable over subsequent years.

**3.2. Student Behavior/Disciplinary Actions**

Unlike with student achievement, the relationship between PA and disciplinary actions was not significant in either model. As with achievement, we begin with graphical presentations of the trends for the seven discipline types. Figure 3, Figure 4, Figure 5, Figure 6, Figure 7, Figure 8, and Figure 9 illustrate that with the exception of Removal from Class, the trend lines for TS and non-TS schools

generally follow similar directions. More specifically, TS schools do not show divergence from non-TS schools following the implementation of PA programs in 2015. As for Removal from Class, although the TS trend differs from non-TS, there does not appear to be a divergence in the trends that could reasonably be associated with the implementation of PA programs from 2015 through 2017.



**Figure 3.** Removal from Class Trends Pre (2012-2014) and Post (2015-2017) TS

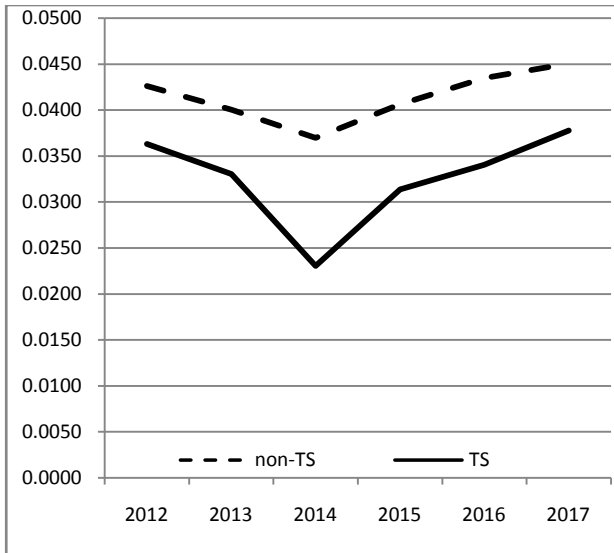


Figure 4. In School Suspension Trends Pre (2012-2014) and Post (2015-2017) TS

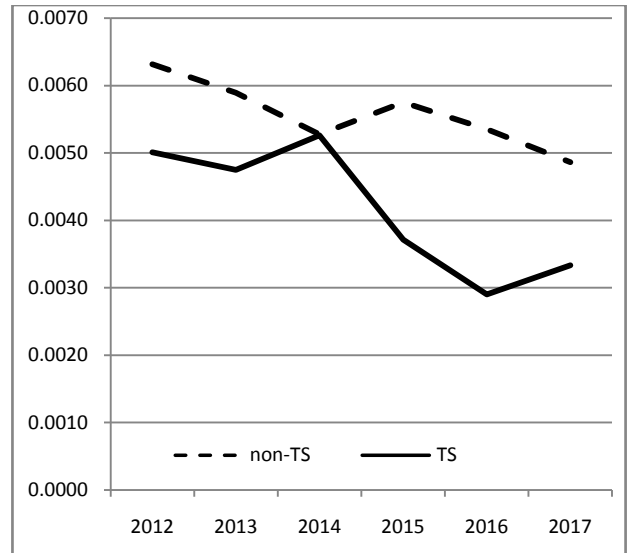


Figure 7. Law Enforcement Referrals Trends Pre (2012-2014) and Post (2015-2017) TS

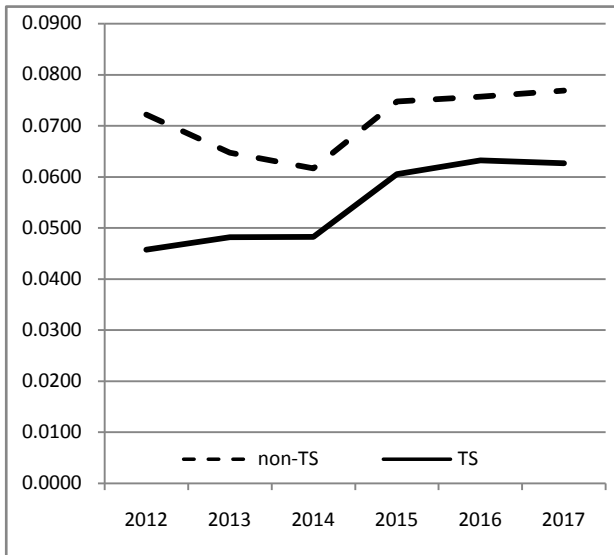


Figure 5. Out of School Suspension Trends Pre (2012-2014) and Post (2015-2017) TS

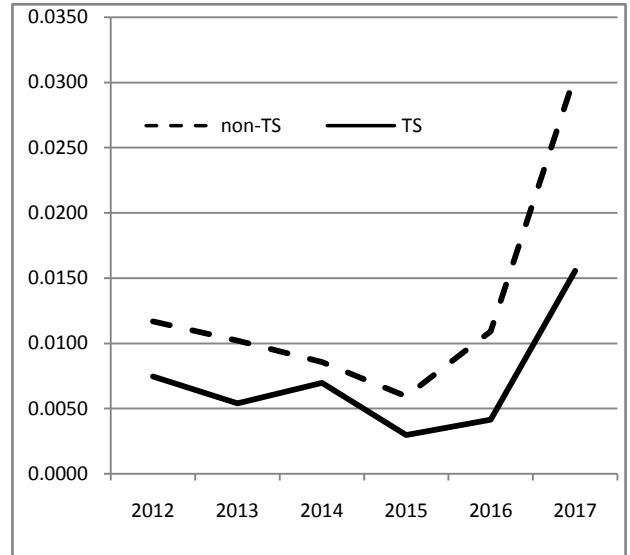


Figure 8. Other Actions Trends Pre (2012-2014) and Post (2015-2017) TS

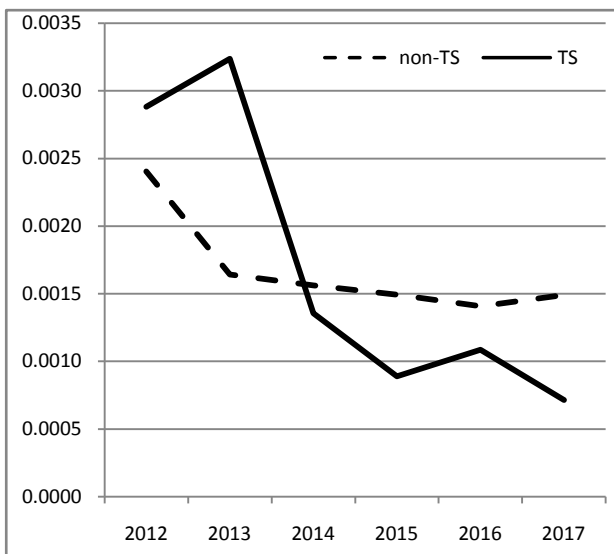


Figure 6. Expulsions Trends Pre (2012-2014) and Post (2015-2017) TS

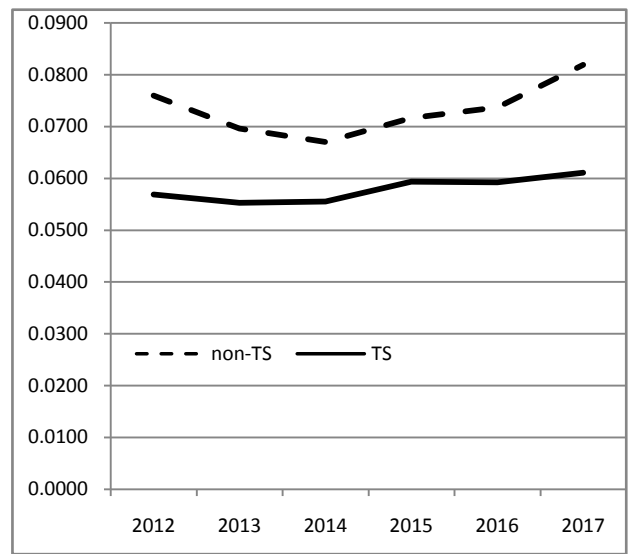


Figure 9. Unduplicated Totals Trends Pre (2012-2014) and Post (2015-2017) TS



When subjected to difference-in-difference analysis, results for the disciplinary actions pre and post do not appear to differ between TS and non-TS schools. Table 5 includes primary results for all seven discipline measures (covariates are omitted for space but are available from the authors), and the interaction variable for all outcome variables indicates non-significance for all. This was true whether or not models included school fixed effects.

**Table 5. Difference-in-Difference Results, Student Behavior/Disciplinary Actions**

	Coef.	se	Coef.	se
<b>Removed from Class</b>				
TSpaticipant	(omitted)		-0.003	0.001
prepost	-0.001	0.002	-0.002	0.002
TSXprepost	-0.001	0.001	-0.001	0.001
<b>In school Suspension</b>				
TSpaticipant	(omitted)		-0.007	0.006
prepost	-0.003	0.003	0.003	0.003
TSXprepost	0.002	0.006	0.001	0.006
<b>Out of school Suspension</b>				
TSpaticipant	(omitted)		-0.016	0.006*
prepost	-0.006	0.003	-0.006	0.003*
TSXprepost	0.008	0.005	0.008	0.005
<b>Expulsions</b>				
TSpaticipant	(omitted)		0.001	0.001
prepost	-0.001	0.000*	-0.001	0.000*
TSXprepost	-0.001	0.001	-0.001	0.001
<b>Law Enforcement Referral</b>				
TSpaticipant	(omitted)		0.00	0.002
prepost	-0.001	0.001*	-0.002	0.000*
TSXprepost	-0.001	0.001	-0.001	0.001
<b>Other Actions</b>				
TSpaticipant	(omitted)		-0.005	0.003
prepost	-0.003	0.003	0.020	0.003*
TSXprepost	-0.004	0.003	-0.004	0.003
<b>Unduplicated Total</b>				
TSpaticipant	(omitted)		-0.012	0.007
prepost	-0.005	0.002	0.003	0.002
TSXprepost	0.000	0.005	0.000	0.005
School fixed effects	x	x		

\*p < .05.

Results from the discontinuity analyses tell much the same story. The variables of specific interest include the interactions of TSpartXnewinterc and TSpartXslope. Across all models, these two variables are consistently not significant, meaning trends before and after the implementation of PA programs did not differ between TS and non-TS schools. Again, this was true with and without school fixed effects.

**Table 6. Discontinuity Results, Student Behavior/Disciplinary Actions**

	Coef.	se	Coef.	se
<b>Removed from Class</b>				
TSpaticipant	(omitted)		-0.005	0.001*
newintercept	0.000	0.002	0.002	0.002
slope	0.000	0.001	0.000	0.001
time	0.000	0.001	-0.001	0.001
TSpartXnewinterc	-0.005	0.002	-0.005	0.002
TSpartXslope	-0.001	0.002	-0.001	0.001
TSpartXtime	0.002	0.001	0.002	0.001
<b>In school Suspension</b>				
TSpaticipant	(omitted)		-0.003	0.009
newintercept	0.001	0.004	0.003	0.004
slope	0.005	0.002*	0.005	0.002*
time	-0.003	0.001*	-0.003	0.002*
TSpartXnewinterc	0.003	0.006	0.002	0.006
TSpartXslope	0.005	0.004	0.005	0.004
TSpartXtime	-0.004	0.004	-0.004	0.004
<b>Out of school Suspension</b>				
TSpaticipant	(omitted)		-0.023	0.008*
newintercept	0.005	0.004	0.005	0.004
slope	0.006	0.002*	0.006	0.002*
time	-0.005	0.001*	-0.006	0.001*
TSpartXnewinterc	-0.001	0.007	-0.001	0.007
TSpartXslope	-0.006	0.005	-0.006	0.005
TSpartXtime	0.007	0.003*	0.007	0.003*
<b>Expulsions</b>				
TSpaticipant	(omitted)		0.001	0.001
newintercept	0.000	0.000	0.000	0.000
slope	0.001	0.000*	0.001	0.000*
time	0.000	0.000*	0.000	0.000*
TSpartXnewinterc	0.001	0.001	0.001	0.000
TSpartXslope	0.000	0.000	0.000	0.000
TSpartXtime	0.000	0.000	0.000	0.000
<b>Law Enforcement Referral</b>				
TSpaticipant	(omitted)		-0.001	0.002
newintercept	0.000	0.001	0.000	0.001
slope	0.000	0.000	0.000	0.000
time	-0.001	0.000*	-0.001	0.000*
TSpartXnewinterc	-0.002	0.001	-0.002	0.001
TSpartXslope	-0.001	0.001	-0.001	0.001
TSpartXtime	0.001	0.001	0.001	0.001
<b>Other Actions</b>				
TSpaticipant	(omitted)		-0.006	0.004
newintercept	-0.017	0.003*	-0.016	0.003*
slope	0.014	0.002*	0.014	0.002*
time	-0.001	0.001	-0.002	0.001
TSpartXnewinterc	0.005	0.006	0.005	0.006
TSpartXslope	-0.006	0.004	-0.007	0.004
TSpartXtime	0.001	0.003	0.001	0.003
<b>Unduplicated Total</b>				
TSpaticipant	(omitted)		-0.016	0.008*
newintercept	-0.002	0.003	-0.001	0.003
slope	0.009	0.001*	0.009	0.001*
time	-0.005	0.001*	-0.005	0.001*
TSpartXnewinterc	0.002	0.006	0.002	0.006
TSpartXslope	-0.007	0.004	-0.007	0.004
TSpartXtime	0.004	0.003	0.004	0.003
School fixed effects	x	x		

\*p < .05.

## 4. Discussion and Conclusion

This study examined the relationship between PA levels, academic achievement, and student behavior, as represented by disciplinary actions, in schools that completed a three-year initiative to increase and improve physical activities for students. Results from difference-in-difference and discontinuity analyses indicated schools that implemented the PA programs saw greater achievement in the years of PA implementation compared to schools that did not participate in the initiative. There were no differences, however, in any of the behavioral measures.

Such results contribute to the growing body of evidence suggesting a positive relationship between PA and achievement. What made these results particularly important in their contribution to the related literature were the causal estimates generated through difference-in-difference and discontinuity analyses. Because random assignment is rare in studies of PA programs, most studies can report only correlational findings. But the use of more rigorous econometrics with panel data enabled us to discuss effects that are more akin to causal estimates.

What made the results for TS schools even more striking was the sharp increase in achievement in the first year of PA program implementation (2015) and that it occurred consistently in both reading and math. Recall that 2015 was also the year when Colorado implemented a new achievement test for its public schools. While it is conceivable a notable difference in test scores could have simply been an artifact of new tests, we would expect to see this effect present in all schools, not just TS schools. Yet, the fact that TS schools saw an increase in test scores in the first year of new tests and PA implementation, while non-TS schools did not, strongly suggests a real effect from the PA programs.

But what is to be made of the significant effects in year one of the PA programming but the lack of achievement growth in subsequent years? The effects of any intervention can, of course, take several forms—an immediate effect with no subsequent increase or decrease in an outcome, an immediate effect with a subsequent increase or decrease, or a gradual increase or decrease over time but no immediate effect. The prior research to date in this space has been dominated by examinations of short-term outcomes, so there is little evidence or theory to suggest which of the three types of effects would be expected. Our finding of immediate effect and then no subsequent increase or decrease may reflect a ceiling on the effects of PA on achievement. Or, like many interventions, the introduction of a new PA program may have a ‘novelty effect’ that creates a new baseline but does not result in continued growth in the outcome. Whatever the reason, it is a unique finding in this literature and worthy of additional research.

By way of implications, these findings suggest schools that reduce or eliminate PA opportunities in their schools to increase academic ‘seat time,’ may be undermining their efforts to increase student achievement. Instead, maintaining or even increasing PA may assist in focusing student learning and increasing achievement levels. Moreover, we note that the TS schools in this study reported offering PA at a frequency close to the levels

recommended by the Physical Activity Guidelines for America (i.e., 60 minutes per day). TS schools also offered a variety of PA opportunities, including PE and recess, before and after school programs, and in-class activities. This was consistent with the CDCs recommendation of a Comprehensive School Physical Activity Program (CSPAP) that includes PA before, during, and after school [16]. For schools considering the implementation of PA programming for the purpose, in part, of helping to increase student achievement, these results suggest the Physical Activity Guidelines for America and the CDC’s CSPAP recommendations are a fine place to start.

Turning to the findings concerning disciplinary actions, it may be tempting to conclude PA is not an efficacious intervention to address negative student behavior. We caution, however, against drawing strong conclusions based on this single study. As mentioned above, the literature linking PA to student behavior is essentially non-existent, and among the few related articles, the outcomes of interest are on-task classroom behavior [6,7]. Whether students fidget in their seats is hardly the same thing as the type of behavior that leads to suspension, expulsion, or referral to law enforcement. Additionally, increased PA alone does not address the multitude of psychological and emotional needs that are often precursors to students’ externalizing behaviors. Our analysis makes an important, unique empirical contribution to understanding the relationship between PA and dysfunctional student behavior. Like any topic, a consensus of research findings is important before discussing implications and drawing firm conclusions. We hope others will pursue studies of their own to consider this important question.

## Funding

This work was supported by the Kaiser Foundation Health Plan of Colorado under grant number RNG200313.

## Conflict of Interest

The authors declare that they have no conflict of interest.

## References

- [1] Howie, E. K. and Pate, R. "Physical activity and academic achievement in children: A historical perspective", *Journal of Sport and Health Science*, 1 (3), 160-169. 2012.
- [2] National Association for Sport and Physical Education and American Heart Association, *Shape of the nation report: Status of physical education in the USA*, Reston, VA, 2006.
- [3] Kohl, H. W. and Cook, H. D. (Eds.). *Educating the student body: Taking physical activity and physical education to school*, National Academies Press, Washington, DC: 2013.
- [4] Centers for Disease Control and Prevention, *The association between school based physical activity, including physical education, and academic performance*, US Department of Health and Human Services, Atlanta, GA, 2010.
- [5] Álvarez-Bueno, C., Pesce, C., Cavero-Redondo, I., Sánchez-López, M., Garrido-Miguel, M. and Martínez-Vizcaíno, V. "Academic achievement and physical activity: A meta-analysis", *Pediatrics*, 140 (6). e20171498. 2017.



- [6] Jarrett, O., Maxwell, D., Dickerson, C., Hoge, P., Davies, G. and Yetley, A. "Impact of recess on classroom behavior: Group effects and individual differences", *Journal of Educational Research*, 92 (2), 121-126. 1998.
- [7] Mahar, M., Murphy, S., Rowe, D., Golden, J., Shields, A. T. and Raedeke, T. D. "Effects of a classroom-based program on physical activity and on-task behavior", *Medicine and Science in Sports and Exercise*, 38 (12), 2086-2094. 2006.
- [8] Kelly, C., Carpenter, D. M., Tucker, E., Luna, C., Donovan, J., and Behrens, T. "A method for evaluating school-based physical activity programming", *Preventing Chronic Disease*, 14, 160607.
- [9] Borman, G. D. and Overman, L. T. (2004). "Academic resilience in mathematics among poor and minority students", *Elementary School Journal*, 104 (3), 177-195. 2004.
- [10] Dell'Angelo, T. "The power of perception: Mediating the impact of poverty on student achievement", *Education and Urban Society*, 48 (3), 245-261. 2016.
- [11] Egalite, A. J., & Kisida, B. (2016). School size and student achievement: A longitudinal analysis. *School Effectiveness and School Improvement*, 27(3), 406-417.
- [12] Fowler, W. J., & Walbert, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis*, 13(2), 189-202.
- [13] Skiba, R. J. (2001). When is disproportionality discrimination? The overrepresentation of black students in school suspension. In W. Ayers, R. Ayers & B. Dohrn (Eds.), *Zero Tolerance: Resisting the Drive for Punishment* (pp. 176-187). New York, NY: New Press.
- [14] Balfanz, R., Byrnes, V., & Fox, J. (2014). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. *Journal of Applied Research on Children: Informing Policy for Children at Risk*, 5(2), article 13.
- [15] Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis*. Oxford, UK: Oxford University Press.
- [16] Centers for Disease Control and Prevention. (2015, September 25, 2015). Comprehensive school physical activity program (CSPAP). Retrieved May 17, 2018, from <https://www.cdc.gov/healthyschools/physicalactivity/cspap.htm>.