Preventive Medicine Reports 3 (2016) 196–202

Contents lists available at ScienceDirect

Preventive Medicine Reports

journal homepage: http://ees.elsevier.com/pmedr

Before-school running/walking club: Effects on student on-task behavior

Michalis Stylianou [a,*], Pamela Hodges Kulinna [b], Hans van der Mars [b], Matthew T. Mahar [c], Marc A. Adams [d], Eric Amazeen [e]

[a] School of Human Movement and Nutrition Sciences, The University of Queensland, St Lucia, Queensland 4072, Australia
[b] Mary Lou Fulton Teachers College, Arizona State University, Mesa, AZ 85212, USA
[c] Department of Kinesiology, East Carolina University, Greenville, NC 27858, USA
[d] School of Nutrition and Health Promotion, Arizona State University, Phoenix, AZ 85004, USA
[e] Department of Psychology, Arizona State University, Tempe, AZ 85281, USA

ARTICLE INFO

Available online 1 February 2016

Keywords:
Physical activity program
Classroom behavior
Elementary school

ABSTRACT

Before-school programs provide a good opportunity for children to engage in physical activity (PA) as well as improve their readiness to learn. The purpose of this study was to examine the effect of a before-school running/walking club on elementary school children's on-task behavior. The study employed a two-phase experimental design with an initial baseline phase followed by an alternating treatments phase, and was first conducted at a private school (School A) and subsequently replicated at a public school (School B). Participants were third and fourth grade children from two schools in the Southwestern U.S. who participated in a before-school running/walking club that met two times each week (School A: 20 min; School B: 15 min) during the 2013/2014 academic year. Participation in the program was monitored using pedometers and on-task behavior was assessed through direct observation. Data analyses included visual analysis, Tau-U index, and multilevel modeling. Results from all analyses indicated that on-task behavior was significantly higher on days the children attended the before-school program than on days they did not. According to multilevel modeling results, mean differences and effect sizes were: School A = 15.78%, *pseudo*-R^2 = .34 [strong effect]; School B = 14.26%, *pseudo*-R^2 = .22 [moderate effect]. Results provide evidence for the positive impact of before-school PA programs on children's classroom behavior and readiness to learn. Such programs do not take time away from academics and may be an attractive option for schools. Results also have implications for the structure of children's school day and the scheduling of PA opportunities.

Introduction

Physical activity (PA) is associated with extensive health benefits (Janssen and LeBlanc, 2010), but a large proportion of American youth do not meet the national PA guidelines (National Physical Activity Plan Alliance, 2014; Troiano et al., 2008). Schools have been identified as primary sites for PA promotion (IOM, 2013), but most of the time youth spend in school is sedentary and PA opportunities in U.S. schools have decreased as a result of the current economic conditions and the heavy emphasis on improving academic performance (CEP, 2008; NASPE and AHA, 2012).

Beyond the health benefits of PA, school-based PA opportunities may also improve students' classroom behavior, cognition, and academic achievement (CDC, 2010; Mahar, 2011; Sibley and Ethier, 2003). Research in this area "is needed to justify the incorporation of physical activity in school settings, especially to teachers and administrators" (Mahar et al., 2006, p. 2086). Indeed, even when teachers and administrators are aware of the health benefits of PA, they may be hesitant to introduce/increase PA opportunities throughout the school day because of pressures to improve academic performance (CEP, 2008; Cothran et al., 2010). Demonstrating the cognitive, academic, and classroom behavior benefits of PA may alleviate some of their concerns and lead to additional PA opportunities throughout the school day.

A possible co-benefit of school PA participation is classroom behavior improvement. Classroom behavior involves various behaviors that may impact student adjustment (e.g., attitude towards school; relationships with students and teachers) and academic performance (CDC, 2010), such as time on/off-task. Acute bouts of PA, including recess (Barros et al., 2009; Jarrett et al., 1998; Pellegrini et al., 1995; Ridgway et al., 2003), classroom PA (Goh et al., 2012; Mahar et al., 2006), and before-school PA (Mahar et al., 2011), have been shown to positively impact different types of classroom behavior.

* Corresponding author.
E-mail addresses: m.stylianou@uq.edu.au (M. Stylianou), pkulinna@asu.edu (P.H. Kulinna), Hans.vanderMars@asu.edu (H. van der Mars), MAHARM@ecu.edu (M.T. Mahar), Marc.Adams@asu.edu (M.A. Adams), Eric.Amazeen@asu.edu (E. Amazeen).

http://dx.doi.org/10.1016/j.pmedr.2016.01.010

Although the mechanism underlying the relationship between PA and classroom behavior is unclear, there is evidence that acute bouts of PA positively influence students' cognitive processes, and particularly executive functioning (e.g., Hillman et al., 2009), which involves processes that make it possible to stay focused (Diamond, 2013). Children who have difficulties with executive functioning are less likely to be able to stay on task in the classroom and succeed academically (St Clair-Thompson and Gathercole, 2006). It is likely, therefore, that improvements in executive functioning mediate the relationship between PA and on on-task behavior.

Before-school PA programs can potentially improve students' readiness to learn without taking time away from academics. However, research in this area is limited. Although preliminary evidence suggests that participation in a before-school program using an interactive multi-media PA training system (i.e., that uses DVR technology to engage children in PA) can benefit children's on-task behavior (Mahar et al., 2011), that specific intervention required a significant financial investment, which may make it an unfeasible option for many schools. Hence, this study's purpose was to examine whether elementary school children's on-task behavior during the first 45 min of instruction improved on days they participated in a before-school running/walking club.

Methods

Participants and settings

Participants were third and fourth grade students from two schools in the Southwestern U.S. Schools were purposively selected and represented two different settings (i.e., private and public). The goal was to replicate the study across the two settings rather than compare the two schools.

School A

School A was a high-achieving K-8 private school (over 95% of students meeting/exceeding standards in 2014–2015 state standardized testing for English Language Arts, Math, and Science), with a total enrollment of 273 students (primarily Caucasian) and an average class size of 18 students. The teachers of the four participating classes (all Caucasian) had a mean teaching experience of 11.50 years (SD = 6.14 years).

School B

School B was an average-performing K-6 public school (26%, 35%, and 60% of students meeting/exceeding standards in 2014–2015 state standardized testing for English Language Arts, Math, and Science, respectively), with a total enrollment of 451 students (57.43% Caucasian, 33.92% Hispanic, 8.65% other) and an average class size of 30 students. Approximately 60% of this school's students were eligible for free or reduced-price lunch. The teachers of the participating classes (Caucasian = 3; Hispanic = 1) had a mean teaching experience of 21.75 years (SD = 9.03 years).

Target behavior

The target behavior was students' on-task behavior during the first 45 min of instruction. Consistent with Mahar et al.'s (2006) definitions, *on-task behavior* was defined as behavior that followed the class rules and was appropriate to the learning situation (e.g., listening to directions, working quietly at one's desk). *Off-task behavior* was defined as any behavior that was not on task (i.e., broke the classroom rules or interrupted the learning situation; making noise, not participating when necessary, staring into space).

Research design

This study used an alternating treatments design (ATD) (Cooper et al., 2007) with an initial baseline phase. ATDs test the relative effectiveness of two (or more) treatments/conditions and originate from Applied Behavior Analysis and single-case designs (Barlow and Hersen, 1984; Cooper et al., 2007). In ATDs, participants receive both treatments/conditions (i.e., a person/group serves as its own control) on an alternating schedule, which helps control for most threats against internal validity (Barlow and Hersen, 1984). The addition of the baseline phase serves to establish a counterfactual, determines the level of stability in behavior, and helps control for regression to the mean as well as for testing effects.

Phases, conditions, and data points

The two conditions compared during the AT phase included a non-treatment (no program as in baseline phase) and a treatment condition. The baseline phase for both schools included five data points (over a week for School A; over two weeks for School B). The AT phase included five data points for each condition (10 total over five weeks) at each school. The number of data points for each phase was determined based on the minimum number required (i.e., two for each condition; Barlow and Hersen, 1984) and practical considerations (i.e., predetermined program starting day, limited program duration, natural school breaks). It was expected that the two phases would be adequate to reveal potential condition differences.

Condition sequencing and discrete conditions

The before-school program for each school occurred on specific weekdays, which did not allow randomly counterbalancing the two conditions. However, the non-treatment data points were manipulated so they occurred both on days before and after the treatment data points (at least two times before and two times after). This, combined with the clearly discrete nature of the two conditions (i.e., non-treatment vs. treatment), helped control for potential order effects and minimized possible carryover effects (Barlow and Hersen, 1984).

Intervention: before-school physical activity program

The before-school program in both schools involved an unstructured running/walking club that took place twice a week (Tuesdays and Thursdays) on each school's oval/field and was supervised by each school's physical education teacher. Students were instructed that they could either walk or run for the duration of the program. On program days in both schools, students went straight to their classrooms at the end of the program to complete their morning work.

School A

At School A, the program lasted 20 min, from 7:50 am to 8:10 am, and classes officially started at 8:30. On days without the program, students arrived to school around 8:00 and the period between 8:00 and 8:30 was homeroom time during which students got ready and completed morning work.

School B

At School B, the program lasted 15 min, from 7:20 am to 7:35 am, and classes officially started at 8:00. Students were not allowed on campus before 7:20 and the period between 7:40 and 8:00 was considered homeroom time.

Data collection & procedures

Institutional Review Board, district, principal, and teacher approvals were obtained prior to starting the study. Also, student assent and parental consent forms were collected. Data were collected on:

(a) participation in the before-school program, (b) anthropometric measures, and (c) on-task behavior.

Participation in before-school program

Program participation was established by meeting a minimum threshold of participation. Specifically, students were considered to have participated in the program each time if they had accumulated at least five minutes of moderate-to-vigorous PA (MVPA) (25.00% and 33.33% of program duration in the two schools, respectively). Attendance was taken by each school's physical education teacher and a research team member. PA levels were monitored using *New Lifestyles NL-1000* pedometers. Detailed information about PA data collection is presented elsewhere (Stylianou, in review).

Anthropometric measures

Height and weight measurements were obtained without shoes and heavy clothing using a calibrated digital scale and stadiometer. These were used to calculate students' body mass index (BMI) [weight (kg)/height^2 (m)] and BMI-for-age percentile using CDC's BMI tool for schools (CDC, n.d.). Subsequently, BMI-for-age percentiles were used to classify students as normal-weight or overweight/obese based on CDC'S BMI-for-age growth charts.

On-task behavior

Students' classroom behavior was observed during the first 45 min of instruction and was classified as either on-task or off-task according to Mahar et al.'s (2006) definitions (see "Target behavior"). The measurement/observation process followed the steps identified by Mahar (2011) as essential for generating credible data (i.e., defining behavior, training observers, determining type/length of recording, and assessing inter-observer reliability).

Observation system

The observation method employed was similar to that used by Mahar et al. (2006, 2011). Like in Mahar et al.'s (2011) study, observation and recording intervals lasted five seconds each; shorter intervals can increase data reliability since there is a lower probability for more behaviors to occur within shorter intervals, and also allow for observation of more intervals in a given amount of time. Each student was observed for one minute (i.e., six observation intervals) before the observer rotated to the next student. The rotation was repeated until each student had been observed for a total of three minutes (i.e., 18 observation intervals).

On-task and off-task behavior recording

On-task behavior was coded using whole interval recording (i.e., when behavior was on-task for the entire interval) and off-task behavior was coded using partial interval recording (i.e., if off-task behavior was exhibited at any point during the interval).

Scoring

A student's score for a specific day was a percentage calculated by summing the number of intervals in which each behavior (i.e., on/off-task) occurred, dividing by the total number of intervals, and multiplying by 100.

Observer training and inter-observer agreement

Two primary observers and one secondary observer assessed classroom behavior. In both schools, each primary observer assessed the same two classes throughout the study. Also, during the study, the secondary observer participated in at least 25% of all observations with the primary observers for determining inter-observer agreement.

Before the study, observers trained in observing/coding on/off-task behavior by watching videotapes and attending live classroom instruction until they reached at least a 90% agreement level five successive times. A total of 16 live observations were conducted to reach the desired level of agreement. Observers also practiced observation/coding in participating classrooms for one week before beginning data collection to get familiarized with the setting and minimize potential reactivity effects.

Content, activities, and rules

The subjects taught during the first 45 min of instruction were consistent within each school and were the same across the different phases and conditions (School A: math; School B: literacy/reading). Class activities were the same within grade levels for each school, but were also very similar across grade levels within each school, and mainly included individual work and direct instruction. Further, both schools had school-wide rule and citizenship systems, which included aspects such as being good listeners, raising hand, waiting for their turn, keeping hands/feet to oneself, and using equipment appropriately.

Data analysis

Inter-observer agreement

Percentages of inter-observer agreement were calculated by dividing the number of intervals with common codes (i.e., agreement) by the total number of observation intervals and then multiplying by 100. Three methods were used for calculating inter-observer agreement: interval-by-interval (I-I) method (uses all intervals), scored-interval (S-I) method (ignores intervals where both observers record the non-occurrence of the behavior), and unscored-interval (U-I) method (ignores intervals where both observers record the occurrence of the behavior). Since behavior frequency can influence the results of each method, Cooper et al. (2007) recommend reporting all three methods. The S-I and U-I methods are stringent tests of agreement and typically produce lower scores than the I-I method. Generally, 80% is considered an acceptable level of agreement for S-I and U-I agreement calculations (Metzler, 1983). For high-rate behaviors, like on-task behavior in this manuscript, the I-I and S-I methods may overestimate inter-observer agreement.

Visual analysis

Initially, individual graphs of on-task behavior were created to examine the extent to which student behavior followed a zero/flat, decreasing, or increasing trend during the baseline phase. For this study, where the desired effect of the treatment would be an enhancement of on-task behavior, zero/flat or negative trends during the baseline phase would be ideal. Subsequently, graphs of average on-task behavior were developed by phase and condition (for each grade level at each school), and were analyzed by examining whether average on-task behavior was consistently higher on days participants attended the before-school program. The unit of analysis was the grade level at each school and, thus, standard deviations are also included in graphs. In these graphs, students were included in treatment data points (i.e., data points for treatment/program days) if they had met the minimum threshold of participation.

The main criteria used to determine experimental effects include variability and trends within and between phases/conditions, data overlap between phases/conditions, immediacy of behavior change from one phase/condition to the next, and the distance between data paths of different conditions (Cooper et al., 2007; Parsonson and Baer, 1978). In general, the greater the immediate change with the introduction of a new phase/condition, the larger the distance between data paths of different conditions, the lower the data overlap between phases/conditions, and the more desirable the trends of the data, the stronger the case for an experimental effect.

Tau-U effect size

The Tau-U index was calculated to supplement the visual analysis. This index reflects the percentage of data that improve over time considering nonoverlap between phases and intervention phase trend, and also allows controlling for baseline trend (Parker et al., 2011). Tau-U has shown robustness against autocorrelation, it can control for baseline monotonic trend, and is applicable even for short phases (Parker et al., 2011). Values of 0–.65, .66–.92, and .93–1, represent small, medium, and large/strong effects, respectively (Parker and Vannest, 2009). Confidence intervals (83.4%) were also calculated to examine differences between the alternating treatments/conditions (non-overlapping 83.4% intervals indicate statistically significant differences ($p < .05$); Parker et al., 2010).

Statistical analysis

Statistical analyses were performed within a multilevel modeling framework with daily observations as level-1 variable and person-level variables (e.g., grade level) as level-2 variables. Analyses were conducted separately for the two schools since the purpose was to replicate the study in two different settings. Also, analyses were conducted by individual grade level as well as for the two grade levels combined (with grade level as a covariate). For this analysis, program/treatment days on which students did not attend the before-school program or did not meet the minimum threshold of participation were coded as non-treatment days.

Effect size

The *pseudo-*R^2 effect size (Raudenbush and Bryk, 2002; Singer and Willett, 2003) was calculated, which is interpreted as the proportion reduction in variance for a parameter estimate that results from comparing the variance component (i.e., residual/level-1 variance, intercept/level-2 variance) in a baseline model to the same variance component in a fuller model. Cohen's (1988) criteria for R^2 values were used to interpret *pseudo-*R^2 values, with .02, .13, and .26 representing small, medium, and strong effects, respectively.

Models tested

Initially, a series of unconditional models (i.e., without predictors) were conducted to obtain mean values of on-task behavior by phase and condition. Subsequently, some preliminary analyses were conducted. These included testing for potential effects of the design-based confounding variables of phase and order (level-1 variables) on day-to-day on-task behavior over and above treatment, to determine if they needed to be included as covariates in subsequent models. These two variables were tested in separate models since order was only present in the alternating treatment phase.

To examine the impact of the treatment on students' on-task behavior, a random intercept model was tested, which allows intercepts (i.e., on-task behavior means) to vary across individual students but assumes constant slopes (i.e., same influence of treatment on on-task behavior). Finally, a random slope model was also tested that allows both intercepts (i.e., on-task behavior means) and slopes (i.e., treatment effect) to vary across individual students.

Results

Participants and program participation

Table 1 includes information about the students from each school who volunteered to participate in the study and who attended the before-school program at least once. At School B, eight students did not attend the before-school program at all (i.e., they either were absent or did not meet the minimum threshold of participation on days they attended the program) and were therefore excluded from analyses. No significant differences were observed between the students who did and did not attend the program in terms of sex (Fisher's exact test $p > .05$), BMI status (Fisher's exact test $p > .05$), and baseline on-task behavior ($t(45) = .88$, $p > .05$; controlling for sex and BMI status).

Table 1
Participant information.

	Volunteered to participate		Attended program at least once	
	Boys	Girls	Boys	Girls
School A				
Grade 3	5	11	5	11
Grade 4	9	14	9	14
	$n = 14$	$n = 25$	$n = 14$	$n = 25$
	$N = 39$		$N = 39$	
Ethnicity	89.74% Caucasian; 10.26% Other			
BMI status	23.68% Overweight/Obese*			
School B				
Grade 3	13	7	12	5
Grade 4	11	15	10	11
	$n = 24$	$n = 22$	$n = 22$	$n = 16$
	$N = 46$		$N = 38$	
Ethnicity	56.52% Caucasian, 30.43% Hispanic, 13.04% Other		55.26% Caucasian, 31.58% Hispanic, 13.16% Other	
BMI status	36.96% Overweight/Obese*		39.47% Overweight/Obese*	

Note. School A: 49.37% of Grade 3 and 4 students volunteered to participate in the study. School B: 38.98% of Grade 3 and 4 students volunteered to participate in the study.
* ≥85th percentile, based on CDC'S BMI-for-age growth charts for boys and girls.

Inter-observer agreement

During the study, the secondary observer participated in 25% and 30% of all observations in Schools A and B, respectively. The resulting inter-observer reliability rates were: for the I-I method, 91.93% for School A (range: 85.00–97.53%) and 92.25% for School B (range: 85.61–98.08%); for the S-I method, 89.32% for School A (range: 82.61–95.8%) and 90.01% for School B (range: 83.18–97.02%); and for the U-I method, 81.38% for School A (range: 70.42–95.05%) and 79.76% for School B (range: 68.92–94.96%).

Visual analysis

According to the individual graphs of on-task behavior during the baseline phase, the behavior of over 80% of participants at both schools followed either zero/flat or decreasing trends (i.e., "counter-therapeutic baseline"). Figs. 1–4 present graphs of average on-task behavior across phases and conditions for each grade level at each school. As can be observed, the data paths for the non-treatment condition during the alternating treatments phase are mostly consistent with the data paths of the initial baseline phase. Also, an immediate increase in on-task behavior is evident with the introduction of the treatment/program. Additionally, there is a distinct level difference between the two conditions, as well as an absence of overlap between the data paths of the two conditions, both of which indicate a condition/treatment effect. Some overlap can be observed between the standard deviation bars of the two conditions, which reflect between-person variability and should be taken into account when interpreting these graphs. Collectively, these graphs provide evidence that students' on-task behavior levels were higher on days they participated in the before-school program compared to days they did not.

Tau-U

Tau-U values are presented in Table 2. The effect sizes for the treatment condition are strong in both schools (>.93). The lack of overlapping confidence intervals between the two conditions indicates statistically significant differences for School A. However, there is a small overlap between the confidence intervals for the two conditions for School B.

Statistical analysis

Mean on-task behavior values by phase and condition are presented in Table 3.

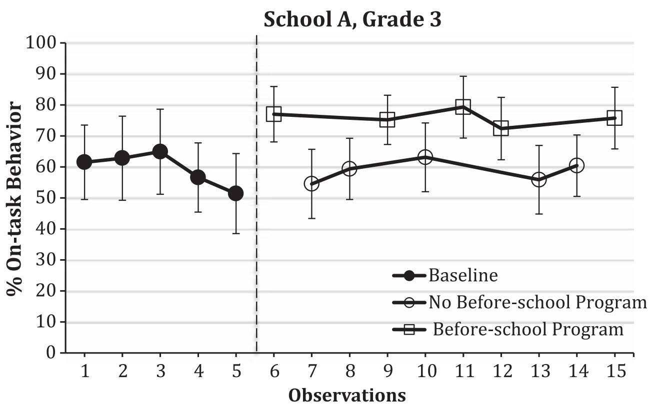

Fig. 1. On-task behavior graph for School A, Grade 3.

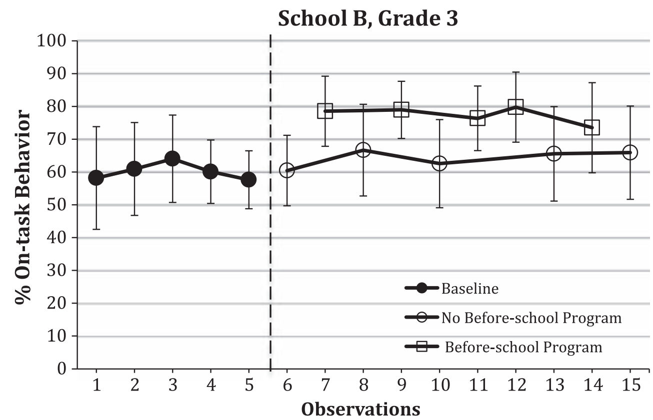

Fig. 3. On-task behavior graph for School B, Grade 3.

Preliminary analyses

According to the results of these analyses, order did not significantly contribute to the prediction of on-task behavior when controlling for treatment. Phase significantly contributed to the prediction of on-task behavior above and beyond treatment only for third grade students at School B. Consequently, order was not included as a covariate in subsequent models, and phase was included as a covariate only for the specific group.

Primary analyses

The results of the random intercept models, which tested the (uniform) effect of the treatment across participants, demonstrated a significant positive effect of the before-school program on students' on-task behavior. The results of these models are presented in Table 4, which presents increases in on-task behavior on days of participation in the before-school program and associated effect sizes.

The models that included a random slope for treatment (see *Equation 4*) did not converge, even when the number of iterations was increased. This possibly indicates "…insufficient information to warrant allowing level-2 residuals for both initial status and rates of change" (Singer and Willett, 2003, p. 156) and suggests a uniform impact of the treatment across participants.

Discussion

This study's purpose was to examine the impact of a before-school running/walking club on students' on-task behavior during the first 45 min of instruction. The visual analysis of the plotted graphic data of on-task behavior provided evidence that students' on-task behavior was consistently higher on days they participated in the before-school PA program than on days they did not. For the most part, Tau-U values and confidence intervals provided support to the visual analysis results. A treatment effect was corroborated by subsequent statistical analyses, which demonstrated significant improvements in on-task behavior on days students attended the before-school program for both individual grade levels and school levels (both grade levels combined). Further, *pseudo-*R^2 effect sizes indicated moderate to strong effects.

This study's results are consistent with the results of Mahar et al. (2011), who found that participation in a before-school PA program had a significant positive impact on students' on-task behavior. Similarly, improvements in classroom on-task behavior were also found in studies that focused on classroom PA (Goh et al., 2012; Mahar et al., 2006) and recess (Barros et al., 2009; Jarrett et al., 1998; Pellegrini et al., 1995; Ridgway et al., 2003). The mechanism underlying the relationship between PA and on-task behavior is yet unclear but may be linked to improvements in cognitive processes, and particularly executive functioning, following acute bouts of PA. In turn, these cognitive improvements have been attributed to various mechanisms, including exercise-induced increases in cerebral blood flow and the release of biochemical substances such as brain-derived neurotrophic factor and dopamine (Chang et al., 2012; Hillman et al., 2009).

In this study, no differential effect of the treatment across individual students was found. This may be a function of the relatively small sample size and should be explored further in future studies. Other studies in which children were grouped based on their baseline on-task behavior or attention levels found differential effects of PA on on-task behavior (Mahar et al., 2006; Ridgway et al., 2003). Collectively, these findings suggest that PA programs before and during the school day may help to

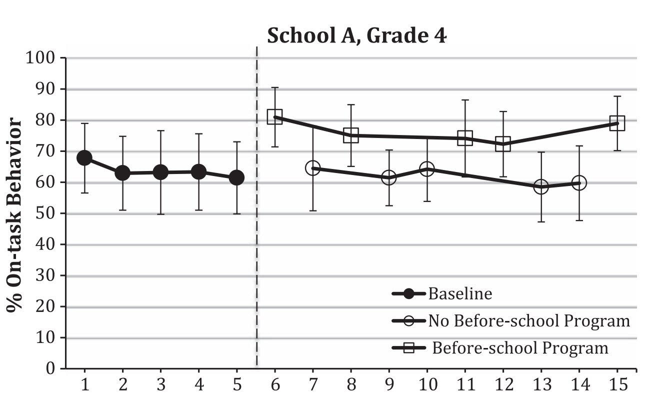

Fig. 2. On-task behavior graph for School A, Grade 4.

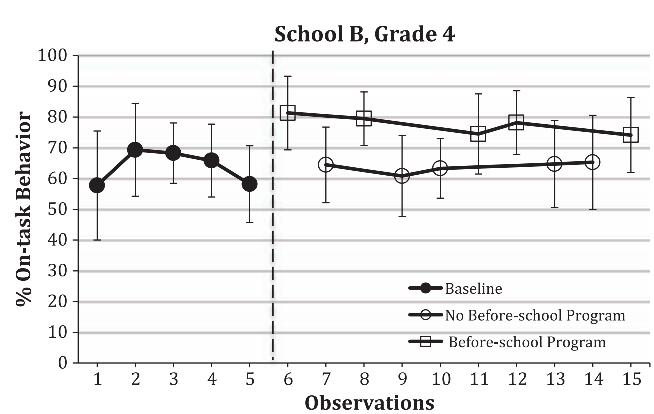

Fig. 4. On-task behavior graph for School B, Grade 4.

Table 2
Tau-U effect sizes.

	Baseline phase vs. AT phase (non-treatment condition)			Baseline phase vs. AT phase (treatment condition)		
	Tau-U	*p*	83.4% CI	Tau-U	*p*	83.4% CI
School A	−.24	.376	−.62–.15	1	<.001	.60–1.36
School B	.28	.301	−.11–.66	1	<.001	.60–1.36

Note. AT = alternating treatments.

increase classroom on-task behavior for all students and may have a greater effect for the students least on-task.

The strengths of this study include its design and the replication at a second setting. Combined, these features provide support both for internal and external validity. However, readers should bear in mind that the two sites were conveniently chosen, which limits generalizability beyond the two schools. Additionally, the program used in this study was cost-effective and did not reduce academic instruction time during the school day, which may make it an attractive option for school administrators and classroom teachers who often hesitate to increase PA during the school day due to the pressures of high-stakes testing.

A limitation of this study was that the observers were not blinded to the conditions during the alternating treatments phase since this was not practically feasible. However, none of the observers knew whether participants had actually attended the program on a given day and/or if they had met the criterion for being considered as having participated in the program (i.e., they observed all participating students who were present in the classroom on a given day). Moreover, the agreement rates between observers provide support for the relative objectivity of the data. Another limitation is that each participant was observed for a small portion of each class period.

Conclusions

This study's findings provide support for the positive effect of before-school PA programs on students' on-task behavior. Although teachers and administrators may hesitate to increase PA programming during the school day, before-school programs take no time away from academic instruction time and can still have a positive impact on both students' on-task behavior during the first part of the school day and health.

This study's findings have implications about the structure of children's day and the school schedule. Typically, children have little to no time and/or opportunities to engage in PA before school, and most PA programs occur after school. School personnel may want to consider later start times and providing various PA opportunities in the morning. Scheduling changes need not be drastic; a 20-min or 30-min delay in the start of the school day may be enough to provide children with a satisfactory amount of PA that can improve their readiness to learn, at least within the first classroom period. Schools should also consider making appropriate arrangements regarding school buses, breakfast offered at school, and other factors that may influence children's participation in PA before school.

Additional research is needed to explore the effects of different types of PA programs on classroom behavior. It is recommended that future studies examine the residual (i.e., lingering) effects of before-school PA programs on different types of classroom behavior as the day progresses (i.e., beyond the first 45 min of instruction), as well as the dose–response relationship between MVPA and on-task behavior. Future studies should recruit secondary school student participants since most relevant studies have been conducted with elementary-aged children, and they should also examine the effects of similar interventions on students with high levels of off-task behavior (e.g., attention-deficit

Table 3
Mean values of on-task behavior by phase and condition.

	Baseline phase			Alternating treatments phase: Non-treatment condition			Alternating treatments phase: Treatment condition		
	Mean	Within-person *SD*	Between-person *SD*	Mean	Within-person *SD*	Between-person *SD*	Mean	Within-person *SD*	Between-person *SD*
School A									
3rd grade	59.47	12.48	4.09	58.82	9.82	5.60	76.56	8.69	5.14
4th grade	63.51	9.61	7.52	61.60	8.17	7.71	77.00	7.69	6.48
Both grade levels	61.82	10.92	6.56	60.43	8.91	6.90	76.77	8.21	5.83
School B									
3rd grade	60.28	10.04	8.14	64.57	10.38	8.65	77.83	9.43	6.88
4th grade	64.14	13.01	6.89	64.32	12.69	5.38	77.58	11.40	5.31
Both grade levels	62.34	11.62	7.65	64.37	11.82	6.79	77.68	10.25	5.94

Note. Non-treatment condition includes days during the alternating treatments phase on which participants either did not attend the before-school program or did not accumulate at least five minutes of MVPA within the program. Treatment condition includes days on which participants attended the before-school program and accumulated at least five minutes of MVPA.

Table 4
Estimates for on-task behavior random intercept models.

	School A				School B			
	Estimate	SE	*p*	*Pseudo-R*^{2}	Estimate	SE	*p*	*Pseudo-R*^{2}
3rd grade								
B_0 (intercept)	59.15	1.53	<.001	–	60.18	2.04	<.001	–
B_1 (treatment)	17.80	1.55	<.001	.37**	13.22	1.69	<.001	.32**
B_2 (phase)	–	–	–	–	4.33	1.47	<.010	
4th grade								
B_0 (intercept)	62.46	1.70	<.001	–	64.23	1.38	<.001	–
B_1 (treatment)	14.10	1.18	<.001	.32**	13.15	1.73	<.001	.17*
Both grade levels								
B_0 (intercept)	59.63	1.82	<.001	–	62.59	1.68	<.001	–
B_1 (treatment)	15.78	.95	<.001	.34**	14.26	1.16	<.001	.22*
B_2 (grade level)	2.46	2.35	.303	–	1.40	2.23	.535	–

Note. Pseudo-R^2 (i.e., effect size) values in this table reflect proportion reduction in residual or level-1 (day-to-day) variance. **Strong effect; *Moderate effect.

disorder). Further, it is recommended that future studies explore potential differential effects of PA across individual students and across subgroups of students (e.g., girls vs. boys, normal-weight vs. overweight). Finally, both students and teachers should be queried regarding the effects of PA on on-task behavior. This would add to the social validity of such PA programming.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Transparency document

The Transparency document associated with this article can be found, in online version.

References

Barlow, D.H., Hersen, M., 1984. Alternating Treatments Design. In: Barlow, D.H., Hensen, M. (Eds.), Single Case Experimental Designs: Strategies for Studying Behavior Change. Pergamon Press, New York, NY, pp. 252–284.
Barros, R.M., Silver, E.J., Stein, R.E.K., 2009. School recess and group classroom behavior. Pediatrics 123, 431–436.
Center on Education Policy (CEP), 2008. Instructional Time in Elementary Schools: A Closer Look at Changes for Specific Subjects. Author, Washington, DC.
Centers for Disease Control and Prevention (CDC), 2010. The Association Between School Based Physical Activity, Including Physical Education, and Academic Performance. U.S. Department of Health and Human Services, Atlanta, GA.
Centers for Disease Control and Prevention (CDC). Children's BMI tool for schools. Retrieved from: http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/tool_for_schools.html
Chang, Y.K., Liu, S., Yu, H.H., Lee, Y.H., 2012. Effect of acute exercise on executive function in children with attention deficit hyperactivity disorder. Arch. Clin. Neuropsychol. 27, 225–237. http://dx.doi.org/10.1093/arclin/acr094.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. second ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
Cooper, J.O., Heron, T.E., Heward, W.L., 2007. Applied Behavior Analysis. second ed. Pearson, Upper Saddle River, NJ.
Cothran, D.J., Kulinna, P.H., Garn, A.C., 2010. Classroom teachers and physical activity integration. Teach. Teach. Educ. 26, 1381–1388. http://dx.doi.org/10.1016/j.tate.2010.04.003.
Diamond, A., 2013. Executive functions. Annu. Rev. Psychol. 64, 135–168. http://dx.doi.org/10.1146/annurev-psych-113011-143750.
Goh, T.L., Hannon, J.C., Fu, Y., Prewitt, S.L., 2012. Children's physical activity and on-task behavior following active academic lessons [Abstract]. Retrieved from http://aahperd.confex.com/aahperd/2013/webprogram/Paper18344.html.
Hillman, C.H., Pontifex, M.B., Raine, L.B., Castelli, D.M., Hall, E.E., Kramer, A.F., 2009. The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children. Neuroscience 159, 1044–1054. http://dx.doi.org/10.1016/j.neuroscience.2009.01.057.
Institute of Medicine (IOM), 2013. Educating the Student Body: Taking Physical Activity and Physical Education to School. The National Academies Press, Washington, DC.
Janssen, I., LeBlanc, A.G., 2010. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int. J. Behav. Nutr. Phys. Act. 7, 40. http://dx.doi.org/10.1186/1479-5868-7-40.
Jarrett, O.S., Maxwell, D.M., Dickerson, C., Hoge, P., Davies, G., Yetley, A., 1998. Impact of recess on classroom behavior: Group effects and individual differences. J. Educ. Res. 92, 121–126.
Mahar, M.T., 2011. Impact of short bouts of physical activity on attention-to-task in elementary school children. Prev. Med. 52, S60–S64. http://dx.doi.org/10.1016/j.ypmed.2011.01.026.
Mahar, M.T., Murphy, S.K., Rowe, D.A., Golden, J., Shields, A.T., Raedeke, T.D., 2006. Effects of a classroom-based program on physical activity and on-task behavior. Med. Sci. Sports Exerc. 38, 2086–2094. http://dx.doi.org/10.1249/01.mss.0000235359.16685.a3.
Mahar, M.T., Vuchenich, M.L., Golden, J., DuBose, K.D., Raedeke, T.D., 2011. Effects of a before-school physical activity program on physical activity and on-task behavior [Abstract]. Med. Sci. Sports Exerc. 43 (Suppl. 1 5), 24. http://dx.doi.org/10.1249/01.MSS.0000402740.12322.07.
Metzler, M., 1983. An Interval Recording System for Measuring Academic Learning Time. In: Darst, P.W., Mancini, V.H., Zakrajsek, D.B. (Eds.), Systematic Observation Instrumentation for Physical Education. Leisure Press, Champaign, IL, pp. 181–195.
National Association for Sport and Physical Education (NASPE), American Heart Association (AHA), 2012P. 2012 Shape of the Nation Report: Status of Physical Education in the USA. AAHPERD, Reston, VA.
National Physical Activity Plan Alliance, 2014. The 2014 United States Report Card on Physical Activity for Children and Youth. National Physical Activity Plan, Columbia, SC.
Parker, R.I., Vannest, K., 2009. An improved effect size for single-case research: Nonoverlap of all pairs. Behav. Ther. 40, 357–367. http://dx.doi.org/10.1016/j.beth.2008.10.006.
Parker, R.I., Vannest, K.J., Davis, J.L., Clemens, N.H., 2010. Defensible progress monitoring data for medium and high stakes decisions. J. Spec. Educ. 44, 221–233. http://dx.doi.org/10.1177/0022466910376837.
Parker, R.I., Vannest, K.J., Davis, J.L., Sauber, S.B., 2011. Combining nonoverlap and trends for single-case research: Tau-U. Behav. Ther. 42, 284–299. http://dx.doi.org/10.1016/j.beth.2010.08.006.
Parsonson, B.S., Baer, D.M., 1978. The Analysis and Presentation of Graphic Data. In: Kratochwill, T.R. (Ed.), Single Subject Research: Strategies for Evaluating Change. Academic Press, Inc., Orlando, FL, pp. 101–165.
Pellegrini, A.D., Huberty, P.D., Jones, I., 1995. The effects of recess timing on children's playground and classroom behaviors. Am. Educ. Res. J. 32, 845–864.
Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. second ed. Sage Publications, Inc., Thousand Oaks, CA.
Ridgway, A., Northup, J., Pellegrin, A., LaRue, R., Hightshoe, A., 2003. Effects of recess on the classroom behavior of children with and without attention-deficit hyperactivity disorder. Sch. Psychol. Q. 18, 253–268.
Sibley, B.A., Ethier, J.L., 2003. The relationship between physical activity and cognition in children: A meta-analysis. Pediatr. Exerc. Sci. 15, 243–256.
Singer, J.D., Willett, J.B., 2003. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford University Press, New York, NY.
St Clair-Thompson, H.L., Gathercole, S.E., 2006. Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. Q. J. Exp. Psychol. 59, 745–759. http://dx.doi.org/10.1080/17470210500162854.
Stylianou, M., van der Mars, H., Kulinna, P.H., Adams, M.A., Mahar, M., Amazeen, E. Before-school running/walking club and student physical activity levels: An efficacy study. (in review).
Troiano, R.P., Berrigan, D., Dodd, K.W., Mâsse, L.C., Tilert, T., McDowell, M., 2008. Physical activity in the United States measured by accelerometer. Med. Sci. Sports Exerc. 40, 181–188. http://dx.doi.org/10.1249/mss.0b013e31815a51b3.