Effects of a health-related physical fitness intervention on middle school students’ academic learning time during physical education

You Fu¹, Ryan D. Burns², Wei Yang¹, Timothy A. Brusseau², James C. Hannon³

¹School of Community Health Sciences, University of Nevada, Reno, Nevada, USA; ²Department of Health, Kinesiology, & Recreation, University of Utah, Salt Lake City, Utah, USA; ³College of Education, Health, and Human Services, Kent State University, Kent, Ohio, USA

Contributions: (I) Conception and design: Y Fu; (II) Administrative support: JC Hannon, TA Brusseau; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: Y Fu, RD Burns; (V) Data analysis and interpretation: RD Burns; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: You Fu, PhD, Assistant professor. School of Community Health Sciences, University of Nevada, Reno, 1664 N Virginia Street, Reno, Nevada 89557, USA. Email: youf@unr.edu.

Background: This study examined the effects of the sports, play and active recreation for kids (SPARK) on middle school students’ academic learning time—physical education (ALT-PE) compared to the traditional physical education (PE) program.

Methods: A total of 12 target middle-school students were randomly selected from a pool of 174 who participated in PE lessons over 9 weeks. Two raters used a 12-second-interval observation/record protocol to record the student’s context level (general, subject matter knowledge, and subject matter motor). In the SPARK school, target students attended the SPARK lessons for 9 weeks that included three curricular sport activities in the order of soccer, flag football, and ultimate Frisbee. Target students’ percentages of time spent in ALT-PE context categories and subcategories were calculated and compared by different PE programs and sport activities. A series of 2 × 3 mixed design ANOVA tests were conducted using SPSS statistical software.

Results: Compare to traditional PE group, SPARK group had significant difference of percentage of time in general content (mean difference =−19.97%, P<0.05, d=2.12) and in subject matter motor (mean difference =+24.64%, P<0.05, d=1.99). Specifically, students in SPARK program had statistically significant higher values in percentage of time spent in the subcategory of skill practice (mean difference =25.57%, P<0.01, d=1.98) and fitness (mean difference =6.22%, P<0.01, d=1.45).

Conclusions: SPARK is an effective pedagogical strategy to increase youth’s ALT in school PE settings.

Keywords: Sports, play and active recreation for kids (SPARK); traditional physical education (traditional PE); context level; academic learning time

Introduction

The health of American children is a major public concern with obesity rates steadily increasing over the past 30 years (1,2) due, in part, to low levels of physical activity (3). Sedentary behaviors and low levels of physical activity in adolescents can track into adulthood, which manifests potential significant consequences for both individuals and society as a whole (4). Because of the increased recognition on the importance of optimizing health in children, numerous research studies have been conducted on the youth population to examine effective strategies to increase healthy behaviors in school settings (5,6).

Evidence has suggested that school physical education
(PE) programs are viable venues to provide children and adolescents the opportunities to participate in physical activity (7). Consequently, school PE programs play a critical role in promoting students’ health and fitness (8,9). The amount of time students spend in academic learning is a key element that contributes to the quality of teaching and learning processes in PE. Optimal time spent in academic learning is also highly correlated with the student’s achievement (10,11). Academic learning time-physical education (ALT-PE) is an application of academic learning time in school PE setting, which has been extensively studied as a measure of teaching effectiveness and students’ learning achievements in school PE settings (12-14). It has been reported that ALT-PE is an important mediator between teaching behavior and learning achievement, therefore improvement in ALT-PE could result in students’ improved performance (15).

Although numerous studies have emphasized the significance of ALT-PE in enhancing the quality of the PE and students’ achievements (16-18), it is still unknown if a health-related physical fitness program, the sports, play and active recreation for kids (SPARK), could significantly affect middle school students’ ALT-PE. In addition, how students spend their time in the various context levels (14) of curricular activities in the SPARK program is unclear. Therefore, this study examined middle-school students’ ALT-PE in various sport activities using the SPARK program compared to a traditional PE model. It is hypothesized that the students in the SPARK group would spend significantly more time in the ALT-PE categories of motor, but significantly less time in the category of general content compared to the traditional group. The SPARK program will provide students significantly more opportunities for participation in the ALT-PE subcategories of fitness, skill practice, and games, but significantly less time spent in transition/break time compared to the traditional group.

Methods

Participants and setting

The present study consisted of a pool of 174 (82 boys, 92 girls, $M_{age} = 12.06$, SD =0.85) sixth to eighth grade middle-school students that were enrolled in two urban private schools from the Mountain West Region of the U.S. The ethnic distribution consisted of 82.3% Caucasian, 12.0% Hispanic, 2.3% Asian or Pacific Islander, 1.7% African American, and 1.7% Other (Indian or Native of America). In the experimental school (SPARK), there were 75 (33 boys, 42 girls) participants from three PE classes, with one class from each grade, including 25 sixth graders, 27 seventh graders, and 23 eighth graders, respectively. In the control school (traditional), there were 99 (49 boys, 50 girls) participants from five PE classes, including 36 sixth graders, 31 seventh graders, and 32 eighth graders.

In the SPARK school, PE lessons were conducted once a week for approximately 60 minutes. The teacher was male, with more than 10 years K-12 PE teaching experience. PE lessons were also conducted once a week in the control school, where the average PE lessons were approximately 60 minutes. The PE teacher was male, with more than 15 years PE teaching experience. Permission to conduct the study was obtained from the University Institutional Review Board, the school administrations, and the PE teachers prior to the start of this study.

It has been reported that 3 or more students would be an acceptable subject sample to obtain a valid estimate through systematic observation (11,12,14,19). In this study, a total of 12 target students (6 from the SPARK school, 6 from traditional school) were randomly selected from the larger pool for data collection and analysis. In the SPARK school, the PE teacher provided two lists of students who met the inclusion criteria to the investigator. Specifically, the first list included all girls from 3 grades, and each girl’s skill level was identified on the list. A second list was created for boys. Skill levels (i.e., low-skilled, average-skilled, or high-skilled) were classified by the PE teacher based on each student’s personal evaluation. The investigator randomly assigned the students into teams of three based on grade and skill level, and ensured that all grades and skill levels were represented. Then the investigator randomly selected one team from each list, and finalized the target students by combining the two teams. The target students in the traditional PE group were selected using an identical protocol. Table 1 indicates each target student’s characteristics.

ALT-PE measure

The outcome variable, the percentage of time students spend in each category of context level during PE class, was measured using the context level of the ALT-PE systematic measurement observation instrument (14), which describes the context that the student is behaving in and refers to the class as a whole. There are 13 subcategories of context from three major categories: general content, subject
matter knowledge, and subject matter motor (see Table 2). Specifically, it is predicted that the three categories of the context level in the ALT-PE instrument would demonstrate how students spend their time in their class (14). The ALT-PE systematic measurement observation instrument has been shown to have satisfactory internal consistency and validity in PE settings (14). Refer to Siedentop et al. (14) for the definition of each category and subcategory in the ALT-PE observational instrument.

**Data collection and procedures**

**Equipment.**

A digital recorder was used to videotape each student’s PE lesson across the entire data collection period. The purpose was to keep a record of events for the study purpose. The digital videotape recorder was set up in one of the corners of the gymnasium or field during the PE class in order to observe the whole class.

**Training of raters**

Two raters were employed to observe and collect data for the purpose of this study. Both raters were university graduate students working on graduate degrees in PE and science studies, respectively. The raters were presented with the definitions of the ALT-PE categories and subcategories at the context level (14) and attended a written test regarding the observational instrument’s categories and subcategories. Raters were required to achieve a minimum score of 90 out of 100.

There were in total 12 target students from the SPARK and traditional PE groups (6 in each group). Two students were observed per PE lesson, with each rater observing one student per lesson. Each target student was observed once a week for 9 weeks. Therefore, each rater completed a total of 54 observations (27 for SPARK, 27 for traditional) throughout the study.

**Interrater agreement reliability**

Interrater agreement is defined as the extent of agreement between human raters in recording and observing the occurrence and nonoccurrence of specific behaviors. This includes determining the reliability of the agreement between recorded data on specific target behaviors by two independent raters. The higher the interrater agreement, the more confident one can be assured that the recorded data are reliable and trustable. For this study, interrater agreement was calculated by dividing the number of instant-by-instant agreement on the occurrences of sublevels of ALT-PE context categories by the number of agreements plus the number of disagreements. The calculated proportion was then multiplied by 100 to obtain an agreement percentage (12,20). Previous research (21,22) has provided interpretative guidelines of the interrater agreement.

### Table 1 Target student information

<table>
<thead>
<tr>
<th>PE group</th>
<th>Student</th>
<th>Grade</th>
<th>Gender</th>
<th>Skill level</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPARK</td>
<td>Subject1</td>
<td>6</td>
<td>Male</td>
<td>High</td>
</tr>
<tr>
<td>SPARK</td>
<td>Subject2</td>
<td>6</td>
<td>Female</td>
<td>Medium</td>
</tr>
<tr>
<td>SPARK</td>
<td>Subject3</td>
<td>7</td>
<td>Male</td>
<td>Medium</td>
</tr>
<tr>
<td>SPARK</td>
<td>Subject4</td>
<td>7</td>
<td>Female</td>
<td>Low</td>
</tr>
<tr>
<td>SPARK</td>
<td>Subject5</td>
<td>8</td>
<td>Male</td>
<td>Low</td>
</tr>
<tr>
<td>SPARK</td>
<td>Subject6</td>
<td>8</td>
<td>Female</td>
<td>High</td>
</tr>
<tr>
<td>Traditional</td>
<td>Subject7</td>
<td>6</td>
<td>Male</td>
<td>Low</td>
</tr>
<tr>
<td>Traditional</td>
<td>Subject8</td>
<td>6</td>
<td>Female</td>
<td>High</td>
</tr>
<tr>
<td>Traditional</td>
<td>Subject9</td>
<td>7</td>
<td>Male</td>
<td>High</td>
</tr>
<tr>
<td>Traditional</td>
<td>Subject10</td>
<td>7</td>
<td>Female</td>
<td>Medium</td>
</tr>
<tr>
<td>Traditional</td>
<td>Subject11</td>
<td>8</td>
<td>Male</td>
<td>Medium</td>
</tr>
<tr>
<td>Traditional</td>
<td>Subject12</td>
<td>8</td>
<td>Female</td>
<td>Low</td>
</tr>
</tbody>
</table>

SPARK, sports, play and active recreation for kids.
agreement: the values for interrater agreement less than 0.40 are poor; values between 0.40 to 0.60 indicate fair agreement; values between 0.60 to 0.75 suggest good agreement; and values greater than .75 suggest excellent agreement. Prior to this study, training sessions were held to ensure at least 75% reliability between the raters. Data for the interrater agreement calculation was obtained by having the second rater observe 33% (n=18) of the lessons that were observed by the first rater using videotaped lessons.

**Observation procedures**

Two raters used a 12-second-interval observation/record protocol. Specifically, a student was observed for the first 6-second period of the interval, and the next 6-second period was used to record the student’s context level (general, subject matter knowledge, and subject matter motor). It has been reported that an interval between 10 to 12 seconds is an acceptable duration to obtain a valid estimate of systematic observation (11,13,19,23).

In the SPARK school, 6 target students attended the SPARK lessons for 9 weeks that included three curricular sport activities in the order of soccer, flag football, and ultimate Frisbee. Each of the 6 students was observed for one lesson weekly by the primary and secondary raters. Therefore, there were nine observations as a whole for the SPARK group, and the raters observed three classes per sport activity. Each observation was completed for all 6 students in a single school day, and each student’s time percentage in all categories and subcategories at the context level of the ALT-PE instrument were recorded and calculated. The 6 students’ percentage of time spent on the three categories and thirteen subcategories for the first 3-week observation period were averaged as their ALT-PE in SPARK soccer class; the percentages of time spent on categories and subcategories for the second 3-week observation were averaged as the students’ ALT-PE in SPARK flag football class; the time percentages of the categories and subcategories for the third 3-week observation were averaged as the students’ ALT-PE in SPARK ultimate Frisbee class. Students’ ALT-PE measurements in the traditional PE group were calculated using the identical protocol.

**Data analysis**

The data were analyzed through the following three steps. First, interrater agreement reliability was calculated to examine the reliability of the observation among two independent raters, as it is necessary to conduct the interrater agreement in single-subject research design experiments that employs human observation as the data collection method. Second, the percentages of lesson time that students spent in the three categories and 13 subcategories at the context level of the ALT-PE instrument (14) were calculated for the three sports (soccer, volleyball, and flag football) for the SPARK and the traditional PE group as descriptive statistics. Third, a series of 2×3 mixed design ANOVA tests were conducted to determine if there were any differences in students’ ALT-PE context levels in three sports (soccer, flag football, and ultimate Frisbee) between the SPARK and traditional PE groups. The dependent variable was the percentage of lesson time that students spent in each of the three categories and subcategory of the context level of the ALT-PE instrument (14). Independent variables were the two PE groups (SPARK and traditional), and the three sports (soccer, flag football, and ultimate Frisbee). Effects sizes were also calculated using Cohen’s delta. Alpha level was set at p≤0.01 to adjust for multiple comparisons and was carried out using the SPSS statistical software package (Armonk, NY, USA).

**Results**

**Interrater agreement**

Interrater agreement for ALT-PE was measured by having the second rater observe 33% (n=18) of the lessons observed by the first rater using videotape. Mean interrater agreement for ALT-PE context level was 89.04 % (ranging from 73.67–99.53%), indicating excellent interrater agreement: the values for interrater agreement less than 0.40 are poor; values between 0.40 to 0.60 indicate fair agreement; values between 0.60 to 0.75 suggest good agreement; and values greater than .75 suggest excellent agreement.

**Table 2** Context level of the ALT-PE observational instrument

<table>
<thead>
<tr>
<th>General content</th>
<th>Subject matter knowledge</th>
<th>Subject matter skill practice</th>
<th>Transition</th>
<th>Technique</th>
<th>Skill practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Break</td>
<td>Rules</td>
<td>Game</td>
<td>Management</td>
<td>Strategy</td>
<td>Scrimmage/routine</td>
</tr>
<tr>
<td>Warm-up</td>
<td>Social behavior</td>
<td>Fitness</td>
<td>Break</td>
<td>Rules</td>
<td>Game</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rules</td>
<td>Game</td>
</tr>
</tbody>
</table>

ALT-PE, academic learning time-physical education.
reliability.

**Descriptive data of ALT-PE context level**

During the 9-week intervention period, students in the SPARK group had significant lower percentage of time in general content (mean difference =19.97%, P<0.05, d=2.12) compared with their Traditional PE group counterparts, the same effects was found in its subcategory of warm-up (mean difference =13.07%, P<0.05, d=1.39). However, students in the SPARK group had significant higher percentage of time in subject matter motor (mean difference =4.64%, P<0.05, d=1.99) as compared to Traditional PE students. Specifically, students in SPARK program had very statistically significant higher values in percentage of time spent in the subcategory of skill practice (mean difference =25.57%, P<0.01, d=1.98) and fitness (mean difference =6.22%, P<0.01, d=1.45). Descriptive information was presented in Figures 1-4, which illustrate the changes of the percentage of lesson time in each ALT-PE context level within time and sport activity in each group.

**Discussion**

The purpose of this study was to examine the effects of a health-related physical fitness program on middle school students’ ALT-PE compared to a traditional PE program. The results indicated that children’s average time percentage spent in general content for the SPARK group was statistically lower compared to the traditional group over the 9-week intervention period within each sport activity. Additionally, children’s time percent in subject matter motor for the SPARK group was statistically higher compared to the traditional PE group across the entire intervention period within each sport type. Therefore, the hypotheses that the students in the SPARK group would have less time in general content, and significantly more time spent in subject matter motor compared to the traditional PE group were partially supported.

SPARK is designed to encourage and promote health-related fitness levels by maximizing physical activity participation and enjoyment in PE (24). The goal of SPARK is achieved by decreasing the time of transition between different class contents, using less direct instruction time, creating more opportunities of physical activity engagement and skill practices, and implementing relatively short but effective warm-up activities. Based upon the researchers’ observations, the traditional PE classes spent more time in class management compared to SPARK. This may have been due to more activity stations in the SPARK classes, where each student at each station had their own practice contents and purpose, which effectively decreases time spent in direct instruction. Compared to the traditional PE teacher, the SPARK teacher did not have to manage the class as frequently, as most of the students practiced and engaged in the lessons in a predetermined order. For example, the warm-up sections in the SPARK lessons were designed according to the lesson content. For example, the warm-up activities in the SPARK soccer lesson mainly focused on students’ lower body movements, and some warm-up activities encouraged students to interact with the soccer ball.

The traditional PE warm-up activities were relatively longer and repeated. For instance, the warm-up activities were almost identical for every single class in that the

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Descriptive statistics of percent time spent in the ALT-PE context level categories and subcategories across different sport activities in SPARK and traditional PE groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context level</td>
<td>Soccer (week 2–week 4)</td>
</tr>
<tr>
<td></td>
<td>Traditional</td>
</tr>
<tr>
<td>General content</td>
<td>49.18 %*</td>
</tr>
<tr>
<td></td>
<td>(30.32–75.36%)</td>
</tr>
<tr>
<td>Subject matter knowledge</td>
<td>18.06%</td>
</tr>
<tr>
<td></td>
<td>(8.87–31.38%)</td>
</tr>
<tr>
<td>Subject matter motor</td>
<td>32.76%</td>
</tr>
<tr>
<td></td>
<td>(0.00–14.01%)</td>
</tr>
</tbody>
</table>
| *, P<0.05; **, P<0.01. ALT-PE, academic learning time-physical education; SPARK, sports, play and active recreation for kids.
students started the class with running five to ten laps along the gymnasium or field, followed by static stretching, which was more time-consuming compared to the SPARK warm-up sections. This may explain the differences in average warm-up time percent between two PE groups. Additionally, most of the class breaks between skill practices and games in the SPARK lessons were in the format of slow walking or jogging compared to the traditional group where children engaged in sedentary activities (i.e. standing and sitting). Taken together, it was not surprising that the SPARK group spent significantly less time in general content compared to the traditional group.

The data related to subject matter knowledge for both groups in this study were relatively lower compared to the other two ALT-PE context level categories. Compared with the findings in the previous studies (12,13), the subject matter knowledge scores in the traditional school in this study were relatively high. For example, Derri et al. (13) found that the average time percent spent in subject matter knowledge among 110 elementary school students during 48 traditional PE lessons was 11.47%. Another study (12) revealed that the average time percent spent in subject matter knowledge among 23 grade six students during 18 cooperative learning team-handball lessons was 8.33%. Placek and Randall (18) examined and compared a sample of elementary students’ ALT-PE from over 49 classes including rope jumping, track and field, soccer, and kickball
and so on. The average time percent spent in subject matter knowledge for PE specialists and non-specialists were 8.8% and 5.2%, respectively. Thus far, most of the studies examining ALT-PE mainly focused on elementary students; few have investigated ALT-PE levels among middle school students. Therefore, the results of the current study augmented the literature in this field of inquiry.

Based upon the observations from the researchers, every SPARK class included two types of activities: (I) activities aimed to develop health-related fitness levels such as cardiovascular capability, muscular strength, and locomotor and non-locomotor skills; and (II) activities aimed to develop skill-related fitness levels such as speed, reaction and agility. Most of these activities were conducted in the format of skill practices, station fitness practices, and an amount time of game playing. In the SPARK classes, a relatively large amount of time (15–20 minutes) was spent in skill practices with the goal of promoting students’ health-related and sport-related fitness. For example, the PE teacher in the SPARK group usually observed students’ performances and reactions during the classes, and provided immediate feedback. In this way, students in the SPARK group would regularly be allowed to repeat skill practice if they found the respective practice enjoyable. In addition, some of the small-sided games in the SPARK classes were developed in the format of skill practices. For example, students were encouraged to throw the Frisbees into the hula-hoops set up by the PE teacher, with the purpose of developing students’ capability of throwing the Frisbee precisely.

In each of the sport types, the traditional PE classes included more game playing compared to the SPARK, because the nature of the traditional PE program is to develop students’ sport-related capabilities through implementing competitive scrimmages, small-sided-games, and full-sided games (25). Also, the time of the skill practice sections in the traditional PE classes were strongly deprived due to the high time percentage spent in games. Therefore, the children in the SPARK group spent significantly more time in skill practice, while less time in games as each of the sport types were explained.

By observing the classes in both PE groups, the researchers found that every SPARK class included activities with the purpose of promoting students’ fitness levels, such as the station fitness practice after the warm-up. Some other fitness activities between games were also regularly implemented during each class, an example being having a 1/4 mile running competition. In addition, short and easy fitness drills and practices were regularly conducted during classes, such as performing three to five push-ups if their team lost a game, and students who wish to take a water break being required to complete three to five jumping-jacks or squat-jumps. These fitness activities significantly promoted students’ time spent in the fitness category. The traditional PE lessons lacked fitness promotion drills or practices due to its sport-oriented nature. The most common fitness activities observed by the researchers in the traditional PE lessons emerged after the warm-up session, including pushups, lateral planks, and jumping jacks.

**Conclusions**

This study provides unique insights on the effectiveness of the SPARK health-related PE program on youths’ ALT-PE. The traditional school had a higher percentage of time spent in general content in PE compared to the SPARK...
school, but the SPARK had a greater percentage of time spent in subject matter motor and a significantly greater percentage of PE time spent in the subcategories of skill practice and fitness. No study to date has investigated the effect of SPARK on students’ ALT-PE compared to the traditional PE model. Therefore, this project provides unique insights on how SPARK can shift the emphasis of PE content to one of sport-related fitness (traditional) to one of health-related fitness (SPARK). This shift over time may provide students with more opportunity to increase physical activity behaviors in PE, and also the potential to increase cardiovascular fitness.

There are some limitations to this study that must be considered before any generalizations can be made. Firstly, only middle-school aged students were targeted in this study; therefore, the external validity of the results is questionable if generalized to younger or older grade levels. Second, the sample consisted of students who were primarily from high social economic status and the majorities were Caucasian, therefore the study was conducted on a homogeneous sample of youth that lacked ethnic diversity. Third, the SPARK intervention was of relatively short duration compared to previous research, therefore the results may have differed if a longer intervention exposure period was implemented.

Despite its limitations, this research supports the notion that the SPARK program is an effective pedagogical strategy to increase middle school children’s ALT-PE context levels in school PE settings. Physical educators may consider SPARK as an alternative instructional program in order to sustain elevated levels of ALT-PE in PE so that children can have a greater probability of achieving recommended daily physical activity amounts and fitness levels as suggested by various health agencies. Future research needs to implement the SPARK program over longer time frames in order to more accurately determine if SPARK can have a long-term effect on ALT-PE as compared to traditional instructional approaches. Considering the obvious health-related benefits of the context level of the ALT-PE, it is well worth the efforts for researchers to continue to examine effective instructional programs such as SPARK to increase these attributes and behaviors in PE settings. Doing so will manifest more effective pedagogical techniques to sustain healthy behaviors and ALT-PE in school settings.

Acknowledgements

The authors would like to thank the graduate research assistants that helped in the data collection process and the students who participated in this study.

Footnote

Conflicts of Interests: The authors have no conflicts of interest to declare.

Ethical Statement: Permission to conduct the study was obtained from the University Institutional Review Board (No. IRB_00061088), the school administrations, and the PE teachers prior to the start of this study.

References

10. Lee AM. How the field evolved. In: Silverman SJ, Ennis CD, editors. Student learning in physical education. Applying research to enhance instruction. Champaign, IL:


