Acute Effects of Classroom Exercise Breaks on Executive Function and Math Performance: A Dose-Response Study

Erin K. Howie, Jeffrey Schatz, and Russell R. Pate
University of South Carolina

Purpose: The purpose of this study was to determine the acute dose-response relationship of classroom exercise breaks with executive function and math performance in 9- to 12-year-old children by comparing 5-min, 10-min, or 20-min classroom exercise breaks to 10 min of sedentary classroom activity. Method: This study used a within-subjects experimental design conducted in the spring of 2012. Ninety-six 4th- and 5th-grade students in 5 classrooms in South Carolina were randomized to receive each of 4 treatments: 5-min, 10-min, or 20-min exercise breaks or 10 min of a sedentary lesson led by research staff. Students completed the Trail-Making Test, an Operational Digit Recall test, and a math fluency test immediately before and after each condition. Planned linear contrasts were used to compare posttest scores between conditions using a repeated-measures mixed model, adjusted for gender, classroom, and the time-varying pretest scores. Potential effect modifiers were added as interaction terms.

Results: Math scores were higher after the 10-min and 20-min exercise breaks compared with the sedentary condition (d ¼ 0.24, p ¼ .04, and d ¼ 0.27, p ¼ .02, respectively), and an interaction was observed with gender, IQ, aerobic fitness, and lower engagement in some of the conditions. There were no improvements in executive function tasks. Conclusions: A 10-min and 20-min classroom exercise break moderately improved math performance in students compared with a seated classroom lesson.

Keywords: academic achievement, children, cognition, school

Because many children are not meeting physical activity recommendations (Eaton et al., 2010), feasible strategies for increasing physical activity in children are important for a broad range of child outcomes. Yet, instead of increasing opportunities for physical activity, many schools have reduced physical activity opportunities in response to budget reductions and increased attention on standardized testing (Center on Education Policy, 2011). Paradoxically, reducing physical activity may decrease the academic achievement that schools are trying to improve (Biddle & Asare, 2011; Centers for Disease Control and Prevention, 2010; Howie & Pate, 2012; Roberts, Freed, & McCarthy, 2010; Ruiz et al., 2010). Recent reviews (Biddle & Asare, 2011; Centers for Disease Control and Prevention, 2010; Howie & Pate, 2012) and cross-sectional studies (Roberts et al., 2010; Ruiz et al., 2010) suggest a positive association between physical activity and academics, but more experimental trials are necessary. Recent experimental studies do support a positive relationship between physical activity and cognition (Davis et al., 2011; Hillman et al., 2009).

Of the cognitive abilities shown to improve with exercise, the strongest effects have been seen in executive function (Hillman, Kamijo, & Scudder, 2011; Tomporowski, Lambourne, & Okumura, 2011). Executive functions are higher-order complex cognitive processes including working memory, inhibition, and cognitive flexibility (Miyake et al., 2000). Executive function has been researched extensively in relation to learning disabilities, including attention-deficit hyperactivity disorder (ADHD), in clinical populations (Willcutt, Doyle, Nigg, Faraone,
& Pennington, 2005), and it has been shown to be highly predictive of academic achievement with early assessments of executive functions predicting later academic success (Gathercole, Brown, & Pickering, 2003).

However, many questions remain unanswered regarding the relationships between physical activity, executive function, and academic achievement, including the appropriate dose of physical activity required to produce optimal outcomes (Biddle & Asare, 2011). Time is a critical resource in schools; therefore, it is crucial to maximize efficiency in implementing physical activity throughout the school day. Schools have begun to implement classroom exercise breaks in their curriculum and practices. Preliminary findings suggest that exercise breaks may improve physical activity and cognitive effects (Donnelly et al., 2009). Yet, only three studies have examined the acute cognitive effects of these exercise breaks (of approximately 10 min) in children (Grieco, Jowers, & Bartholomew, 2009; Kubesch et al., 2009; Mahar et al., 2006). No previous studies have evaluated the differences in the acute cognitive effects of classroom exercise breaks of various durations in children.

The purpose of this study was to determine the acute dose–response relationship between 5-, 10-, or 20-min classroom exercise breaks and these cognitive functions in 9- to 12-year-old children. The exercise durations were based on prior research reporting improvements in executive functions and academic performance following 10 min and 20 min of physical activity (Hillman et al., 2009; Mahar et al., 2006). Schools and teachers, however, tend to use shorter durations of 5 min as this duration is more practical to implement in school practice. Currently, it is unknown if these shorter breaks have acute cognitive benefits. Finally, the study examined whether the relationship between duration of acute classroom exercise breaks and cognitive functions was moderated by gender, IQ, aerobic fitness, body mass index (BMI), behavior, school engagement, and/or level of participation during the intervention.

METHODS

Study Design

This study used a within-subjects experimental design. Students participated in each of the four conditions: 10 min of sedentary classroom activity, and 5-min, 10-min, and 20-min classroom exercise breaks. Each classroom participated in the intervention at a consistent time and day each week. To account for sequencing or practice effects, a balanced Latin Square design was used to randomize the four treatments at the classroom level. The primary dependent variables of executive function and math performance were assessed before and after each experimental condition. This pretest–posttest design was chosen to account for daily variation in cognitive abilities within each child. All participants provided parental consent and student assent. This study was approved by the institutional review board at the University of South Carolina and the research board of the school district.

Participants and Setting

Participants (n = 96) were aged 9 to 12 years old and were drawn from four fourth-grade and four fifth-grade classrooms at an elementary school in South Carolina. Students from two classrooms who provided consent were combined into a single classroom and the intervention was delivered to five classroom groups. All data were collected during the spring of 2012.

Intervention

Brain BITES (Better Ideas Through ExerciSe) was a simple classroom exercise break intervention developed for this study and was designed to maintain moderate-to-vigorous aerobic activity for the duration of the exercise break. Research staff led students in activities performed in minimal space including stationary marching with arm movements and various forms of jumping and running in place. Each break began with 30 s of low-intensity warm-up and ended with a brief deep-breathing and stretching cool-down. Activities performed were similar throughout the intervention; only the duration of activities varied. The control condition was an attention control, as the social interaction with research staff may influence performance on tests. To simulate a typical classroom environment, the research staff delivered a brief lesson about exercise science while the students remained seated.

To encourage participation and high-intensity activity, the instructor physically participated, gave verbal encouragement, which has been shown to increase child activity (Donnelly et al., 2009). Additionally, students were instructed on how to take heart rates using their carotid or radial pulse and aimed to get their heart rates to 150 beats per minute based on previous work that has shown this is an achievable goal for similarly aged students (Davis et al., 2007). The self-assessed heart rates were used solely as a motivational tool to increase the intensity of the exercise break. To objectively measure intervention fidelity, the exercise sessions were videotaped and observed as described below in the “Fidelity of Intervention” section.

Measures

Information was collected on potential confounding variables and factors that have been shown to influence the relationship between exercise and cognition. These included socioeconomic status (SES), gender, age, student
engagement, ADHD and problem behavior symptoms, IQ, BMI, daily physical activity, and aerobic fitness. Students completed a simple demographic questionnaire developed for the current study to obtain age, and they completed questions from a previously used school engagement questionnaire (e.g., I enjoy school/learning; O’Farrell & Morrison, 2003). Parents completed a questionnaire to obtain SES (household income, parent education) and the Conners’ Parent Rating Scales Revised short version, a 27-item checklist to assess ADHD and disruptive behavior symptoms (Conners, 2008). The Conners’ Parent Rating Scales has an internal reliability for the four subscales ranging from .816 to .944 for children ages 9 to 11 years old as published in the testing manual.

Prior to beginning the intervention, participants completed baseline aerobic fitness assessment conducted by research staff (the Progressive Aerobic Cardiorespiratory Endurance Run fitness test from the FITNESSGRAM testing battery administered by research staff during physical education class [Meredith & Welk, 2005]). Height and weight measures were taken by the school nurse within 1 month of the intervention. The Kaufmann Brief Intelligence Test-Second Version measure of abbreviated IQ was administered to each child individually by research staff, and composite IQ scores were used (testing manual internal reliability is .87 for the Riddles subscale and .84 for the Verbal subscale in 9-year-old children [Kaufman & Kaufman, 2004]).

**Fidelity of Intervention**

Videotapes of all four conditions were coded for intensity of physical activity using a modified System for Observing Fitness Instruction Time (SOFIT) as modified by Donnelly to capture activity occurring in the elementary school classroom (unpublished). Observations of participating individual children were made at consecutive 10-s intervals during the exercise or sedentary condition, not including cognitive testing. Each child’s average activity level during the 10-s interval was coded using a scale from 1 to 5 where 1 is equal to lying down and 5 is equal to being very active (e.g., running in place, jumping). Videos were viewed and coded three times for a total of 4,212 observations. To assess reliability, 10% of the intervals were recoded 4 months after the initial coding (α = .74 24). Intervals were randomly selected in groups of 10. The percent agreement was 91.0% with a weighted kappa of .95.

Each video was watched three times. Different participants were observed during each interval in each of the three viewings. Each participant was observed for an average of 16.8 intervals during the sedentary condition, 7.5 intervals during the 5-min exercise break, 12.5 intervals during the 10-min break, and 25.5 intervals during the 20-min break. Mean participant physical activity intensity scores were calculated for each condition.

**Cognitive Measures**

The testing battery was selected for age appropriateness and was assessed in pilot work to establish feasibility and acceptability.

**Trail-Making Test**

The Trail-Making Test (TMT) was selected as a valid, feasible, and appropriate measure of executive function for children in the target age range of this study (Lezak, Howieson, & Loring, 2004). Additionally, performance on the TMT has previously shown to be affected by exercise, but only in adults (Lambourne & Tomporowski, 2010). The TMT has reliability of .56 for 6 months in children aged 4 to 12 years old (Neyens & Aldenkamp, 1997). Two alternative forms, mirror images (to maintain path distance but alter the search pattern), were used to decrease practice effects. The test was modified to be self-timed for group administration. A subsample of participants (approximately 20%) was timed by researchers from the videotapes. The correlation between participant-timed and researcher-timed was .90 with a mean difference of 2.34 s (SD = 10.9). The TMT has two parts, Part A and Part B. Part A consists of connecting number sequences, while Part B involves alternating between numbers and letter sequences. As suggested by Sanchez-Cubillo et al. (2009), the difference between the TMT-A and TMT-B (TMT-BA) represents executive control. Therefore, TMT-BA was used as a measure of executive function in the current study (Sanchez-Cubillo et al., 2009). The correlation of the TMT-B with the TMT-BA in the current study was .94. The TMT-BA pretests had a one-way intraclass correlation of .65, which is considered good for the assessment of group-level outcomes (Nunnally & Bernstein, 1994).

**Digit Recall**

Operational digit recall is a validated measure of working memory (Gathercole et al., 2003). To modify the task for increased validity during group administration, participants were read a list of three to seven numbers (e.g., 5, 7, 3, 9) and then were given 5 s to write them in chronological order from the lowest to highest. The digit recall score was the sum of sequences the student answered correctly, adjusted for the length of the sequence. The digit recall pretests had an intraclass correlation of .63, which is considered good for the assessment of group-level outcomes (Nunnally & Bernstein, 1994).

**Timed Math Test**

To assess ecological validity with academic performance and provide a tangible outcome for teachers and school administrators, a timed math test was given, similar to a previous study (Maeda, 2003). Students completed as many grade-appropriate (based on state curriculum standards)
arithmetic problems as possible within 1 min. Such tests are a measure of math fluency and are considered a good indicator of individual differences in math ability (Durand, Hulme, Larkin, & Snowling, 2005). The math score was the number of problems correctly answered. The pretest math scores had an intraclass correlation of .95, which is considered excellent for the assessment of group-level outcomes (Nunnally & Bernstein, 1994).

Data Analysis

Descriptive statistics were calculated for the total group and for each gender using SAS Version 9.2. As the cognitive tests may be prone to practice effects, the Time x Exercise condition interaction was examined in an initial analysis of variance to test for order effects between weeks. This interaction was only statistically significant for the digit recall scores. However, when the raw scores were examined, the scores decreased rather than increased at one time point. Due to this lack of practice effects, coupled with the randomization to the order of conditions to counterbalance practice effects, the primary analyses included all classroom groups together, adjusted for group. Participants were included in the analysis if they completed at least two assessments.

To test a-priori hypotheses of interest, thereby reducing Type I error from testing hypotheses that are not of meaningful interest (including an omnibus F test), planned linear contrasts were made comparing each exercise break condition to the sedentary condition and the three exercise conditions combined with the sedentary condition (Howell, 2013). A repeated-measures mixed model (PROC MIXED) was used to compare posttest scores between conditions. Models were adjusted for gender, classroom group, and the time-varying covariate of pretest scores. This method of analysis was chosen to account for the within-subjects correlation in repeated measures, the ability to use all available data, and the ability to adjust for a time-varying covariate. Separate analyses were conducted for the dependent variables of TMT-BA, digit recall, and math scores. As simple effect sizes could not be calculated due to the importance of pretest measures for each condition, Cohen’s d effect sizes were calculated using the adjusted predicted means and standard errors were used to derive estimated standard deviations.

Finally, to test whether the effects differed by baseline student characteristics or by participation in the intervention, interaction terms were added to the model adjusting for classroom group, sex, and pretest scores. Associations between potential effect modifiers were tested using Spearman correlations. Potential effect modifiers of abbreviated IQ, aerobic fitness levels, BMI, behavior problems from the Conners Parent Rating Scales, and school engagement were split based on median scores. Median scores were used for exploratory effect modification analyses. However, the split for BMI corresponded with the 85th percentile, according to the Centers for Disease Control and Prevention BMI age distributions (U.S. Department of Agriculture/Agricultural Research Service Children’s Nutrition Center, 2003). The IQ and aerobic fitness score splits corresponded with the respective 50th percentiles. Interaction terms between the condition and the dichotomous effect modifier were added to the model in separate analyses.

RESULTS

A total of 96 students participated in the study, and 94 completed assessments. Demographics and baseline descriptive variables are shown in Table 1. The average physical activity intensity as coded by the SOFIT observation was higher during all three exercise conditions compared with the sedentary condition (sedentary/4.01, SD 4 0.05; 5-min 4 4.00, SD 4 0.43; 10-min 4 4.35, SD 4 0.33; 20-min 4 4.26, SD 4 0.37). There were no differences in intensity between exercise conditions.

Effect of 5 Min, 10 Min, or 20 Min of Exercise

The change in math scores was statistically higher after 10 min (estimated difference of 1.07, 95% CI [0.03, 2.12], p 4 .04) and 20 minutes (1.2, 95% CI [0.15, 2.26], p 4 .02) of exercise compared with the sedentary condition as seen in Table 2. The estimated effect sizes were d 4 0.24 and d 4 0.27, respectively. When the three exercise conditions were combined, math scores were statistically greater than after the sedentary condition (p 4 .02). There were no other statistically significant differences between any durations of exercise and the sedentary condition in digit recall scores or performance on the TMT-BA.

Effect Modification

Only the correlation between aerobic fitness and BMI was found to be moderately correlated (r 4 2.51, p 4 .001), and remaining correlations were not correlated to each other (r ranging from 2.18 to .17). Students who had higher aerobic fitness had higher math scores across the sedentary condition and all exercise conditions, including when adjusted for gender, race, parent education, and parent income (B 4 0.30, p 4 .02). To test whether student characteristics influenced their responses to the exercise, interaction terms between the exercise dose and student characteristics were tested. The only overall statistically significant interactions were between BMI and condition for digit recall (p 4 .01, students with lower BMI improved after 20 min, while students with higher BMI decreased performance after 5 min).
The results for the comparisons between the exercise doses and the sedentary condition for math scores are described in Table 3. The only differences in math scores were found by gender, IQ, and engagement. After 10 min and 20 min, girls had statistically significantly higher math scores than in the sedentary condition (d ¼ 0.37, p ¼ .01, and d ¼ 0.21, p ¼ .04), while boys had no statistically significant changes (d ¼ 0.04, p ¼ .80, and d ¼ 0.12, p ¼ .40). After 10 min, participants with lower IQ had higher math scores than in the sedentary condition (d ¼ 0.39, p ¼ .01). After 20 min, students with lower engagement had higher math scores than in sedentary condition (d ¼ 0.29, p ¼ .01). Analyses of the digit recall scores showed that after 5 min of exercise, students with lower aerobic fitness (d ¼ 0.35, p ¼ .01) and higher BMI (d ¼ 0.41, p ¼ .004) had lower digit recall scores compared with their scores after the sedentary condition. After 20 min of exercise, students with lower BMI (d ¼ 0.45, p ¼ .001) had higher digit recall scores compared with their scores in the sedentary condition. The only statistically significant differences in TMT-BA scores were for students with low engagement who decreased their performance after 5 min of exercise compared with their scores in the sedentary condition (d ¼ 0.41, p ¼ .01).

DISCUSSION

This is the first study to directly compare the acute effects of varying doses of classroom exercise breaks on acute cognitive effects. Although the current study did not find a significant overall effect between all four conditions, 10-min and 20-min classroom exercise breaks moderately improved math scores in students compared with a sedentary classroom lesson. These findings are largely consistent with previous research that has shown improvements in diverse measures of cognitive functions following at least 10 min of physical activity (Budde, Voelcker-Rehage, Pietrabyk-Kendziorra, Ribeiro, & Tidow, 2008; Hillman et al., 2009). Researchers have yet to see significant improvements in cognition or math performance with doses...
of less than 10 min, although few studies have examined these shorter durations (Kubesch et al., 2009; Maeda, 2003). In one of the few studies to directly compare multiple doses, Kubesch et al. (2009) found improvements in cognitive performance after a 20-min physical education class but found no improvements after a 5-min classroom exercise break.

In the current study, effects were seen in math scores but not in working memory or the TMT-BA. This may have been due to the much lower reliability (higher unexplained variation) in the operational recall and TMT-BA scores compared with the math test. However, the reliability of TMT-BA scores in this study was consistent with previous studies (Neyens & Aldenkamp, 1997). Executive function is a difficult construct to measure (Miyake et al., 2000); nonetheless, executive function may be the cognitive function most responsive to exercise (Hillman et al., 2011; Tomporowski et al., 2011). To counter the lower reliability of executive function measures, this study used a within-subjects design and included pretest measures for each condition, including using them as a covariate in the model. Additionally, very recent work has suggested that select executive functions may be more sensitive to acute physical activity, such as selective attention and inhibition, rather than working memory (Drollette, Shishido, Pontifex, & Hillman, 2012).

Many hypotheses exist for the mechanisms underlying improvements in cognitive performance after acute exercise (Hillman et al., 2011; Tomporowski et al., 2011). These mechanisms may respond differently to different doses, intensities, and types of physical activity, but a clear dose-response pattern has not yet emerged (Lambourne & Tomporowski, 2010). Castelli et al. suggested that time in high-intensity exercise may be necessary for cognitive improvements (Castelli, Hillman, Hirsch, Hirsch, & Drollette, 2011). While the current study measured intensity at the classroom level, future studies are needed to examine the effect of varying intensities of physical activity on cognitive functions.

The findings from this study suggest that different students may react differently to classroom exercise breaks. Previous studies have shown that numerous factors potentially influence the relationship between exercise and cognition (Hillman et al., 2009; Roberts et al., 2010). In the current study, participants with lower IQ, higher aerobic fitness, or lower school engagement had more improvement in math scores with the classroom exercise breaks. Additionally, girls had greater improvements in math scores. Classroom tracking, or the practice of grouping of students with similar academic abilities, allows for tailored recommendations to specific classrooms. For example, a classroom of students with lower academic ability may benefit from a 10-min classroom exercise break, while higher-achieving students may seek alternative physical activity opportunities such as an outdoor recess activity break. It is important to note, however, there were no differences in how students with higher BMI and poorer behavior responded, suggesting that classroom exercise breaks are appropriate for a wide range of students. This study is unable to examine the effect of individual intensity

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Effect Modification for Posttest Math Performance After 10 Min of Sedentary Classroom Activity or 5 Min, 10 Min, or 20 Min of Classroom Exercise Breaks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom Exercise Break Condition</td>
<td>Girls</td>
</tr>
<tr>
<td></td>
<td>Sedentary</td>
</tr>
<tr>
<td>Mean(SE)</td>
<td>24.3 (0.6)</td>
</tr>
<tr>
<td>ES (d)</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td>Mean(SE)</td>
</tr>
<tr>
<td>ES (d)</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>Mean(SE)</td>
</tr>
<tr>
<td>ES (d)</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>Mean(SE)</td>
</tr>
<tr>
<td>ES (d)</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td>Mean(SE)</td>
</tr>
<tr>
<td>ES (d)</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>Mean(SE)</td>
</tr>
<tr>
<td>ES (d)</td>
<td>.13</td>
</tr>
</tbody>
</table>

Note. Adjusted for classroom group, gender, and pretest scores; n = 4/4 for analysis.
*Statistically different than sedentary condition within strata at p < .05.
of physical activity on the cognitive effects. The SOFIT observation was used to obtain a classroom estimate of intervention fidelity, but it is not representative of an individual’s intensity. For the short duration of exercise in the classroom field setting, it was unpractical to use heart rate monitors, though future studies may want to evaluate the effect of differing intensities of exercise on acute cognitive effects.

This study was an efficacy study, implemented by research staff in the classroom environment. This approach ensured high implementation fidelity, with high participation in moderate-to-vigorous physical activity, but it is not easily sustainable. However, this intervention can be easily implemented by classroom teachers using few resources, and schools are implementing similar practices. Additionally, these results can only be generalized to similar classroom breaks as effects may differ by type or intensity of exercise (Pesce, Crova, Cereatti, Casella, & Bellucci, 2009). The different findings between the different outcome measures used also suggests that the findings on the effect of exercise on cognitive and academic performance may be highly reliant on the specific outcome measures selected. Researchers should carefully select multiple measures to represent a more complete representation of the constructs of cognitive or academic performance. The cumulative effect of these exercise breaks on academic performance over longer periods of time (e.g., a school year) was not addressed by this study, but longer-term outcome studies will be important for understanding the ultimate value of such breaks for student outcomes.

CONCLUSION

Although this study did not find cognitive improvements after 5 min of classroom exercise breaks, breaks of 10 min and 20 min were sufficient to elicit small improvements in math performance. Importantly, students participated in quality physical activity with numerous potential health benefits from all doses, and there were no detrimental effects on cognitive or academic performance.

WHAT DOES THIS ARTICLE ADD?

While many studies have examined the association between exercise and cognition in children, no previous studies have examined the dose-response relationship to identify the optimal duration required for positive effects. Our study is the first to directly study the dose-response relationship of classroom exercise breaks and field measures of cognitive effects (TMT, digit recall, and math performance) using a controlled, within-subject, experimental design. The current study revealed that 10 min and 20 min of acute classroom exercise breaks moderately improved performance on a math test, and 5, 10, or 20 min did not negatively affect performance on the math or executive function measures. The findings have immediate implications for teachers to implement classroom exercise breaks of at least 10 min to achieve potential academic benefits. Unfortunately, with strict school schedules and curricula, most exercise breaks currently being implemented in schools last less than 10 min. Additional training and resources may help teachers and administrators conduct 10-min classroom exercise breaks. If conducting classroom exercise breaks for at least 10 min is not feasible, schools can implement other physical activity opportunities of similar durations to receive moderate acute academic benefits.

ACKNOWLEDGMENTS

We sincerely thank the school district, principal, teachers, parents, and students who made this study possible.

FUNDING

This study was funded by a Doctoral Research Grant from the American College of Sports Medicine.

REFERENCES


