Cognitively engaging chronic physical activity, but not aerobic exercise, affects executive functions in primary school children: A group-randomized controlled trial

Mirko Schmidt, Katja Jäger, Fabienne Egger, Claudia M. Roebers, and Achim Conzelmann

University of Bern

Accepted manuscript: October 3, 2015

Mirko Schmidt, Institute of Sport Science, University of Bern, Bremgartenstrasse 145, 3012 Bern, Switzerland, mirko.schmidt@ispw.unibe.ch; Katja Jäger, Department of Psychology, University of Bern, Hochschulzentrum vonRoll, Fabrikstrasse 8, 3012 Bern, Switzerland, katja.jaeger@psy.unibe.ch; Fabienne Egger, Institute of Sport Science, University of Bern, Bremgartenstrasse 145, 3012 Bern, Switzerland, fabienne.egger@ispw.unibe.ch; Claudia M. Roebers, Department of Psychology and Center of Cognition, Learning and Memory, University of Bern, Hochschulzentrum vonRoll, Fabrikstrasse 8, 3012, Switzerland, claudia.roebers@psy.unibe.ch; Achim Conzelmann, Institute of Sport Science, University of Bern, Bremgartenstrasse 145, 3012 Bern, Switzerland, achim.conzelmann@ispw.unibe.ch.

This study was supported by a grant (#13-01) from the Federal Office of Sport (FOSPO) Magglingen, Switzerland.

We would like to thank the participating teachers and students and we also gratefully acknowledge the help of Josias Bucher, Claudia Buser, Laura Dapp, Caroline Köchli, Seraina Röllin, and Stefan Wieland with the data collection.

Correspondence concerning this article should be addressed to Mirko Schmidt, Institute of Sport Science, University of Bern, Bremgartenstrasse 145, 3012 Bern, Switzerland, mirko.schmidt@ispw.unibe.ch; Phone number: +41 31 631 83 52.

Abstract

Although the positive effects of different kinds of physical activity (PA) on cognitive functioning have already been demonstrated in a variety of studies, the role of cognitive engagement in promoting children’s executive functions is still unclear. The aim of the present study was therefore to investigate the effects of two qualitatively different chronic PA interventions on executive functions in primary school children. 181 children aged between 10 and 12 years were assigned to either a 6-week physical education program with a high level of physical exertion and high cognitive engagement (*team games*), a physical education program with high physical exertion but low cognitive engagement (*aerobic exercise*), or to a physical education program with both low physical exertion and low cognitive engagement (*control condition*). Executive functions (updating, inhibition, shifting) and aerobic fitness (multistage 20-meter shuttle run test) were measured before and after the respective condition. Results revealed that both interventions (team games and aerobic exercise) have a positive impact on children’s aerobic fitness (4-5 % increase in estimated VO_2max). Importantly, an improvement in shifting performance was found only in the team games and not in the aerobic exercise or control condition. Thus, the inclusion of cognitive engagement in PA seems to be the most promising type of chronic intervention to enhance executive functions in children, providing further evidence for the importance of the qualitative aspects of PA.

Keywords: cognition, chronic exercise, physical education, intervention

Cognitively engaging chronic physical activity, but not aerobic exercise, affects executive functions in primary school children: A group-randomized controlled trial

The beneficial effects of regular physical activity (PA) on physical and mental health throughout life are well documented (Penedo & Dahn, 2005). Besides these effects, there has been growing evidence in recent years that children's cognitive functions also benefit from chronic PA (Hillman, Kamijo, & Scudder, 2011; Khan & Hillman, 2014; Sibley & Etnier, 2003; Verburgh, Königs, Scherder, & Oosterlaan, 2014). This empirical evidence is what almost all school-based PA programs refer to when justifying daily physical education lessons or enhanced PA at school (e.g., Let's Move! Active Schools). To date, however, the question as to which specific type of chronic PA affects which specific cognitive ability remains unanswered. Most studies have targeted quantitative aspects of PA, i.e., they have focused on aerobic activity without regard for the cognitive demands of the PA itself. Investigations addressing qualitative aspects of PA are very rare and the impact of *cognitive engagement*[1] during PA on cognitive outcomes is largely unknown (Pesce, 2012). The answer to this question of the specificity of various forms of PA could be of great practical importance in the educational setting and in particular for designing school-based PA programs or selecting specific physical education contents that target the promotion of cognitive performance in school.

Studies examining the effects of chronic PA on children's cognition have used various measures of cognitive performance to test the whole range from more global components of cognition, e.g., intelligence and academic achievement, to more specific components of cognition, e.g., memory and executive functions (Tomporowski, Davis, Miller, & Naglieri, 2008). In children, as in adults (Colcombe & Kramer, 2003), the effects

---

[1] In order to distinguish it from behavioral and emotional engagement, cognitive engagement can be defined as the degree to which the allocation of attentional resources and cognitive effort is needed to master difficult skills (Tomporowski, McCullick, Pendleton, & Pesce, 2015).

were largest when specific cognitive processes such as executive functions (EFs) were targeted (Tomporowski, Lambourne, & Okumura, 2011). The umbrella term EFs refers to a multi-component construct that consists of several distinct, yet interrelated, processes which encompass the cognitive processes responsible for controlling and organizing goal-directed behavior (Anderson, 2008; Zelazo & Carlson, 2012). EFs have been shown to predict school readiness in young children (Blair & Diamond, 2008; Roebers et al., 2014) and to account for substantial amounts of variance in individual differences in school achievement in elementary school children (Bull, Espy, & Wiebe, 2008; Willoughby, Blair, Wirth, & Greenberg, 2012). Testing the possibility that PA may positively influence EFs is therefore of central and practical importance in the educational field. As suggested by Miyake et al. (2000) and widely established in research (Diamond, 2013), EFs can be divided into three subdimensions: updating (keeping relevant information in working memory and processing this information further), inhibition (the ability to avoid dominant, automatic, or prepotent responses or to resist distractor interference), and shifting (shifting attention back and forth between multiple tasks, operations, rules, or mental sets). Studying these three different EF subdimensions separately seems reasonable because it might increase our understanding of EFs and their development. In fact, the measurement of highly specific constructs might be more sensitive than global measurements as a means of detecting changes (Tomporowski et al., 2008). However, most research in the field focuses on inhibition and, unfortunately, studies assessing all three subdimensions simultaneously while often called for have rarely been documented so far (Barenberg, Berse, & Dutke, 2011).

In the search for theoretical explanations concerning the underlying mechanisms contributing to the positive effects of chronic PA on cognitive performance, two different hypotheses seem to dominate the literature: the *cardiovascular fitness hypothesis* (North, McCullagh, & Tran, 1990) and the "*cognitive stimulation hypothesis*" – labeled according to

Tomporowsi et al. (2008), Best (2010), and Pesce (2012). As the name indicates, the cardiovascular fitness hypothesis assumes that it is the increased cardiovascular fitness, caused by regular PA, which mediates the relationship between PA and EFs. Although cross-sectional findings consistently find fitter children to display better EFs than less fit children (Castelli, Hillman, Buck, & Erwin, 2007; Chaddock, Hillman, Buck, & Cohen, 2011; Pontifex, Scudder, Drollette, & Hillman, 2012), and the few existing interventional studies show that aerobic training results in increased aerobic fitness and improved EFs (Davis et al., 2007; Davis et al., 2011; Kamijo et al., 2011), the cardiovascular fitness hypothesis was not supported by a meta-regression analysis systematically examining the relationship between aerobic fitness and cognitive performance (Etnier, Nowell, Landers, & Sibley, 2006). However, the results of these studies only appear to contradict each other, because whereas the differences found in the cross-sectional studies can also be explained by potential confounding variables, such as differences in educational or nutritional habits, the effects reported in the aforementioned intervention studies need not necessarily be interpreted as a causal effect of the increased aerobic fitness, but may in certain circumstances be a result of other factors associated with these age-appropriate PA programs. Obviously, none of the intervention programs consisted of "pure" aerobic exercise training, but involved group activities such as playing tag, soccer or basketball. In that sense, enhanced social interactions or cognitive engagement through learning new skills could have affected children's EFs in the experimental group. It might therefore be worthwhile "to focus on variables other than aerobic fitness" (Etnier et al., 2006, p. 126) when investigating the relation between PA and EFs.

One factor that is often overlooked seems to play a central role in the relationship between PA and cognitive performance: the cognitive demands inherent in many forms of PA (Best, 2010). The assumption of cognitive stimulation hypothesis is that coordinatively demanding and non-automated sport-related activities activate the same brain regions that are used to control higher-order cognitive processes (Best, 2010; Diamond & Lee, 2011). Based on this theoretical assumption of shared information processes in both motor and cognitive control (Roebers & Kauer, 2009), this hypothesis explains intervention effects in terms of the specific activation of these processes during PA, which then leads to cognitive benefits in these circumscribed domains of executive functioning. Sedentary cognitive training using adaptive cognitively complex computerized games has repeatedly been shown to promote children's EFs (Bergman Nutley et al., 2011; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; Thorell, Lindqvist, Nutley, Bohlin, & Klingberg, 2009). Imagining that PA can be designed to be cognitively demanding and constantly adapting to the actual performance of the participating children, one could hypothesize that this combination would produce effects beyond those of pure aerobic exercise.

To activate and exercise higher-order cognitive processes through PA, Tomporowski, McCullick, and Pesce (2015) suggest applying three principles of mental engagement: contextual interference, mental control, and discovery. Contextual interference arises, for example, when the context and conditions of a PA game are constantly changed and unpredictable sequences of actions have to be performed. Mental control can be induced when PA games are chosen which challenge specific EF subdimensions, such as updating, inhibition and shifting. The principle of discovery can be applied in open-ended games, in which emerging movement problems can be solved in multiple ways. Those two chronic PA studies that used one or more of these three principles in their interventions appear to corroborate the cognitive stimulation hypothesis. In a recent study, Pesce et al. (2013) showed that the receptive attention, defined as the ability to shift attention between different stimulus dimensions, of 5–10-year-old typically developing children improved after a 6-month intervention with cognitively enriched PA, while an intervention without a special focus on cognitive demands produced no effects. Further, Crova and colleagues (2014) showed that a 6-month physical education program including cognitively demanding PA benefitted inhibition in 9–10-year-old overweight children, whereas a curricular physical education program did not. Taken together, according to the cognitive stimulation hypothesis, PA interventions should be cognitively demanding to challenge higher-order cognitive processes (Pesce, 2012).

Contrasting these two hypotheses, the question arises what kind of PA is most suitable for improving the cognitive performance of children. Interventions consisting of coordinatively demanding, child-appropriate and playful contents certainly meet the curricular requirements of physical education better than pure endurance exercises (National Association for Sport and Physical Education, 2004). From a practical point of view, therefore, PA interventions that could be implemented during physical education lessons and that nevertheless have a positive effect on the children’s cognition would be more appropriate for the school setting. Interestingly, aside from acute PA interventions (Best, 2012; Budde, Voelcker-Rehage, Pietraßyk-Kendziorra, Ribeiro, & Tidow, 2008; Gallotta et al., 2012; Jäger, Schmidt, Conzelmann, & Roebers, 2014; Pesce, Crova, Cereatti, Casella, & Bellucci, 2009), attempts to systematically use cognitively challenging PA to promote EFs are rarely found (Crova et al., 2014; Pesce et al., 2013) and to the best of our knowledge no study has ever combined and contrasted the two hypotheses in a single study design.

Given large inter-individual differences in many physical and personality traits, it seems reasonable to assume that not every individual will profit to the same extent from the same PA, and empirical evidence supports this claim. For example, a meta-analysis (Chang Labban, Gapin, & Etnier, 2012) and three recent experimental studies (Chang et al., 2014; Hogan et al., 2013; Jäger, Schmidt, Conzelmann, & Roebers, 2015) indicate that subjects with higher fitness seem to benefit more from acute exercise in terms of cognitive performance.

Thus, physical fitness could be a potential moderating variable for the effects of chronic PA interventions on children's EFs. Turning to cognitive performance as a moderating variable, the studies on differential effects of PA on EFs have produced mixed results: for the most part, children with the poorest baseline performance were found to benefit most from acute and chronic PA interventions (Diamond & Lee, 2011; Drollette et al., 2014; Sibley & Beilock, 2007), whereas a recent study has revealed that only children displaying higher academic achievement (and/or higher physical fitness) profit from an acute PA intervention in a "real-world setting" (Jäger et al., 2015). Considering the previously mentioned intervention studies which primarily report effects on overweight children (Crova et al., 2014; Davis et al., 2007; Davis et al., 2011) and the fact that children's BMI itself is associated with variables such as their PA level, physical fitness, and maturation (Armstrong, 2013; Armstrong, & Welsman, 2001), an exploratory analysis of these potential moderators might help to explain which individuals may benefit most from which specific PA interventions.

Based on the literature presented and the as yet unanswered questions, the aim of the present study was to investigate the effects of two qualitatively different chronic PA interventions with diverging amounts of cognitive engagement on the three EF subdimensions (updating, inhibition, & shifting, defined by Miyake et al., 2000) in prepubertal children. Therefore, a physically demanding intervention focusing on augmenting children's aerobic fitness (aerobic exercise condition) and one with the same aim but with large amounts of cognitive engagement on top of physical exertion (team games condition) were developed and compared with a curricular physical education program without specific targets (in terms of its physical or cognitive demands). Three different EF subdimensions were included because there is a particular paucity of results regarding the effects of PA on updating and shifting in children and because it is not yet clear whether certain aspects of EFs can be affected more easily than others.

Method

Design

Two 6-week interventions in physical education with a high level of physical exertion but different amounts of cognitive engagement were compared to a control condition with respect to their effects on children's EFs: *team games* with high amounts of both cognitive engagement and physical exertion, *aerobic exercise* with low cognitive engagement and high physical exertion, and the *control condition* meeting the curricular requirements, with both low cognitive engagement and low physical exertion. Altogether, twelve classes were randomly assigned to one of the three experimental conditions. The teachers were informed about the basic aims of the study, but were blinded with respect to the specific hypotheses. As usual in classical comparison group pretest-posttest designs, each intervention was preceded and followed by a measurement point for data collection (of the dependent variables: EFs and aerobic fitness). Prior to the intervention, information about PA level, pubertal and socioeconomic status was collected using questionnaires, heights and weights (for calculating the body mass index, BMI) were measured, and academic achievement (math, language) was tested using three standardized tests.

Subjects

A total of 181 children ranging from 10 to 12 years of age (*M* = 11.35 years, *SD* = .60; 54.9 % girls) from the region of Bern, Switzerland, participated in the study. To maximize the generalizability of the results, the eight children with a diagnosed attention deficit hyperactivity disorder (ADHD) were included in the study. Throughout the entire study and during cognitive testing, these participants took their medication as usual. There was some loss of data due to sick leave, non-participation in the aerobic fitness test because of injury, or incomplete questionnaires. The percentage of pupils with incomplete values was 6.3% at pre-test and 8.6% at post-test. Since the MCAR test according to Little was not significant (χ^2 (131) = 155.38, p = .072), the resulting missing values were imputed with the help of the expectation–maximization (EM) algorithm, so that it was possible to work with a complete set of data.

There were no significant differences between the three experimental conditions with respect to age (F(2, 178) = .09, p = .915, η_p^2 = .001), gender distribution (χ^2(2) = 2.91, p = .233, Cramer's V = .127), children with ADHD (χ^2(2) = 1.40, p = .496, Cramer's V = .088), PA level (F(2, 178) = .67, p = .513, η_p^2 = .007), pubertal status (F(2, 178) = .61, p = .542, η_p^2 = .007), socioeconomic status (F(2, 178) = .95, p = .388, η_p^2 = .011), zBMI (F(2, 178) = 1.72, p = .182, η_p^2 = .019), and academic achievement (F(2, 178) = 2.32, p = .101, η_p^2 = .025). Both the principals of the schools and the parents of the children signed an informed consent form approved by the Institutional Review Board prior to participating in the study. All the children were asked before the first data collection session whether they wanted to participate and informed that they could discontinue at any time during the study. All data were treated confidentially.

General Procedure

The interventions were carried out by the regular physical education teachers of the respective classes in two physical education lessons per week, over a period of 6 weeks. Hence, the entire intervention extended over 12 lessons (45 minute each). Prior to the study, teachers completed a half-day training program instructing them in the basic principles, aims and purposes of the intervention program and demonstrating the specific contents with special teaching materials. To test the exposure component of implementation fidelity (Dane & Schneider, 1998), teachers had to report the number of lessons effectively carried out. Program exposure seems to have been as intended, since the teachers of both interventions reported that they had implemented M = 11.13 of the prescribed 12 lessons (range = 11–12). The number of physical education lessons carried out during the six weeks did not differ between the three conditions ($M_{team\ games}$ = 11.25, *SD* = 0.5 vs. $M_{aerobic\ exercise}$ = 12.3, *SD* = 0.5 vs. $M_{control\ condition}$ = 11.25, *SD* = 0.5; *F*(2, 9) = .50, *p* = .622, η_p^2 = .10).

Manipulation check variables were collected by two graduate students (one male, one female) of developmental psychology during weeks 3 and 4 of the experiment (i.e. during lessons 5-8). The students were familiar with the construct of EFs, but blinded with respect to the experimental conditions. In order to cover the entire scope of the interventions, for each condition, a different class was chosen for the manipulation check for each of the four lessons 5, 6, 7 and 8. The chosen class can be considered to be representative, as the three ANOVAs for the three different experimental conditions, with class as the independent variable and children's heart rate during "one PE lesson" as the dependent variable, showed that the heart rate did not differ between different lessons of the same condition: team games (*F*(3, 65) = 1.77, *p* = .162, η_p^2 = .084); aerobic exercise (*F*(3, 53) = 1.21, *p* = .315, η_p^2 = .062); control condition (*F*(3, 51) = 1.31, *p* = .283, η_p^2 = .076). The two observers were in the same class at the same time, fitted all the children with heart rate belts to assess their mean heart rate during the selected lesson, assessed physically active time of the group of children, and rated the involvement of the three EF subdimensions in every observed task the children had to carry out.

The children completed the same cognitive testing, which took place between 10.00 a.m. and 12.00 p.m. for all participants, before (pre-test) and after (post-test) the intervention. During the two lessons before this testing they were taught language skills in their mother tongue. The cognitive testing took place in a quiet room in small groups of four children. First, one investigator gave some general instructions. The children were encouraged to work quietly but to ask questions about the test whenever something was not clear. All the cognitive tasks were then completed on a computer and children received the instructions both in writing on the screen and simultaneously over headphones, which at the same time served as sound absorbers. Children were placed so that they could not see the screens of the other children and they completed the tasks at their own pace. The investigator was present during the entire testing procedure, but was blind with respect to the experimental condition to which the participants had been allocated.

Experimental Conditions[2]

Team games (high cognitive engagement, high physical exertion; *n* = 69): This intervention consisted of specifically designed team games (floorball and basketball) tailored to challenge EFs. First, these two team games were chosen because according to ACSM exercise guidelines (American College of Sports Medicine, 2010) they are appropriate to induce moderate to vigorous PA intensity, which should promote aerobic fitness when performed regularly. Second, both team games contain large amounts of prospective control and complex eye-hand coordination, and require goal-directed behavior. Third, these team games were suitable for combining sport-specific skill development (as required by the curriculum) with enriched cognitive engagement. Based on the three principles of mental engagement elaborated by Tomporowski, McCullick, and Pesce (2015), the second principle *mental control* was systematically used to increase children's cognitive engagement while playing these team games. For example, while children were playing basketball, the teacher suddenly blew the whistle, meaning that some rules of the game changed immediately. In order to ensure that such rule changes remained cognitively demanding, from the moment when the children learned to adapt their actions to the acoustic signal, it was later linked to an additional visual signal. For example, the combination of hearing a whistle and seeing a red card meant a change in the rules whereas the combination with a green card meant that the learned rules remained in force. Besides the focus on skill development in these two team games, each lesson started with a warm-up, which was characterized by high demands on

---

[2] Detailed information and the precise schedule of the interventions can be obtained from the authors.

EFs. For example, a form of tag was played, in which the children had to keep in mind different rules, react appropriately to acoustic cues, inhibit prepotent movements, as well as shift between different situations and rules.

Aerobic exercise (low cognitive engagement, high physical exertion; *n* = 57): This condition consisted of different group-oriented and playful forms of aerobic exercises, whose main aim was to promote children's aerobic fitness. Although it is not possible to exclude cognitive engagement entirely from chronic PA interventions, the attempt was made to choose exercises that were not cognitively demanding. For example, the children were instructed to run a marathon as an entire class, whereby each child was allowed to cross off one box from a joint list after each circuit. With a circuit of 200 m, it was therefore necessary to cross off 211 boxes in total in order to achieve the joint goal. Music was played in the background. Such exercises were chosen so that the motivation of the children could be maintained for as long as possible for an aerobic exercise training and to guarantee that this kind of PA did not differ from the PA in the team games intervention regarding physical intensity and/or the amount of social interaction.

Control condition (low cognitive engagement, low physical exertion; *n* = 55): Teachers in the control group continued to teach according to the national curriculum for physical education (Federal Office of Sport, 1997). This curriculum requires that the five areas fitness, athletics, gymnastics, dance, and team games are given a balanced amount of time during the lessons. To check whether the teachers in the control group adhered to these requirements and did not devote a disproportionate amount of time to one of the areas, they had to document the contents and goals of their teaching in a table every week. The analysis of these tables showed that they carried out their physical education in line with the requirements of the national curriculum.

Manipulation check variables

In order to estimate the *physical exertion* in the three experimental conditions, first, the children’s heart rate (HR) was measured using Suunto Dual Comfort Belts®. These belts transmitted the children’s heart rate wirelessly to a laptop, where the data was monitored and saved in real-time for each participant. The mean heart rate was used in the analyses. Second, the physically active time was recorded by two independent observers using the method of event sampling (Reis & Gable, 2000). Activity codes 1 to 5 were used to denote lying down, sitting, standing, walking, and being very active (McKenzie, Sallis, & Nader, 1992). To calculate the physically active time, only the time spent on activities coded with a 4 or a 5 was considered. Interrater reliability was assessed using two-way, mixed consistency, average-measures intraclass correlation coefficients (ICC). The resulting ICC was in the excellent range, $ICC = .98$ (Cicchetti, 1994), indicating a high degree of agreement between the two observers. The mean of the two recorded times in minutes was used in the analyses.

Cognitive engagement was rated by the same two aforementioned observers. They rated the involvement of the three EF subdimensions in every task the children had to perform on a 4-point Likert scale (*1* = not used at all, *2* = slightly used, *3* = moderately used, *4* = highly used). Updating was coded, for example, when children had to keep in mind relevant information or rules to fulfill a given motor task. Inhibition was coded, for example, when children had to interrupt an already initiated motor response to an external stimulus. Shifting was coded, for example, when children had to disengage attention from one cue and refocus on a different one. The rating of the involvement of each EF subdimension was multiplied by the measured time (in minutes) during which the respective task was performed. The ICCs for all three EF subdimensions were in the excellent range, $ICC_{updating} = .84$, $ICC_{inhibition} = .88$, $ICC_{shifting} = .89$ (Cicchetti, 1994), again indicating a high degree of agreement between the two raters. The mean of the two raters’ products was used in the analyses.

Cognitive Assessment

EFs were measured in two computer-based tasks using E-Prime Software (Psychology Software Tools, Pittsburgh, PA). Each task took about 10 minutes to complete and the order of the two tasks was counterbalanced between participants. *Updating* was assessed by means of a non-spatial n-back task (adapted from a spatial n-back task in Drollette, Shishido, Pontifex, & Hillman, 2012). Several pictures of fruit were presented one after another on the screen. Children were instructed to press the right button in front of them when the fruit on the screen was similar to the second to last fruit presented (target) and the left button in all other cases (non-targets). The task consisted of two test blocks containing 24 trials each, with one third of all trials being targets. The total number of correct answers was used as the dependent measure.

Inhibition was measured by means of a child-adapted Flanker task (Jäger et al., 2014) consisting of a block with 20 congruent trials (“pure” block) and a block with 20 congruent and 20 incongruent trials in a randomized order (“standard” block). The conflict score between trials with the highest rate of distraction (incongruent trials standard block) and trials with the lowest rate of distraction (congruent trials pure block) was calculated as the dependent measure for inhibition (Rueda, Posner, & Rothbart, 2005).

Shifting was assessed by an additional block (“mixed” block) included in the Flanker task (Jäger et al., 2014). In this block, again, 20 congruent and 20 incongruent trials were shown and an additional rule was introduced – cued by a different color of the trials. Children had to adapt their response depending on the color of the trials, requiring a switch between the two rules whenever the color of the trials changed. Global switch costs were calculated as the dependent variable (Chevalier & Blaye, 2009). Since trials in the mixed block not only required the child to shift between different tasks, but also contained inhibitory demands, the difference between the mixed and the standard block was calculated to control for the inhibition component.

Aerobic Fitness Assessment

Children's aerobic fitness was assessed using the Multistage 20-Meter Shuttle Run test (Léger, Mercier, Gadoury, & Lambert, 1988). Subjects have to run back and forth on a 20 m course and touch the 20 m line with their foot, and at the same time, a sound signal is emitted from a pre-recorded tape. The frequency of the sound signal increases by 0.5 km/h every minute, indicating the next stage (level) and starting with a speed of 8.5 km/h. The test ends when subjects fail to reach the line before the signal. Maximal oxygen uptake (VO_2max; $mL \cdot kg^{-1} \cdot min^{-1}$) was estimated from the number of the last stage reached as: $[31.025 + (3.238 \cdot \text{velocity}) – (3.248 \cdot \text{age}) + (0.1536 \cdot \text{age} \cdot \text{velocity})]$. Evidence for the reliability and validity of the 20 m shuttle run test has been provided by Liu, Plowman, and Looney (1992) and McVeigh, Payne, and Scott (1995).

Background Variables

The Physical Activity Questionnaire for Children (PAQ-C; Crocker, Bailey, Faulkner, Kowalski, & McGrath, 1997) was used to measure general levels of PA. The PAQ-C is a 7-day self-administered recall measure that provides a summary PA score derived from nine items. The response format varies by item, but each is scored on a 5-point scale, a sample item being: "In the last 7 days, on how many evenings did you do sports, dance, or play games in which you were very active?" Response options range from: "None" (1 point) to "6 or 7 times last week" (5 points). Evidence for the reliability and validity of the questionnaire in 8- to 16-year-olds has been provided by Crocker et al. (1997).

The German version (Watzlawik, 2009) of the Pubertal Development Scale (PDS; Petersen, Crockett, Richards, & Boxer, 1988) was used to assess *pubertal status*. For each gender, three questions are used to determine the pubertal status, a sample question for boys being: “Have you noticed a deepening of your voice?” Response options were: not yet started (1 point); barely started (2 points); definitely started (3 points); seems complete (4 points). The puberty index (ranging from 3 to 12) was calculated from the sum of the three items. Evidence for the reliability and validity of the German version used in 9- to 13-year-olds has been provided by Watzlawick (2009).

The Family Affluence Scale II (FAS II; Boudreau & Poulin, 2009) was used to assess the *socioeconomic status*. The scale consists of 4 questions asking children about things they are likely to know about in their family (number of family-owned cars, computers, number of family holidays in the past year, and having an own bedroom at home). A sample item is: “Does your family own a car, van or truck?” Response options are: no (0 points); yes, one (1 point); yes, two or more (2 points). The response format varies by item. The prosperity index (ranging from 0 to 9) was calculated from the sum of the three items. Evidence for the reliability and validity has been provided by Boudreau and Poulin (2009).

The *BMI* was calculated as the body weight (in kg) divided by the square of the height (in m). As recommended by Field et al. (2003), age- and gender-specific *z* scores of BMI (zBMI) were used in all statistical analyses.

Academic achievement (math, language) was assessed using three standardized academic achievement tests. Math performance was measured using the three subscales (arithmetic, geometry, and solving written math problems) of the German math test for 5th graders (*Deutscher Mathematiktest für fünfte Klassen DEMAT 5+*; Götz, Lingle, & Schneider, 2013). Writing was assessed using the *Hamburger Schreib-Probe 1-10* (*HSP 1-10*; May, 2012) and reading using the *Salzburger Lese-Screening für die Klassenstufen 5-8* (*SLS 5-6*; Auer, Gruber, Mayringer, & Wimmer, 2005). The correlations between the three tests were all significant (*r* between .47 and .55). The z-standardized values of the three tests were aggregated to form a general academic achievement score.

Statistical Analyses

Outlier analysis: Trials with a reaction time under 150 ms were excluded (interindividual outliers, Flanker task: 0.1%, N-Back task: 0.1%). In a next step, trials with reaction times deviating by more than 3 *SD* from the child's mean (intraindividual outliers) were excluded as well (Flanker: 1%, N-Back: 0.4%.). Only correct trials were included in the calculation of reaction times. Subsequently, blocks with an accuracy of less than or equal to 50% were deleted (N-Back: 2.92%, Flanker: 0.73%) because those children seemed to have either not understood the task or to have done it incorrectly due to a lack of motivation.

To account for the hierarchical data structure due to the children being clustered within classes, multilevel analyses were conducted (using the mixed models module of the Statistical Package for Social Sciences; SPSS 21.0). Since each class came from a different school, a two-level structure was applied, with children (*n* = 181) at the first level and class (*n* = 12) at the second level. To test whether the full model including class as a second-level factor fitted the data significantly better than the model in which only the intercepts were included, a χ^2 difference test was used, with - 2 Log Likelihood as the information criterion. In *preliminary analyses*, potential between-group differences in background variables (physical activity level, pubertal status, socioeconomic status, zBMI, academic achievement) and in pre-test values of the dependent variables (updating, inhibition, shifting, estimated VO_2max) were tested using multilevel analyses. Due to convergence problems, univariate analyses of variance were conducted in the manipulation check analyses (Tabachnick & Fidell, 2013). Partial eta square (η_p^2) was reported as an estimate of effect size. When the overall ANOVA proved significant, Bonferroni-corrected post-hoc comparisons were used to determine the specific differences between the three groups. In the *main analyses*, multilevel analyses were conducted. When the main fixed effect was significant, Bonferroni-corrected post-hoc comparisons were reported. To test the more explorative assumption that not every individual benefitted equally from the same intervention, all background variables and the children's aerobic fitness were inserted into each multilevel model as a separate covariate. The level of significance was set at $p < .05$ for all analyses.

Table 2 shows means and standard deviations for accuracy and reaction times in the different blocks of the Flanker task at pre- and post-test for the three groups. Because there was an expected ceiling effect concerning accuracy in the Flanker task (mean accuracy was between 86% and 98%; see Table 2 for raw data), the mean reaction times were included in the subsequent analyses.

Results

Preliminary analyses

At pre-test, the multilevel analyses revealed no significant group differences in background variables (physical activity level, pubertal status, socioeconomic status, zBMI, academic achievement), cognitive performance (updating, inhibition, shifting) nor in aerobic fitness (estimated VO_2max), Fs(2, 178) < 2.89, n.s. Aside from inhibition, χ^2 (1, N = 181) = 4112.28 – 4111.63 = .65, $p > .05$, the full model was significantly better than the intercepts-only model in all tested variables, e.g., PA level, χ^2 (1, N = 181) = 4913.05 – 4902.23 = 10.82, $p < .05$. This indicates that being part of a certain class explains a notable part of the variance in the data, and multilevel analyses therefore seem justified.

To test whether the *physical exertion* in the three experimental conditions differed, the three groups were compared with regard to the physically active time and the heart rate during the one tested lesson. Even if the physically active time (F(2, 21) = .52, p = .604, η_p^2 = .047) did not differ between the three groups (see Table 1), the mean heart rate did (F(2, 178) = 13.23, $p < .0005$, η_p^2 = .135). Post hoc tests revealed that both the team games ($p < .0005$) and the aerobic exercise intervention ($p < .0005$) led children to higher physical exertion than the control condition with regular physical education contents, whereas the two interventions did not differ from each other ($p = .802$).

To check whether the *cognitive engagement* differed between the three experimental conditions, the three groups were compared with regard to the rated involvement of updating, inhibition and shifting. As intended, the three conditions differed in updating ($F(2, 21) = 13.68$, $p < .0005$, $\eta_p^2 = .566$), inhibition ($F(2, 21) = 11.69$, $p < .0005$, $\eta_p^2 = .527$) and shifting ($F(2, 21) = 12.54$, $p < .0005$, $\eta_p^2 = .544$). Post hoc tests revealed that all three EF subdimensions were represented significantly more strongly in the team games than in the aerobic exercise (*p*s < .0005) and control condition (*p*s < .021), whereas the aerobic exercise and control condition did not differ from each other (*p*s > .257). Taken together, the results show a successful manipulation of the three experimental conditions.

Main analyses

To test the main hypotheses of the study, the three groups were compared regarding their change in the different EF subdimensions between pre- and post-test. Parameter estimates and statistics are presented in Table 3. A χ^2 difference test revealed that the full model fitted the data significantly better than the intercepts-only model for updating, χ^2 (1, $N = 181$) = 1929.84 – 1925.93 = 3.91, $p < .05$, and inhibition, χ^2 (1, $N = 181$) = 4200.36 – 4195.86 = 4.50, $p < .05$, but not for shifting, χ^2 (1, $N = 181$) = 4879.85 – 4877.06 = 2.79, $p > .05$. The change in updating ($F(2, 178) = .81$, $p = .470$) and inhibition ($F(2, 178) = .06$, $p = .947$) did not differ significantly between the three groups. However, the change in shifting differed significantly between the groups ($F(2, 178) = 4.93$, $p = .027$), with post hoc tests revealing a stronger improvement in shifting performance in the team games condition than in the aerobic exercise ($t(178) = 2.33$, $p = .039$) and control condition ($t(178) = 2.95$, $p = .012$). The aerobic exercise and control condition did not differ from each other ($t(178) = -.62$, $p = .544$). The results are depicted in Figure 1.

For aerobic fitness, too, the full model was significantly better than the intercepts-only model, χ^2 (1, N = 181) = 2178.15 – 2053.74 = 124.41, $p < .05$. The pre-post changes in aerobic fitness did differ significantly between the three groups (F(2, 178) = 7.57, p = .001), with post hoc tests revealing both the team games (t(178) = 3.69, $p < .0005$; 4.69 % increase in estimated VO_2max) and the aerobic exercise intervention (t(178) = 3.02, p = .003; 3.79 % increase in estimated VO_2max) to have a greater impact on children's aerobic fitness than the control condition (-.14 % increase in estimated VO_2max). The two interventions did not differ from each other (t(178) = .06, p = .950).

To reveal potential moderating variables for the effects of the interventions on children's EFs, all background variables and the baseline levels of aerobic fitness were inserted as covariates in each of the aforementioned multilevel models. All six full models were significantly better than the respective full model without additional covariate for updating, e.g., SES, χ^2 (1, N = 181) = 1928.09 – 1922.28 = 5.81, $p < .05$ and inhibition, e.g., SES, χ^2 (1, N = 181) = 4158.54 – 4152.20 = 6.34, $p < .05$, but not for shifting, e.g., SES, χ^2 (1, N = 181) = 4831.62 – 4829.00 = 2.62, $p > .05$. Interestingly, none of the six covariates had a significant main or interaction effect on the four dependent variables, Fs < 2.05, n.s., indicating no differential effects of the interventions investigated.

Discussion

The aim of the present study was to investigate the effects of two qualitatively different PA interventions with distinguishable degrees of cognitive engagement on primary school children's EFs. In summary, the results showed (1) that both interventions enhanced children's aerobic fitness more than regular physical education (control condition), but (2) that only the cognitively engaging intervention (team games) fostered pronounced increases in children's shifting performance. The two EF subdimensions updating and inhibition remained unaffected.

The main results showed that cognitive engagement on top of physical exertion affects EFs differently than physical exertion alone. In general, the combination of physical exertion and cognitive engagement in the team games condition seems to have the strongest effect on EFs, since the group exposed to the team games condition improved most in its shifting performance between pre- and post-test compared with the aerobic exercise and the control condition. This result supports the cognitive stimulation hypothesis, whereby interventions including both high amounts of cognitive engagement and physical exertion are thought to have stronger effects on EFs than physically demanding PA with low cognitive engagement. This finding is in line with the rare intervention studies demonstrating the cognitive benefits derived from PA interventions with high amounts of cognitive engagement (Crova et al., 2014; Pesce et al., 2013).

Considering the existing literature, which mainly focuses on endurance-oriented interventions (Davis et al., 2007; Davis et al., 2011; Kamijo et al., 2011), we would have expected both the aerobic exercise and the team games intervention to have a positive effect on children’s EFs, but that the effects of the cognitively enriched intervention would be stronger. Surprisingly, the aerobic exercise did not differ from the control condition with respect to changing switching performance, possibly due to a less pronounced cognitive stimulation compared to the team games condition, as supported by the manipulation check analyses. The present data cannot answer the question whether the cognitive stimulation was mainly induced by the high levels of prospective control and complex eye-hand coordination required by the team games, or by the modification of these team games to specifically challenge EFs through the principle of mental control (Tomporowski, McCullick, & Pesce, 2015). Further studies might, for example, compare an intervention using cognitively enriched team games (like the one in our design) with a traditional team games intervention without any “add-ons”.

The cognitive stimulation hypothesis is, moreover, supported by the fact that the children in both experimental conditions improved their aerobic fitness, but only the ones in the team games condition ameliorated their shifting performance. Thus, the pure improvement of aerobic fitness does not automatically lead to improved cognitive performance, as suggested by the cardiovascular fitness hypothesis (North et al., 1990). Our results are therefore in line with the conclusions of the meta-regression analysis carried out by Etnier et al. (2006), which found that the empirical literature does not support the cardiovascular fitness hypothesis. Perhaps the association between children's aerobic fitness and their cognitive performance (Castelli et al., 2007; Chaddock et al., 2011; Pontifex et al., 2012) could be better explained by the fact that many forms of PA that lead to an improved aerobic performance are themselves cognitively engaging activities (Best, 2010). To answer this speculative question, however, more studies systematically examining the qualitative characteristics of the PA and controlling for cognitive engagement are essential.

The aforementioned improvement of aerobic fitness in both experimental conditions is, in addition to the results of the manipulation check, another indicator that the experimental manipulation of physical exertion with the help of specifically designed contents has succeeded. This improvement, however, seems due not so much to an increase in the physically active time (which did not differ between the three conditions) but rather to a higher intensity during the same time, as represented by a higher mean heart rate in the two intervention groups. The measured mean heart rate in both experimental groups corresponds to moderate to vigorous physical activity (MVPA), whereas that of the control group represents moderate physical activity (Ainsworth et al., 1993). Training studies in children show that MVPA twice a week is necessary to improve prepubertal children's aerobic capacity by 5-6 % in the peak VO_2 (Baquet, van Praagh, & Berthoin, 2003), which is in line with the findings of the present study. Therefore, it is not surprising that the aerobic exercise intervention improved children’s aerobic fitness. But the fact that the same improvement could also be achieved with the help of the team games intervention, which largely complies with the physical education curriculum, is quite a novel result, especially bearing in mind the relatively short time period of six weeks. This calls for high-quality physical education to foster children’s aerobic fitness and thereby their physical health.

The fact that only shifting was positively affected by the team games intervention and updating and inhibition were not, needs to be discussed in detail regarding the selectivity for the effects of chronic PA. To date, no effects on *updating* have been documented in child and adolescent samples, but updating has hardly ever been examined in this age group (Barenberg et al., 2011). Thus, the inclusion of all three EF subdimensions is an added value of the present study, indicating that in the age group studied not all EF subdimensions may be equally prone to changes through chronic PA. As in previous studies (e.g., Drollette et al., 2012; Jäger et al., 2014), updating was measured via the accuracy score. However, a random effects meta-analysis for the effects of acute exercise on working memory has shown that the effect sizes for reaction time and accuracy differ significantly, with a beneficial effect size for reaction time and a detrimental one for accuracy (McMorris, Sproule, Turner, & Hale, 2011). Their explanation of increased catecholamine concentrations in the brain due to acute exercise does not apply for the effects of chronic PA interventions, but possibly the discovered difference in effect sizes indicates a different sensitivity of reaction time and accuracy, respectively, as two outcome variables of the same task. Future studies could therefore use tasks including both measures to better detect possible effects of chronic PA interventions on children’s updating performance.

In contradiction to the findings by Crova et al. (2014) when testing 9- to 10-year-olds, no intervention effects were found on *inhibition*. From a developmental perspective, this lack of effect might be because inhibition is the first EF subdimension to be fully developed in children (Davidson, Amso, Cruess Anderson, & Diamond, 2006) and might therefore be less easily affected than other, not yet fully developed EF subdimensions, such as shifting (Diamond, 2013). Considering studies with adult samples, however, this explanation seems unlikely, since positive effects of PA on inhibition have consistently been found (e.g., Kamijo et al., 2011; Kramer, Erickson, & Colcombe, 2006) and inhibition actually seems to be the dimension which can be affected most easily (Barenberg et al., 2011). A methodological explanation of the disparate findings may lie in the duration of the interventions. Since the cognitively enriched intervention conducted by Crova et al. (2014) lasted six month and the intervention in the present study only lasted six weeks, it could be that producing significant improvements in inhibition, as a more stable EF subdimension in the age group investigated, takes time. However, based on the existing empirical evidence, including the present study, no firm conclusions can yet be drawn regarding different effects due to different intervention durations.

The selective effect on *shifting* exerted by the team games intervention is in line with the study by Pesce et al. (2013) demonstrating the impact of a six-month, cognitively enriched physical education intervention on shifting performance in children aged 5-10 years. Comparing the two studies concerning temporal extension, it should be noted that in the present study positive effects were already achieved after a short period of six weeks. This demonstrates on the one hand the effectiveness of the contents applied in the intervention, but on the other hand, that shifting (compared to inhibition) is an EF subdimension which is prone to positive changes through PA also in later stages of child development. According to recent studies, younger children and older adults tend to exercise their EFs by responding to environmental demands (reactively), while older children and young adults do this more through planful and anticipatory tasks (proactively) (Diamond, 2013; Munakata, Snyder, & Chatham, 2012). Considering that team games make high demands on prospective control and anticipatory abilities, the selected contents appear to offer an ideal match between cognitive development and cognitive demands. Bearing this developmental rationale in mind, it makes absolute sense that studies using the Cognitive Assessment System in 7–11-year-olds (Davis et al., 2007; Davis et al., 2011) have found selective effects only on “higher-order EFs” (Diamond, 2013) such as planning. Although shifting is not considered a higher-order EF, it still seems to be more complex than inhibition and updating and to build upon these two EF subdimensions (Diamond, 2013). Taking into account the finding that larger effects can be expected when higher-order EFs are targeted (McMorris & Hale, 2012), the selective effect on shifting found in the present study is not surprising. So one might speculate that aspects of EFs that are not fully developed (at a certain developmental stage) should be easier to change using PA interventions.

Regarding the selective effect on children’s shifting performance, one might ask how significant this finding is for the educational setting in general and what possible consequences it may have for PA at school. EFs have been shown to predict academic achievement in children and adolescents from ages 5 to 17 (Best, Miller, & Naglieri, 2011). Shifting, as one EF subdimension, also seems to predict academic performance, for example in reading, math and science (e.g., Bull et al., 2008; Latzman, Elkovitch, Young, & Clark, 2010; Yeniad, Malda, Mesman, van IJzendoorn, & Pieper, 2013). This relationship is explained as follows: better shifting abilities can help children to choose and switch between two different problem-solving strategies, to flexibly shift attention to features relevant to the task and to move back and forth between different types of task. These all are requirements needed, for example, when trying to solve a complex math problem.

Although it is still too early to make clear recommendations for the selection of specific contents for physical education or PA interventions at school, the findings of the present study suggest that activities should be chosen that are both physically and cognitively demanding, in order to promote both physical fitness and cognitive performance. That physical fitness is related to academic achievement has been repeatedly demonstrated (Castelli et al., 2007; Chaddock et al., 2011; Pontifex et al., 2012) and a recent study has even discovered EFs to be a mediator in the relationship between physical fitness and academic achievement (van der Niet, Hartman, Smith, & Visscher, 2014). Therefore, the aim to increase children's physical fitness through PA interventions and physical education should undoubtedly be maintained. Nevertheless, it is less obvious how the qualitative characteristics of PA can be used systematically to induce cognitive engagement. Cognitively enriching team games by including certain tasks that address the three EF subdimensions seems to be one viable way, but maybe not the best, since only shifting was enhanced. Building upon all three principles of mental engagement (Tomporowski, McCullick, & Pesce, 2015) could be a promising way of promoting EFs, which may in the end lead to better academic achievement.

Finally, yet importantly, the results of the analyses into potential differential effects need to be briefly discussed, even if none of the analyzed variables were discovered to be moderators. In contrast to previous results, showing that physical aspects like fitness (Chang et al., 2014; Jäger et al., 2015; Hogan et al., 2013) and BMI (Crova et al., 2014; Davis et al., 2007; Davis et al., 2011) as well as cognitive performance (Diamond & Lee, 2011; Drollette et al., 2014; Sibley & Beilock, 2007) moderated the effects of PA interventions on EFs, the results of the present study indicate that the beneficial effect of the team games condition was independent of the specific characteristics of the participants. Thus, positive effects of cognitively engaging PA can be expected in a broad range of typically developing children.

Like any study, the present also has certain limitations, which need to be addressed. First, the assessment of each EF subdimension using only one task was not ideal. However, if we had tried to include all three EF subdimensions, as proposed by Miyake et al. (2000), using two tasks per subdimension, the entire assessment – including aerobic fitness tests, background variables questionnaires, and manipulation checks – would probably have been too long and stressful for both the children and their teachers. In our view, including three EF subdimensions has increased our understanding of the selective nature of different EF subdimensions prone to changes through PA interventions. Second, one could argue that the cognitive engagement in the team games condition was not strong or individualized enough to exercise all three EF subdimensions to the same extent, since there might be an optimal challenge point depending on the joint moderating effect of the complexity of the movement tasks and children’s individual skill level (Pesce et al., 2013). However, the intervention was designed in collaboration with two physical education teachers to ensure that the cognitive and physical requirements were age-appropriate. Furthermore, the manipulation check on cognitive engagement revealed that all three EF subdimensions were represented more highly in the team games than in the aerobic exercise and control condition, whereas the aerobic exercise and control condition did not differ from each other. Therefore, a specific sensitivity to improvement through PA of the three different EF subdimensions seems probable. Third, the manipulation check of the cognitive engagement was performed by two trained observers who were blinded with respect to the experimental conditions, but no standardized instrument was used. However, even though we are not aware of any validated instrument for testing cognitive engagement in PA and therefore the approach of using observational ratings seems justified, further research might develop methods to infer the cognitive engagement inherent in qualitatively different forms of PA. Finally, the randomization was done at a class level rather than on an individual level, as would be required by a true experimental design. Although we tried to deal with this problem using suitable multilevel analyses, a definite causal conclusion between the manipulated variables and their effect is not permissible. To avoid having to resort to quasi-experimental designs in a school setting, one could carry out PA interventions in additional, extra-curricular sports settings. Here, a randomization on the individual level would not pose a problem, but results would not be ecologically valid for physical education.

In conclusion, the present study supports the cognitive stimulation hypothesis in that cognitively engaging PA leads to a stronger improvement in EFs than PA without cognitive engagement, although not all EF subdimensions seem to be affected equally. However, since no detrimental effects of cognitive engagement on top of physical exertion emerged, a combination of these two aspects seems to be a promising approach to affect children's EFs in chronic PA interventions. Thus, when aiming to improve cognitive performance through physical education or PA during a school day, it should be taken into consideration that children not only need enough PA, but also high-quality PA which integrates games and activities requiring cognitive engagement.

References

Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, D. R., Montoye, H. J., Sallis, J. F., & Paffenbarger, R. S. (1993). Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise, 25*(1), 71–80. doi:10.1249/00005768-199301000-00011

American College of Sports Medicine. (2010). *ACSM's guidelines for exercise testing and prescription*. Philadelphia: Lippincott Williams & Wilkins.

Anderson, P. J. (2008). Towards a developmental model of executive function. In V. Anderson, R. Jacobs, & P. J. Anderson (Eds.), *Executive functions and the frontal lobes* (pp. 3–21). New York: Psychology Press.

Armstrong, N. (2013). Aerobic fitness and physical activity in children. *Pediatric Exercise Science, 25*(4), 548-560.

Armstrong, N., & Welsman, J. (2001). Peak oxygen uptake in relation to growth and maturation in 11- to 17-year-old humans. *European Journal of Applied Physiology, 85*(6), 546-551. doi:10.1007/s004210100485

Auer, M., Gruber, G., Mayringer, H., & Wimmer, H. (2005). *Salzburger Lese-Screening für die Klassenstufen 5-8* [Salzburg Reading Screening for Grades 5–8]. Bern: Huber.

Baquet, G., van Praagh, E., & Berthoin, S. (2003). Endurance training and aerobic fitness in young people. *Sports Medicine, 33*(15), 1127–1143. doi:10.2165/00007256-200333150-00004

Barenberg, J., Berse, T., & Dutke, S. (2011). Executive functions in learning processes: Do they benefit from physical activity? *Educational Research Review, 6*(3), 208–222. doi:10.1016/j.edurev.2011.04.002

Bergman Nutley, S., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental Science, 14*(3), 591-601. doi:10.1111/j.1467-7687.2010.01022.x

Best, J. R. (2010). Effects of physical activity on children’s executive function: Contributions of experimental research on aerobic exercise. *Developmental Review, 30,* 331–351. doi:10.1016/j.dr.2010.08.001

Best, J. R. (2012). Exergaming immediately enhances children's executive function. *Developmental Psychology, 48*(5), 1501–1510. doi:10.1037/a0026648

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences, 21,* 327–336. doi:10.1016/j.lindif.2011.01.007

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*(3), 899–911. doi:10.1017/s0954579408000436

Boudreau, B., & Poulin, C. (2009). An examination of the validity of the Family Affluence Scale II (FAS II) in a general adolescent population of Canada. *Social Indicators Research, 94,* 29–42. doi:10.1007/s11205-008-9334-4

Budde, H., Voelcker-Rehage, C., Pietraßyk-Kendziorra, S., Ribeiro, P., & Tidow, G. (2008). Acute coordinative exercise improves attentional performance in adolescents. *Neuroscience Letters, 441,* 219–223. doi:10.1016/j.neulet.2008.06.024

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*(3), 205–228. doi:10.1080/87565640801982312

Castelli, D. M., Hillman, C. H., Buck, S. M., & Erwin, H. E. (2007). Physical fitness and academic achievement in third- and fifth-grade students. *Journal of Sport & Exercise Psychology, 29*(2), 239–252.

Chaddock, L., Hillman, C. H., Buck, S. M., & Cohen, N. J. (2011). Aerobic fitness and executive control of relational memory in preadolescent children. *Medicine and Science in Sports and Exercise, 43*(2), 344–349. doi:10.1249/MSS.0b013e3181e9af48

Chang, Y. K., Labban, J. D., Gapin, J. I., & Etnier, J. L. (2012). The effects of acute exercise on cognitive performance: A meta-analysis. *Brain Research, 1453*, 87–101. doi:10.1016/j.brainres.2012.02.068

Chang, Y.-K., Chi, L., Etnier, J. L., Wang, C.-C., Chu, C.-H., & Zhou, C. (2014). Effect of acute aerobic exercise on cognitive performance: Role of cardiovascular fitness. *Psychology of Sport and Exercise, 15*(5), 464–470. doi:10.1016/j.psychsport.2014.04.007

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology, 45*(3), 782–797. doi:10.1037/a0015409

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290. doi:10.1037/1040-3590.6.4.284

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14*(2), 125–130. doi:10.1111/1467-9280.t01-1-01430

Crocker, P. R. E., Bailey, D. A., Faulkner, R. A., Kowalski, K. C., & McGrath, R. (1997). Measuring general levels of physical activity: Preliminary evidence for the Physical Activity Questionnaire for Older Children. *Medicine and Science in Sports and Exercise, 29*(10), 1344–1349. doi:10.1097/00005768-199710000-00011

Crova, C., Struzzolino, I., Marchetti, R., Masci, I., Vannozzi, G., Forte, R., & Pesce, C. (2014). Cognitively challenging physical activity benefits executive function in overweight children. *Journal of Sports Sciences, 32*(3), 201–211. doi:10.1080/02640414.2013.828849

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23–45. doi:10.1016/S0272-7358(97)00043-3

Davidson, M. C., Amso, D., Cruess Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. doi:10.1016/j.neuropsychologia.2006.02.006

Davis, C. L., Tomporowski, P. D., Boyle, C. A., Waller, J. L., Miller, P. H., Naglieri, J. A., & Gregoski, M. (2007). Effects of aerobic exercise on overweight children's cognitive functioning: A randomized controlled trial. *Research Quarterly for Exercise and Sport, 78*(5), 510–519. doi:10.1080/02701367.2007.10599450

Davis, C. L., Tomporowski, P. D., McDowell, J. E., Austin, B. P., Miller, P. H., Yanasak, N. E., . . . Naglieri, J. A. (2011). Exercise improves executive function and achievement and alters brain activation in overweight children: A randomized, controlled trial. *Health Psychology, 30*(1), 91–98. doi:10.1037/a0021766

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology, 64,* 135–168. doi:10.1146/annurev-psych-113011-143750

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959–964. doi:10.1126/science.1204529

Drollette, E. S., Scudder, M. R., Raine, L. B., Moore, R. D., Saliba, B. J., Pontifex, M. B., & Hillman, C. H. (2014). Acute exercise facilitates brain function and cognition in children who need it most: An ERP study of individual differences in inhibitory control capacity. *Developmental Cognitive Neuroscience, 7,* 53–64. doi:10.1016/j.dcn.2013.11.001

Drollette, E. S., Shishido, T., Pontifex, M. B., & Hillman, C. H. (2012). Maintenance of cognitive control during and after walking in preadolescent children. *Medicine and Science in Sports and Exercise, 44*(10), 2017–2024. doi:10.1249/MSS.0b013e318258bcd5

Etnier, J. L., Nowell, P. M., Landers, D. M., & Sibley, B. A. (2006). A meta-regression to examine the relationship between aerobic fitness and cognitive performance. *Brain Research, 52*, 119–130. doi:10.1016/j.brainresrev.2006.01.002

Federal Office of Sport (1997). *Lehrmittel Sporterziehung* [Teaching material for physical education]. Bern: EDMZ.

Field, A. E., Laird, N., Steinberg, E., Fallon, E., Semega-Janneh, M., & Yanovski, J. A. (2003). Which metric of relative weight best captures body fatness in children? *Obesity Research, 11*(11), 1345-1352. doi:10.1038/oby.2003.182

Gallotta, M. C., Guidetti, L., Franciosi, E., Emerenziani, G. P., Bonavolontà, V., & Baldari, C. (2012). Effects of varying type of exertion on children's attention capacity. *Medicine and Science in Sports and Exercise, 44*(3), 550–555. doi:10.1249/MSS.0b013e3182305552

Götz, L., Lingel, K., & Schneider, W. (2013). *DEMAT 5+ – Deutscher Mathematiktest für fünfte Klassen* [German math test for 5th graders]. Göttingen: Hogrefe.

Hillman, C. H., Kamijo, K., & Scudder, M. (2011). A review of chronic and acute physical activity participation on neuroelectric measures of brain health and cognition during childhood. *Preventive Medicine, 52*, 21–28. doi:10.1016/j.ypmed.2011.01.024

Hogan, M., Kiefer, M., Kubesch, S., Collins, P., Kilmartin, L., & Brosnan, M. (2013). The interactive effects of physical fitness and acute aerobic exercise on electrophysiological coherence and cognitive performance in adolescents. *Experimental Brain Research, 229*(1), 85–96. doi:10.1007/s00221-013-3595-0

Jäger, K., Schmidt, M., Conzelmann, A., & Roebers, C. M. (2014). Cognitive and physiological effects of an acute physical activity intervention in elementary school children. *Frontiers in Psychology, 5:1473.* doi:10.3389/fpsyg.2014.01473

Jäger, K., Schmidt, M., Conzelmann, A., & Roebers, C. M. (2015). The effects of qualitatively different acute physical activity interventions in real-world settings on executive functions in preadolescent children. *Mental Health and Physical Activity, 9*, 1–9. doi:10.1016/j.mhpa.2015.05.002

Kamijo, K., Pontifex, M. B., O'Leary, K. C., Scudder, M. R., Wu, C.-T., Castelli, D. M., & Hillman, C. H. (2011). The effects of an afterschool physical activity program on working memory in preadolescent children. *Developmental Science, 14*(5), 1046–1058. doi:10.1111/j.1467-7687.2011.01054.x

Khan, N. A., & Hillman, C. H. (2014). The relation of childhood physical activity and aerobic fitness to brain function and cognition: A review. *Pediatric Exercise Science, 26*(2), 138–146. doi:10.1123/pes.2013-0125

Kramer, A. F., Erickson, K. I., & Colcombe, S. J. (2006). Exercise, cognition, and the aging brain. *Journal of Applied Physiology, 101*, 1237–1242. doi:10.1152/japplphysiol.000500.2006

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *Journal of Clinical and Experimental Neuropsychology, 32*(5), 455–462. doi:10.1080/13803390903164363

Léger, L. A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sports Sciences, 6*(2), 93–101. doi:10.1080/02640418808729800

Liu, N. Y. S., Plowman, S. A., & Looney, M. A. (1992). The reliability and validity of the 20-Meter Shuttle Test in American students 12 to 15 years old. *Research Quarterly for Exercise and Sport, 63*(4), 360–365. doi:10.1080/02701367.1992.10608757

May, P. (2012). *HSP – Hamburger Schreib-Probe 1-10* [Hamburg Writing Test 1–10]. Hamburg: vpm.

McKenzie, T. L., Sallis, J. F., & Nader, P. R. (1992). SOFIT – System for Observing Fitness Instruction Time. *Journal of Teaching in Physical Education, 11*(2), 195–205.

McMorris, T., & Hale, B. J. (2012). Differential effects of differing intensities of acute exercise on speed and accuracy of cognition: A meta-analytical investigation. *Brain and Cognition, 80*(3), 338–351. doi:10.1016/j.bandc.2012.09.001

McMorris, T., Sproule, J., Turner, A., & Hale, B. J. (2011). Acute, intermediate intensity exercise, and speed and accuracy in working memory tasks: A meta-analytical comparison of effects. *Physiology & Behavior, 102*(3-4), 421–428. doi:10.1016/j.physbeh.2010.12.007

McVeigh, S. K., Payne, A. C., & Scott, S. (1995). The reliability and validity of the 20-meter shuttle test as a predictor of peak oxygen uptake in Edinburgh school children, age 13 to 14 years. *Pediatric Exercise Science, 7*(1), 69–79.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. doi:10.1006/cogp.1999.0734

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science, 21*(2), 71–77. doi:10.1177/0963721412436807

National Association for Sport and Physical Education. (2004). *Moving into the future: National standards for physical education* (2nd ed.). Reston, VA: National Association for Sport and Physical Education.

North, T. C., McCullagh, P., & Tran, Z. V. (1990). Effect of exercise on depression. *Exercise and Sport Sciences Reviews*, *18,* 379-415. doi:10.1249/00003677-199001000-00016

Penedo, F. J., & Dahn, J. R. (2005). Exercise and well-being: A review of mental and physical health benefits associated with physical activity. *Current Opinion in Psychiatry, 18*(2), 189–193. doi:10.1097/00001504-200503000-00013

Pesce, C. (2012). Shifting the focus from quantitative to qualitative exercise characteristics in exercise and cognition research. *Journal of Sport & Exercise Psychology, 34*(6), 766–786.

Pesce, C., Crova, C., Cereatti, L., Casella, R., & Bellucci, M. (2009). Physical activity and mental performance in preadolescents: Effects of acute exercise on free-recall memory. *Mental Health and Physical Activity, 2,* 16–22. doi:10.1016/j.mhpa.2009.02.001

Pesce, C., Crova, C., Marchetti, R., Struzzolino, I., Masci, I., Vannozzi, G., & Forte, R. (2013). Searching for cognitively optimal challenge point in physical activity for children with typical and atypical motor development. *Mental Health and Physical Activity, 6,* 172–180. doi:10.1016/j.mhpa.2013.07.001

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence, 17*(2), 117–133. doi:10.1007/bf01537962

Pontifex, M. B., Scudder, M. R., Drollette, E. S., & Hillman, C. H. (2012). Fit and vigilant: The relationship between poorer aerobic fitness and failures in sustained attention during preadolescence. *Neuropsychology, 26*(4), 407–413. doi:10.1037/a0028795

Reis, H. T., & Gable, S. L. (2000). Event-sampling and other methods for studying everyday experience. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 190-222). New York, NY: Cambridge University Press.

Roebers, C. M., & Kauer, M. (2009). Motor and cognitive control in a normative sample of 7-year-olds. *Developmental Science, 12*(1), 175–181. doi:10.1111/j.1467-7687.2008.00755.x

Roebers, C. M., Roethlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., & Jäger, K. (2014). The relation between cognitive and motor performance and their relevance for children's transition to school: A latent variable approach. *Human Movement Science, 33*, 284–297. doi:10.1016/j.humov.2013.08.011

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*(2), 573–594. doi:10.1207/s15326942dn2802_2

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America, 102*(41), 14931–14936. doi:10.1073/pnas.0506897102

Sibley, B. A., & Beilock, S. L. (2007). Exercise and working memory: An individual differences investigation. *Journal of Sport & Exercise Psychology, 29*(6), 783–791.

Sibley, B. A., & Etnier, J. L. (2003). The relationship between physical activity and cognition in children: A meta-analysis. *Pediatric Exercise Science, 15*(3), 243–256.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Boston: Pearson.

Thorell, L. B., Lindqvist, S., Nutley, S. B., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*(1), 106-113. doi:10.1111/j.1467-7687.2008.00745.x

Tomporowski, P. D., Davis, C. L., Miller, P. H., & Naglieri, J. A. (2008). Exercise and children’s intelligence, cognition, and academic achievement. *Educational Psychology Review, 20,* 111–131. doi:10.1007/s10648-007-9057-0

Tomporowski, P. D., Lambourne, K., & Okumura, M. S. (2011). Physical activity interventions and children's mental function: An introduction and overview. *Preventive Medicine, 52,* 3–9. doi:10.1016/j.ypmed.2011.01.028

Tomporowski, P. D., McCullick, B., Pendleton, D. M., & Pesce, C. (2015). Exercise and children's cognition: The role of exercise characteristics and a place for metacognition. *Journal of Sport and Health Science, 4*(1), 47-55. doi:10.1016/j.jshs.2014.09.003

Tomporowski, P. D., McCullick, B., & Pesce, C. (2015). *Enhancing children's cognition with physical activity games.* Champaign, IL: Human Kinetics.

van der Niet, A. G., Hartman, E., Smith, J., & Visscher, C. (2014). Modeling relationships between physical fitness, executive functioning, and academic achievement in primary school children. *Psychology of Sport and Exercise, 15*(4), 319-325. doi:10.1016/j.psychsport.2014.02.010

Verburgh, L., Königs, M., Scherder, E. J., & Oosterlaan, J. (2014). Physical exercise and executive functions in preadolescent children, adolescents and young adults: A meta-analysis. *British Journal of Sports Medicine, 48*(12), 973–979. doi:10.1136/bjsports-2012-091441

Watzlawik, M. (2009). Die Erfassung des Pubertätsstatus anhand der Pubertal Development Scale. *Diagnostica, 55*(1), 55–65. doi:10.1026/0012-1924.55.1.55

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment, 24*(1), 226–239. doi:10.1037/a0025361

Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences, 23*, 1–9. doi:10.1016/j.lindif.2012.10.004

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives, 6*(4), 354–360. doi:10.1111/j.1750-8606.2012.00246.x

Tables and Figure

Table 1

Means (and standard deviations) for the background, the manipulation check and the dependent variables in the three experimental conditions

	Team games	Aerobic exercise	Control condition
Sample characteristics			
Age (years)	11.32 (.56)	11.33 (.61)	11.40 (.62)
Gender distribution (male/female)	26/43	28/29	28/27
ADHD distribution (with/without)	2/67	4/53	2/53
Physical activity level	2.56 (1.81)	2.70 (1.80)	2.95 (1.78)
Pubertal status	4.97 (1.96)	4.69 (1.57)	5.00 (1.53)
Socioeconomic status	6.44 (1.62)	6.35 (1.74)	6.04 (1.64)
BMI (kg · m^{-2})	18.16 (2.77)	17.37 (2.46)	17.56 (2.61)
Academic achievement	.17 (.78)	-.09 (.76)	-.12 (.90)
Manipulation check variables			
Physical exertion			
Mean heart rate (bpm)*	147.85 (16.62)	150.29 (14.41)	132.00 (28.04)
Physically active time (min/lesson)	22.81 (5.01)	18.69 (10.92)	22.00 (8.75)
Cognitive engagement			
Updating*	60.31 (22.84)	20.18 (13.30)	26.06 (11.20)
Inhibition*	51.13 (14.58)	18.44 (11.21)	30.75 (14.87)
Shifting*	59.63 (13.78)	23.75 (18.13)	30.69 (13.21)
Pre-post data			
Pre-post Δ updating (accuracy)[a]	.91 (3.34)	1.12 (3.91)	.33 (3.24)
Pre-post Δ inhibition (RT)[b]	-15.09 (88.46)	-14.98 (85.78)	-13.25 (63.19)
Pre-post Δ shifting (RT)[b],*	-202.63 (217.52)	-125.90 (184.72)	-107.28 (209.61)
Pre-post Δ estimated VO$_2$max (mL · kg^{-1} · min^{-1})*	2.24 (4.59)	1.91 (3.65)	-.07 (6.32)

Note. ADHD = (diagnosed) attention deficit hyperactivity disorder, BMI = body mass index, VO$_2$max = maximal oxygen consumption, RT = reaction time.

[a] Accuracy corresponds to the number of correct responses.

[b] Reaction time is given in milliseconds.

*p < .05

Table 2

Means and standard deviations for accuracy and reaction times in the Flanker task at pre- and post-test for the three different experimental conditions

	Accuracy[a] *M*(*SD*)			Reaction time[b] *M*(*SD*)		
	Team games	Aerobic exercise	Control condition	Team games	Aerobic exercise	Control condition
Pure (congruent)						
Pre-test	19.21(.98)	18.56(1.65)	18.85(1.29)	421(77)	405(71)	428(84)
Post-test	19.56(1.00)	18.67(2.73)	19.80(2.71)	417(75)	411(83)	437(89)
Standard (congruent)						
Pre-test	19.57(.91)	19.29(1.13)	19.60(.78)	480(106)	473(117)	487(94)
Post-test	19.77(.88)	19.15(2.68)	20.00(2.41)	459(91)	464(132)	484(103)
Standard (incongruent)						
Pre-test	19.37(1.46)	19.08(1.27)	19.50(.90)	504(112)	483(117)	508(99)
Post-test	19.56(1.11)	18.87(2.78)	19.85(2.72)	485(102)	473(119)	503(105)
Mixed (non-switch)						
Pre-test	16.32(1.86)	15.89(2.23)	16.16(1.60)	1042(231)	947(260)	979(234)
Post-test	16.38(1.68)	15.58(3.29)	16.61(2.59)	843(197)	837(231)	882(202)
Mixed (switch)						
Pre-test	17.77(1.65)	16.89(2.27)	17.83(1.08)	1060(299)	959(298)	995(248)
Post-test	17.94(1.17)	17.02(3.27)	18.60(2.58)	845(178)	822(269)	892(195)

Note. [a]Accuracy corresponds to the number of correct responses. [b]Reaction times are given in milliseconds.

Table 3

Results of the four multilevel models with experimental condition as the independent variable and updating, inhibition, shifting and estimated VO_2max as dependent variables

Random Effects

Level	**Effect**	**Parameter Estimate**	**Standard Error**	**Wald *Z***	***p***	**95 % Confidence Interval** *Lower*	*Upper*
Updating							
Class	Intercept	.385	.321	1.20	.230	.075	1.970
Inhibition							
Class	Intercept	219.115	167.811	1.30	.192	48.839	983.048
Shifting							
Class	Intercept	641.989	815.829	.78	.431	53.190	7748.558
Estimated VO_2max							
Class	Intercept	8.877	3.848	2.30	.021	3.795	20.763

Fixed Effects

Effect	**Parameter Estimate**	**Standard Error**	**Approx df**	***t* ratio**	***p***	**95 % Confidence Interval** *Lower*	*Upper*
Updating							
Team games	.569	.622	178	.91	.379	-.792	1.929
Aerobic exercise	.779	.635	178	1.23	.242	-.597	2.155
Control condition[a]	0	0	.	.	.	.	.
Inhibition							
Team games	-4.832	14.572	178	-.332	.746	-36.392	26.727
Aerobic exercise	-2.776	14.865	178	-.187	.855	-34.699	29.147
Control condition[a]	0	0	.	.	.	.	.
Shifting							
Team games	-93.093	31.549	178	-2.951	.012	-162.019	-24.167
Aerobic exercise	-20.197	32.465	178	-.622	.544	-90.172	49.777
Control condition[a]	0	0	.	.	.	.	.
Estimated VO_2max							
Team games	2.307	.625	178	3.687	.000	1.076	3.538
Aerobic exercise	1.974	.654	178	3.017	.003	.687	3.261
Control condition[a]	0	0	.	.	.	.	.

Note. [a]In the main analyses, the control condition served as the reference group.

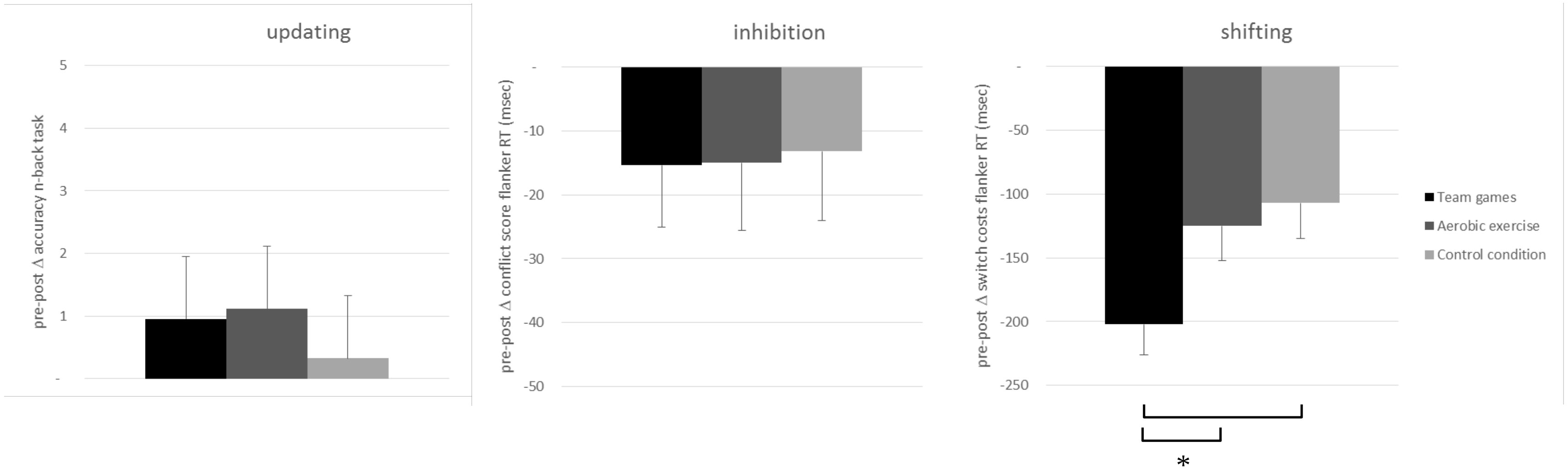

Figure 1. Means and error bars (representing the standard error of the mean) for the change (Δ) in the three EF subdimensions (updating, inhibition, and shifting) in the three experimental conditions between pre- and post-test. *RT* = reaction time.

**p* < .05